

Институт химической физики им Н.Н. Семенова Российской Академии Наук

На правах рукописи

Померанцев Алексей Леонидович

**Методы нелинейного регрессионного
анализа для моделирования кинетики
химических и физических процессов**

01.04.17 – Химическая физика, в том числе физика горения и взрыва

Диссертация

на соискание ученой степени

доктора физико-математических наук

Москва 2003

Оглавление

| | |
|---|-----------|
| ОГЛАВЛЕНИЕ | 2 |
| ВВЕДЕНИЕ | 5 |
| НЕКОТОРЫЕ ТЕОРЕТИЧЕСКИЕ ВОПРОСЫ НЕЛИНЕЙНОГО РЕГРЕССИОННОГО АНАЛИЗА | 16 |
| 1. Основы регрессионного анализа | 17 |
| 1.1. Модель и данные | 17 |
| 1.2. Метод максимума правдоподобия | 21 |
| 1.3. Точность оценивания | 24 |
| 1.4. Проверка гипотез | 27 |
| 1.5. Результаты главы 1 | 33 |
| 2. Последовательное байесовское оценивание | 34 |
| 2.1. Метод максимума правдоподобия с учетом априорной информации | 35 |
| 2.2. Апостериорная информация | 38 |
| 2.3. Общие и частные параметры | 40 |
| 2.4. Обратное последовательное байесовское оценивание | 46 |
| 2.5. Пример применения ПБО | 47 |
| 2.6. Практическое использование метода ПБО | 52 |
| 2.7. Результаты главы 2 | 54 |
| 3. Учет нелинейности регрессии | 56 |
| 3.1. Традиционные методы построения доверительных интервалов | 57 |
| 3.2. Новые методы построения доверительных интервалов | 62 |
| 3.3. Модельный пример построения интервалов | 65 |
| 3.4. Коэффициент нелинейности | 73 |
| 3.5. Результаты главы 3 | 80 |
| ВЫЧИСЛИТЕЛЬНЫЕ АСПЕКТЫ НЕЛИНЕЙНОГО РЕГРЕССИОННОГО АНАЛИЗА | 82 |
| 4. Алгоритмы | 83 |
| 4.1. Минимизация целевой функции | 83 |
| 4.2. Вычисление модели и ее производных | 89 |
| 4.3. Тестирование программ | 100 |

| | |
|---|------------|
| 4.4. Мультиколлинеарность | 107 |
| 4.5. Результаты главы 4 | 114 |
| 5. Описание программы Fitter | 116 |
| 5.1. Основные свойства, возможности, требования и ограничения | 116 |
| 5.2. Данные | 119 |
| 5.3. Модель | 121 |
| 5.4. Параметры | 127 |
| 5.5. Априорная информация | 128 |
| 5.6. Главный Диалог Fitter | 129 |
| 5.7. Регистратор настроек | 131 |
| 5.8. Регистратор данных | 137 |
| 5.9. Регистратор модели | 139 |
| 5.10. Регистратор априорной информации | 141 |
| 5.11. Диалог дополнительных действий | 143 |
| 5.12. Функции Fitter | 145 |
| 5.13. Результаты главы 5 | 165 |
| ПРИЛОЖЕНИЯ НЕЛИНЕЙНОГО РЕГРЕССИОННОГО АНАЛИЗА К ПРАКТИЧЕСКИМ ЗАДАЧАМ | 167 |
| 6. Анализ термограмм полимеров | 170 |
| 6.1. Прогнозирование старения ПВХ методом ТМА | 170 |
| 6.2. Анализ структуры сетки в радиационно модифицированном ПЭ методом ТМА | 177 |
| 6.3. Оценка активности антиоксидантов методом ДСК | 184 |
| 6.4. Результаты главы 6 | 193 |
| 7. Оценивание кинетических параметров по спектральным данным | 194 |
| 7.1. Моделирование «экспериментальных» данные | 195 |
| 7.2. Метод ПБО для спектральных данных | 198 |
| 7.3. Обработка модельных данных | 201 |
| 7.4. Проверка метода ПБО на модельных данных | 206 |
| 7.5. Реальный пример | 207 |
| 7.6. Результаты главы 7 | 210 |
| 8. Прогнозирование старения эластомерных материалов | 212 |
| 8.1. Эксперимент | 214 |
| 8.2. Модель | 217 |

| | |
|--|------------|
| 8.3. Обработка данных УИ | 219 |
| 8.4. Прогнозирование | 222 |
| 8.5. Результаты главы 8 | 224 |
| 9. Моделирование диффузии | 227 |
| 9.1. Нормальная диффузия | 227 |
| 9.2. Аномальная диффузия | 233 |
| 9.3. Релаксационная модель аномальной диффузии | 234 |
| 9.4. Конвекционная модель аномальной диффузии | 240 |
| 9.5. Кинетика цикла «увлажнение–сушка» | 248 |
| 9.6. Другие модели сорбции | 254 |
| 9.7. Результаты главы 9 | 257 |
| 10. Обработка кривых титрования | 259 |
| 10.1. Линейное титрование | 261 |
| 10.2. Потенциометрическое титрование | 265 |
| 10.3. Результаты главы 10 | 270 |
| 11. Хемилюминесцентный метод оценки эффективности ингибиторов | 271 |
| 11.1. Модель | 273 |
| 11.2. Устройство шаблона | 276 |
| 11.3. Программирование шаблона | 280 |
| 11.4. Результаты главы 11 | 287 |
| ЗАКЛЮЧЕНИЕ | 288 |
| ПРИЛОЖЕНИЕ 1. СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ | 292 |
| ПРИЛОЖЕНИЕ 2. СПИСОК ИСПОЛЬЗОВАННЫХ СОКРАЩЕНИЙ | 294 |
| ЛИТЕРАТУРА | 295 |

Введение

Эта работа посвящена применению различных математико-статистических методов и, прежде всего, нелинейного регрессионного анализа (НЛРА) для обработки и интерпретации данных, получаемых в физико-химических экспериментах [1, 2, 3]. Термины «эксперимент» и «экспериментальные данные» являются одним из основных понятий регрессионного анализа, поэтому разъясним, что под ними понимается. Прежде всего, мы будем полагать, что *состояние* исследуемой системы можно исчерпывающе описать некоторым (возможно бесконечным) набором *детерминированных* величин. Часть этих величин известна априори (например, условия эксперимента), а другая часть – неизвестна. Известные величины принято называть *предикторами* (x), а неизвестные – *параметрами* (a). В результате *эксперимента* мы получаем другой (уже конечный) набор величин (y) – экспериментальные данные, которые являются реализацией случайных величин, т.е. выборкой из некоторой гипотетической генеральной совокупности. Случайность результатов измерений – это результат действия многих неизвестных факторов, действующих на исследуемую систему, которые принято называть *ошибкой* или *шумом* (ϵ). Если удалить шум из данных, то оставшиеся детерминированные значения будут являться *сигналом* (f) – полезной информацией, получаемой в эксперименте. Принципиально важно, что различие между сигналом и шумом не является абсолютным и зависит от поставленной задачи и от возможностей прибора. То, что в одной задаче можно рассматривать как шум, в другом случае будет уже полезной информацией – сигналом.

Результаты эксперимента, называемые *откликами*, зависят от набора величин, характеризующих состояние системы – как от предикторов, так и параметров. В общем случае эту зависимость можно представить некоторым оператором

$$y = T(x, a, f, \epsilon),$$

который, собственно, и является математическим аналогом физического *прибора*. Этот оператор может представлять простую, линейную зависимость, но в чаще всего это – сложная, нелинейная функция.

Большинство приборов устроено таким образом, что оператор T можно записать в виде

$$y = f(x, a) + \epsilon,$$

(абсолютная ошибка измерения) или в виде

$$y = f(x, a)(1 + \epsilon)$$

(относительная ошибка измерения). Кроме того, обычно можно предполагать, что шум ϵ является несмещенным, некоррелированным случайным процессом, т.е. ошибки в среднем равны нулю и в разных точках несвязанны друг с другом. Величины этих ошибок, естественно, неизвестны. Такое представление связи между измеряемым откликом и неизвестным сигналом называется *регрессией*, а математические методы анализа этих зависимостей носят название *регрессионных*.

Если модель сигнала $f(\mathbf{x}, \mathbf{a})$ известна, то для решения задачи обработки экспериментальных данных нужно только оценить неизвестные параметры \mathbf{a} , входящие в эту функцию. Для этого применяют различные методы, которые можно объединить одной схемой. Строится некоторая функция $Q(\mathbf{y}, \mathbf{a})$, называемая *целевой*, которая зависит от откликов \mathbf{y} и от неизвестных параметров \mathbf{a} . Затем ищется минимум этой функции по параметрам \mathbf{a} при фиксированных значениях \mathbf{y} . Точка $\hat{\mathbf{a}}$, в которой достигается этот минимум, и является искомой оценкой. Эта величина зависит от экспериментальных значений \mathbf{y} , которые представляют реализацию случайных величин, поэтому и сама оценка является случайной. Выбирая разные целевые функции можно получать разные оценки. При этом предпочтение отдается оценкам, являющимся состоятельными, несмещенными и эффективными. *Состоятельность* означает, что при неограниченном увеличении числа экспериментов значение оценки $\hat{\mathbf{a}}$ сходится к истинному значению \mathbf{a} . Оценка называется *несмещенной*, если ее математическое ожидание равно истинному значению. Наконец оценка будет *эффективной* в классе некоторых оценок, если у нее будет минимальное отклонение от истинного значения.

Выбор вида регрессионной модели, т.е. функция $f(\mathbf{x}, \mathbf{a})$ является центральным моментом при обработке экспериментальных данных. Если эта функция строится на основе базовых представлений о природе процессов, происходящих в исследуемой системе, то она, как правило, является сложной нелинейной зависимостью. Такой подход называется *содержательным* моделированием [2, 111] (*hard modeling*). Другой подход, называемый *формальным* моделированием [122, 123, 127] (*soft modeling*), используется в тех случаях, когда физико-химическое содержание исследуемого процесса либо неизвестно, либо слишком сложно. Тогда строится простейшая линейная зависимость сигнала от неизвестных параметров.

Оба подхода широко применяются на практике. При этом западные исследователи, в основном, предпочитают формальное моделирование. Этот факт подтверждается простым сравнением числа статей [4] об использовании линейного регрессионного анализа (во всех

его разновидностях – PCA, PLS) с немногими публикациями о НЛРА. В то же время у российских ученых издавна существует традиция использовать именно содержательные модели для обработки результатов эксперимента [2, 95, 114]. Это связано с тем, что только нелинейное моделирование дает реальную возможность прогнозировать поведение сложной системы в условиях, которые сильно отличаются от условий, наблюдаемых в эксперименте.

Противопоставление линейного и нелинейного моделирования имеет важное методическое значение. В Табл. 1.1 в схематичном виде представлены некоторые ключевые свойства того и другого метода. Их сравнительный анализ помогает понять, в чем состоят особенности, недостатки и преимущества каждого подхода. Заметим, что существует стойкое убеждение, заключающееся в том, что использование линейного моделирования значительно проще, чем нелинейного. Одна из задач этой работы состоит в том, чтобы опровергнуть такое мнение. Мы покажем, что нелинейный регрессионный анализ данных ничуть не сложнее, а в некотором смысле даже проще, чем линейный.

Табл. 1.1 Свойства линейной и нелинейной регрессионных моделей

| <i>Свойства</i> | <i>Линейное моделирование</i> | <i>Нелинейное моделирование</i> |
|----------------------|---|--------------------------------------|
| Формула | $f = a_1 \varphi_1(x) + \dots + a_p \varphi_p(x)$ | любая $f(x, a)$ |
| Различие | $\partial^2 f / \partial^2 a \equiv 0$ | $\partial^2 f / \partial^2 a \neq 0$ |
| Модель | формальная | содержательная |
| Построение | легкое | трудное |
| Интерпретация | хорошо известна | плохо исследована |
| Назначение | интерполяция | экстраполяция |
| Обработка | обращение матрицы | минимизация |
| Размерность | большая | маленькая |
| Мультиколлинеарность | избыток параметров | нехватка данных |
| Программы | много | мало |

Рассмотрим, в чем проявляется различие между линейным и нелинейным моделированием. Линейная модель представляется уравнением

$$f = a_1 \varphi_1(x) + \dots + a_p \varphi_p(x)$$

в котором a_i – это неизвестные параметры, а x_i – это известные независимые переменные или их функции. Существенно то, что модель линейна именно по параметрам т.е.

$$\partial^2 f / \partial^2 a \equiv 0.$$

При этом она зависимость от предикторов x может быть не линейной. Например, модель $f = a \exp(-20x)$ линейна, поскольку она линейна по параметру a , несмотря на нелинейную зависимость от x .

Формально говоря, линейная модель – это точка в пространстве всех возможных функций от p аргументов. Все остальное пространство занято нелинейными моделями. Поэтому понятно, что основная проблема нелинейного моделирования – это выбор функции для описания эксперимента. Только физико-химические соображения помогают понять, какая модель должна быть использована, и построить содержательное описание. В этой работе почти не затрагиваются вопросы построения содержательных физико-химических моделей. Исключение составляет глава 9, посвященная диффузионным процессам. Большинство моделей, использованных в работе, не является оригинальными. Они либо хорошо известны, либо создавались совместно с коллегами, имеющими более глубокие знания в областях, где эти задачи актуальны.

Работа имеет следующую структуру (см. Оглавление). Она разбита на три части, посвященные соответственно теоретическим, алгоритмическим и практическим проблемам НЛРА. В первой части имеются три главы (1-3), из которых первая посвящена введению в проблему нелинейного оценивания, а вторая и третья содержат оригинальные теоретические разработки в этой области. Вторая часть содержит две главы (4 и 5) – описание алгоритмов и системы Fitter. В третьей части работы, включающей шесть глав (6–11), представлены результаты применения разработанных методов и программ при решении некоторых практических задач.

Методы построения оценок – это центральная проблема статистической теории НЛРА. В главе 1 рассказывается о традиционных, хорошо известных подходах к этой проблеме – методе наименьших квадратов (МНК), методе максимума правдоподобия (ММП). Следующая глава 2 содержит описание оригинального метода *последовательного байесовского оценивания* (ПБО) [30]. Этот метод позволяет решать самые сложные проблемы обработки данных, разбивая одну большую задачу на последовательность маленьких задач, связанных между собой априорной информацией, передаваемой по цепочке. При этом можно решать как задачи с большим числом экспериментальных данных ($n \gg 1$), так и задачи, в которых имеется большое число неизвестных параметров ($p \gg 1$). Метод ПБО – это наш главный инструмент, применяемый при нелинейном статистическом моделировании. Многочисленные примеры таких приложений сосредоточены в третьей части работы. В

главе 2 излагаются основные идеи этого метода, доказывается основная теорема, обосновывающая его применение, приводится простейший пример, и обсуждаются его преимущества и недостатки.

Нахождение оценок неизвестных параметров модели – это только половина работы. Необходимо также и *интерпретировать* полученные результаты, то есть вычислить точность оценок (стандартные отклонения, ковариации), проверить качество подгонки (проверить статистические гипотезы) и построить доверительные интервалы. Классическая теория линейной регрессии дает простые решения [1, 20] для всех этих задач. В случае нелинейной регрессии соответствующая теория разработана еще недостаточно. Здесь мы сталкиваемся с дилеммой: либо использовать линейное приближение, либо применять методы *статистического моделирования* [27, 56]. Первый вариант прост в вычислениях, но не гарантирует точных результатов. Второй вариант, как показала практика, дает очень точные результаты, но требует длительного времени для выполнения. В главе 3 изложен новый подход к учету нелинейности регрессионных оценок – построению доверительных интервалов, оценки степени нелинейности, который отличается от известных тем, что в нем моделируются не исходные данные, а оценки параметров. При этом, что этот метод дает ту же точность, что и традиционное статистическое моделирование, он примерно в 1000 раз быстрее. Важность решения задачи доверительного оценивания объясняется тем, что основное предназначение содержательных, нелинейных моделей – это *прогнозирование* на условия, значительно отличающиеся от условий эксперимента. Хорошо известно, что при такой экстраполяции ошибка предсказания резко возрастает. Поэтому правильное определение границ доверия необходимо для принятия практически важных решений при прогнозе.

В главе 4 рассказывается о том, какие *вычислительные алгоритмы* применяются в нелинейном моделировании. Опираясь на имеющийся опыт, мы можем утверждать, что поиск оценок параметров нелинейной модели не намного сложнее, чем линейной. Это определяется тем, что процедуры *оптимизации* целевой функции детально разработаны. Метод Марквардта [67] является сейчас самым популярным, но есть и более интересные методы, например алгоритм минимизации, основанный на обращении матричной экспоненты, который был предложен Б.В. Павловым и А.Я. Повзнером [74]. Специфика нелинейного моделирования проявляется в двух проблемах: выбор начальных значений параметров и расчет производных модели по параметрам. Проблема выбора стартовой точки не имеет простого решения (и, по-видимому, никогда не будет решена). Здесь можно рассчитывать на

успешный выбор опытного исследователя, который понимает сущность проблемы. Кроме того, можно полагаться и на стабильность алгоритма минимизации, который сходится из широкой окрестности точки минимума целевой функции.

Суть проблемы *вычисления производных* состоит в следующем. Физико-химический эксперимент – это сложная и дорогостоящая процедура, в которой, подчас, имеются принципиальные ограничения, не позволяющие провести измерения в той области, которая была бы желательной с точки зрения математических доводов метода планирования эксперимента. Поэтому ключевым моментом, определяющим успех обработки данных, является точность вычисления модели и особенно ее производных. Производные модели по параметрам играют важную роль при поиске минимума целевой функции – чем точнее они вычисляются, тем точнее определяются оценки. Практический опыт и теоретические выкладки показывают, что для решения задач нелинейного моделирования физико-химических систем, необходимо проводить вычисления с точностью 10-12 десятичных порядков. Разумеется, достичь такой точности, используя разностные методы вычисления производных, невозможно. В этой главе излагаются оригинальные алгоритмы, позволяющие решить одновременно две важные задачи – упростить пользователю задание моделей и добиться высокой точности вычислений

При построении этих процедур мы исходили, прежде всего, из того, что исследователь, применяющий программное обеспечение, построенное на этих алгоритмах, должен иметь максимальные удобства для задания самых сложных моделей. Основное требование – это возможность вводить *явные, неявные*, а также функции, заданные *дифференциальными уравнениями*, в естественной форме, которая должна мало отличаться от общепринятых математических обозначений. Второе условие – это возможность использования промежуточных переменных и подстановок, которые сильно упрощают вид модели. И, наконец, третье требование сводилось к тому, что вычисление производных от функции $f(x, a)$ по параметрам a должно производиться автоматически, без участия человека, даже для самых сложных уравнений. Эти требования удалось удовлетворить, разработав специальные процедуры, интерпретирующие текст модели, задаваемый пользователем. Эти процедуры не только вычисляют (*компилируют*) результат, но и проводят *символьное дифференцирование* модели с последующей компиляцией результата дифференцирования.

Простота тривиального выбора линейной модели иллюзорна, так как эти модели всегда имеют большое число неизвестных параметров. Такая избыточность описания приводит к тому, что все эти параметры невозможно оценить и задача становится *мультиколлинеар-*

ной. Мультиколлинеарность [2, 3, 83] означает вырожденность регрессионной информационной матрицы. Такая проблема встречается и в нелинейной регрессии, но ее интерпретация совершенно другая. Это похоже на классический спор между пессимистом и оптимистом – эта бутылка наполовину пуста или наполовину полна? Линейный анализ представляет оптимистическую точку зрения. В нем всегда предполагается, что модель слишком полна, так что необходимо сократить число параметров [19, 123, 127] любыми способами (PLS, PCA). С другой стороны, в нелинейной модели, как правило, нет лишних параметров, так как все эти параметры продиктованы природой исследуемого процесса [54, 92]. Вот почему при использовании НЛРА мы выбираем пессимистическую точку зрения и предполагаем нехватку экспериментальных данных. Такой подход ведет к специфическим методам борьбы с мультиколлинеарностью в нелинейных моделях (например, байесовский подход), что, тем не менее, не мешает нам использовать и традиционные методы. В конце главы 4 рассматриваются некоторые приемы, позволяющие простыми методами преодолеть проблему мультиколлинеарности.

Как правило, обычного пользователя мало интересуют теоретические и алгоритмические вопросы. Прежде всего, ему нужен простой *инструмент*, позволяющий быстро получить правильный результат. При этом у него должна быть возможность легко менять регрессионную модель «на лету», редактируя ее как текст. Кроме того, нужно предусмотреть и возможность «пакетной» обработки стандартных данных в рутинном процессе, который можно доверить даже неопытному оператору. Все это подводит к задаче создания соответствующего программного обеспечения, удовлетворяющего всем этим требованиям. Разработчики программного обеспечения решили все эти задачи, предоставив пользователям большой выбор программ для линейного моделирования [4], но вот с нелинейным делом обстоит значительно хуже. Конечно, существует несколько программных продуктов [84-90], но они, в большинстве, не отвечают подобным требованиям.

В главе 5 представлен новый инструмент НЛРА [5, 6], который практически реализует все теоретические и алгоритмические разработки, представленные в работе. Он называется *Fitter*, от английского глагола «to fit» – «подгонять, приспособливать». При проектировании этой программы, мы следовали правилу “чем проще - тем лучше”, и не стали создавать собственный интерфейс, а вместо этого воплотили все математические методы как *настройку* для популярной программы Microsoft Excel. В некоторых аспектах система *Fitter* устроена подобно хорошо известному приложению Solver Add-In. Так же, как и в Solver все данные, необходимые для построения регрессии, размещаются на листе стан-

дартной рабочей книги и затем регистрируются посредством диалоговых окон. Внутренний язык системы Microsoft Office – Visual Basic for Applications (VBA) [7] является очень медленным, поэтому все вычислительные процедуры системы Fitter написаны на языке C++ и собраны в отдельной, динамически подключаемой библиотеке (DLL). Таким образом, достигнута быстрота, удовлетворяющая пользователей. Что касается размера экспериментальных данных и числа неизвестных параметров в модели, то Fitter не имеет ограничений на эти величины – все зависит только от возможностей компьютера, который используется для расчетов.

Прототипом и, в какой-то степени, аналогом системы Fitter, является интегрированная компьютерная система *Kinetic Trunk* [8-10, 50]. Эта программа, работающая в среде DOS, была закончена в 1994 году и, до сих пор, используется несколькими научно-исследовательскими и производственными организациями, такими как, например, ВНИИ Эластомерных Материалов и Изделий, Московский Институт Тонкой Химической Технологии, НИИ Шинной Промышленности, Алтайский Университет, Кировский шинный завод, НИИ Кабельной Промышленности, ЦНИИ Точного Машиностроения, НИИ Приборов, Охтинский НПО Пластполимер, Казанский Инженерно Строительный Институт. Опыт эксплуатации программы *Kinetic Trunk* помог при разработке и написании более современной системы – Fitter Add-In, хотя нужно отметить, что многие практические задачи, описанные в третьей части работы, исходно решались еще с использованием системы *Kinetic Trunk*.

В этой части представлены практические приложения методов и алгоритмов, описанные в работе. Подбор этих примеров проводился, в основном, по соображениям методического характера. При этом задачи следуют в *дидактическом* порядке – по возрастанию сложности, с точки зрения использования приемов НЛРА. Каждая из этих глав раскрывает один или несколько методических приемов, применяемых в нелинейном моделировании.

Содержательная, физико-химическая суть этих примеров очень разная. В них рассматриваются и традиционные методы аналитической химии (глава 10), и сложные задачи обработки кинетических данных, полученных с помощью современных инструментальных методов (глава 7), и проблемы прогнозирования эксплуатационной устойчивости полимеров (главы 6 и 8). Все эти задачи наглядно показывают, как проводится нелинейное моделирование в различных ситуациях.

Для всех численных расчетов использовалась система Fitter, поэтому экспериментальные данные и модели почти всегда представляются в форме, которая является стандартной для этой программы. Рабочие книги, в которых содержатся решения этих задач, можно найти в [126].

В главе 6 собраны три практических примера, объединенные общей темой – обработка *термограмм* (ТГА, ТМА и ДСК) полимеров. Эти примеры служат цели введения в проблематику нелинейного моделирования и демонстрации практических приемов работы в системе Fitter. Во всех этих задачах активно используется метод последовательного байесовского оценивания (ПБО).

Глава 7 посвящена актуальной задаче оценки кинетических параметров по *спектральным данным*. Метод последовательного байесовского оценивания, примененный к этой проблеме, позволил получить результаты с высокой точностью. С методической точки зрения пример, рассмотренный в этой главе, важен, прежде всего, потому, что он показывает, как проводится обработка однородных данных в задачах с большим количеством параметров.

Другую сторону метода ПБО раскрывает задача, разобранный в главе 8, где демонстрируется, как обрабатываются разнородные экспериментальные данные, описываемые разными моделями, зависящими от общих параметров. Это сделано на примере практически важной задачи *прогнозирования старения* эластомерных материалов.

Следующая глава 9 посвящена проблемам моделирования нормальной и аномальной *диффузии*. Хорошо известно [129, 130], что модели, описывающие эти процессы, имеют сложную математическую форму и требуют специальных усилий по их программированию. Прямыми выкладками удалось получить точные и достаточно удобные формулы для расчета кинетики сорбции в нефиковских моделях релаксационной и конвекционной диффузии, а также кинетики цикла «увлажнение-сушка». Оригинальная форма этих моделей позволяет использовать их в системе Fitter, что радикально облегчает процесс подбора параметров. Материал, изложенный в этой главе, показывает, что разработанное программное обеспечение может легко оперировать с очень большими и сложными моделями.

Глава 10 стоит несколько в стороне от главной проблематики работы. В ней рассматривается задача обработки данных *титрования*, актуальная для аналитической химии.

В последней главе 11 показано, как, используя разработанное программное обеспечение – систему Fitter, можно создавать стандартные программы-шаблоны, которые автоматиче-

ски строят модели и проводят всю необходимую обработку в скрытом от оператора режиме. Это продемонстрировано на примере оценки *эффективности ингибиторов* хемилюминесцентным методом.

Для правильного моделирования исследуемого процесса или системы важна, прежде всего, постановка задачи. Именно она определяет выбор надлежащего подхода. Предсказание (строго говоря, интерполяция) значений сигнала внутри экспериментальной области – это первая постановка. Прогнозирование (экстраполяция) значений модели на область переменных, которая лежит далеко от области наблюдений – это вторая возможная постановка. Необходимо подчеркнуть существенную разницу этих двух задач. Хорошо известно, что ошибка предсказания (интерполяции) не зависит от вида модели – она, в основном, определяется ошибкой измерений. Напротив, ошибка прогноза (экстраполяции) зависит, прежде всего, от модели и уже во вторую очередь от ошибки измерения. Поэтому, «формальные» модели (в основном, линейные) пригодны для предсказания, и только «содержательные» модели (в основном, нелинейные) пригодны для прогнозирования.

Выбор модели остается главной проблемой для использования НЛРА. Вопрос о том, откуда взять модель, относится, по мнению автора, к категории вечных вопросов, на которые не возможно дать однозначного ответа. Всякая модель отражает только актуальную глубину нашего понимания сути изучаемого явления. Чем больше мы знаем, тем легче составить модель. При этом, нелинейная модель может быть простой, а линейная модель – сложной. Это связано с тем, что содержательная модель использует, как правило, минимальный набор неизвестных параметров, который обусловлен природой системы, тогда как формальная модель вынуждена привлекать большой набор параметров, базирующийся только на структурных особенностях системы.

Пример такой модели – простой по сути, но сложной по форме, приведен в разделе 6.1, где разобран пример прогнозирования старения ПВХ. Эта задача замечательно иллюстрирует наш тезис о том, что теперь, когда разработаны нужные методы и написаны соответствующие программы, нелинейное моделирование ни чуть не сложнее, чем традиционное, линейное. Очень часто модель может быть построена как комбинация содержательной и формальной части. При этом формальная модель используется только для интерполяции данных, так как это делается в примере ТМА–анализа в разделе 6.2.

Проверка адекватности модели является очень острой проблемой в нелинейном моделировании. Традиционные статистические методы, применяемые для линейных задач (про-

верка гипотез, тестовая проверка и перекрестная проверка), здесь подходят плохо. Дело в том, что в линейных моделях в качестве альтернативы всегда выступает тоже линейная модель, но с другим количеством параметров, тогда как для нелинейной модели альтернатив – бесконечно много. Часто встречающаяся постановка о выборе наилучшей модели из конечного списка заявленных альтернатив не имеет ни практического смысла, ни теоретического обоснования. Критерий минимальности невязки, который часто используют в таких задачах, неверен, т.к. цель моделирования не минимальное, а истинное (увы, неизвестное) значение. Легко видеть, что при такой постановке нелинейное моделирование сводится просто к подбору подходящих полиномов, которыми можно описать все с любой точностью.

Нам представляется, что правильным подходом к проблеме адекватности может быть только «внутренняя перепроверка» модели. Поясним эту мысль на примере прогнозирования старения эластомерных материалов, разобранным в главе 8. В этой задаче использовалось несколько моделей, описывающих различные свойства резин. Эти модели содержали несколько общих кинетических параметров – предэкспонент и энергии активации. Именно тот факт, что эти модели удалось построить с использованием единого набора общих параметров, и является подтверждением их адекватности, базирующимся на внутренних, кинетических свойствах изучаемой системы. В тоже время, не следует, разумеется, отказываться от некоторых традиционных статистических приемов, таких, как, например, критерий серий, доказавший свою эффективность в нелинейных моделях.

Эта работа выполнялась в тесном сотрудничестве со многими коллегами. Прежде всего, следует отметить роль Е.В. Быстрицкой (ИХФ), которой принадлежат большинство оригинальных физико-химических моделей, использованных в работе. Большой вклад в разработку алгоритмов и написание программ внесла О.Е. Родионова (ИХФ), которая является соавтором системы Fitter. Неоценимую помощь при разработке алгоритмов минимизации оказал автору Э.Ф. Брин (ИХФ). Экспериментальные данные, использованные в этой работе, были получены А.А. Крючковым (НИИКП), О.В. Старцевым (АГУ), О.В. Платоновой (НИИР), Б.М. Марьяновым (ТГУ), А. Smilde (University of Amsterdam) и N. Overbergh (Raychem Corp.). Особо следует отметить роль О.Н. Карпухина (ИХФ), который оказал огромное влияние на выбор темы и направления этого исследования.

Некоторые теоретические вопросы нелинейного регрессионного анализа

Описать экспериментальные данные с помощью известной модели и предсказать поведение системы в новых условиях, где эксперимент невозможен или затруднен – вот главная цель регрессионного анализа. Существует несколько монографий, посвященных этой области прикладной математики [1, 3, 19, 54]. В этой части работы рассматриваются теоретические аспекты нелинейного регрессионного анализа (НЛРА). Она состоит из трех глав.

Глава 1 предназначена, в основном, для введения необходимых понятий и обозначений. В ней представлен обзор известных подходов к задачам регрессионного анализа.

Следующая глава 2 содержит оригинальный метод последовательного байесовского анализа, разработанный автором в сотрудничестве с Г.А. Максимовой [30]. В ней излагаются основные идеи этого алгоритма, доказывается основная теорема, обосновывающая его применение, приводится простейший пример, и обсуждаются его преимущества и недостатки.

Последняя в этой части, 3 глава, посвящена исследованию наиболее сложных вопросов нелинейной регрессии – проблеме достоверного прогноза и точности оценок. В ней излагается новый метод построения доверительных интервалов для прогноза по регрессии и вводится коэффициент нелинейности. Показывается, как с помощью этой величины принимается решение об использовании того или иного метода доверительного оценивания.

1. Основы регрессионного анализа

Эта глава является введением в проблематику нелинейного регрессионного анализа, поэтому в ней не содержится новых идей. Здесь собраны все основные математические определения величин и методов, которые используются в дальнейшем изложении.

В первом разделе вводятся главные объекты регрессионного анализа: *модель*, экспериментальные и априорные *данные*. Во втором разделе рассматриваются традиционные инструменты, используемые для построения оценок – метод наименьших квадратов и метод максимума правдоподобия. Последние два раздела содержат краткий обзор классических методов анализа *точности* оценок и проверки статистических *гипотез* о данных и модели.

1.1. Модель и данные

Регрессионная модель – это функция

$$\mathbf{f} = f(\mathbf{x}, \mathbf{a}) \quad (1.1)$$

связывающая вектор входных переменных

$$\mathbf{x} = (x_1, \dots, x_m)^t, \quad (1.2)$$

которые называются *предикторами* или *факторами*, с вектором выходных переменных

$$\mathbf{f} = (f_1, \dots, f_k)^t, \quad (1.3)$$

которые называется *сигналом* (или *откликом*). Здесь и далее везде вектор – это столбец чисел, а символ \mathbf{x}^t обозначает операцию транспонирования – т.е. превращение строки в столбец и наоборот. Число предикторов в модели – размерность вектора \mathbf{x} – мы будем обозначать символом m

$$m = \dim(\mathbf{x}), \quad (1.4)$$

Число сигналов в модели – размерность вектора \mathbf{f} – будет обозначаться символом

$$k = \dim(\mathbf{f}). \quad (1.5)$$

В работе рассматривается, в основном, однооткликовая регрессия, т.е. случай $k=1$. Многооткликовая регрессия ($k>1$) будет исследована в разделе 2.3.

В модель регрессии (1.1) обязательно входят неизвестные *параметры*

$$\mathbf{a} = (a_1, \dots, a_p)^t, \quad (1.6)$$

образующие вектор размерности

$$p = \dim(\mathbf{a}). \quad (1.7)$$

Кроме того, модель, конечно, может содержать и известные заранее величины, т.е. константы.

Формальная запись уравнения модели (1.1) в явном, т.е. разрешенном относительно отклика f , уравнении совершенно не означает, что на практике модель будет иметь именно такой простой вид. Кроме того, часто модель можно представить только в виде нескольких выражений (системы), которые в совокупности определяют процедуру вычисления значения отклика. Или же в модель специально вводят новые величины (*промежуточные переменные*), которые упрощают запись уравнений. Приведем некоторые примеры.

Ниже представлена явная функция, которая используется (см. раздел 8.2) для описания зависимости относительного удлинения ELB от времени t и температуры T .

$$ELB(t, T) = a_0 + a_1 e^{-K_1 t} + a_2 e^{-K_2 t}$$

$$K_1 = e^{-k_1 - E_1 X}, \quad K_2 = e^{-k_2 - E_2 X}$$

$$X = \frac{1000}{T + 273} - X_0$$

$$X_0 = 2.77$$

Здесь величина ELB – это сигнал, величины t и T – это предикторы. Вектор неизвестных параметров имеет размерность $b - a_0, a_1, a_2, k_1, k_2, E_1$ и E_2 . В модели участвуют также четыре промежуточных переменных: K_1, K_2, X, X_0 и три константы $1000, 273$ и 2.77 .

Неявная функция используется в примере, представленном в разделе 6.2 для описания данных термомеханического анализа.

$$0 = \frac{S}{P} A \left(Y - \frac{1}{Y^2} \right) - 1$$

$$A = G + G_1 \exp(-k_1 t) + G_2 \exp(-k_2 t)$$

$$S = 0.8659$$

Здесь величина Y – это сигнал, величины P и t – это предикторы. В модели участвуют пять неизвестных параметров G, G_1, G_2, k_1 и k_2 , одна константа S и одна промежуточная величина A .

Пример модели, имеющей вид дифференциального уравнения, можно найти в разделе 6.1. Там приведена модель, которая описывает термогравиметрические кривые, используемые для анализа десорбции пластификатора.

$$\frac{dy}{dt} = -k \left(1 - \frac{I - C_0}{y} \right) \quad y(0) = y_0$$

$$k = F \exp \left(-k_0 - \frac{E}{RT} \right)$$

$$T = T_0 + vt$$

$$R = 1.98717$$

Величина y является откликом, величины t , C_0 , F , v и T_0 – это предикторы. Модель зависит от трех неизвестных параметров y_0 , k_0 и E , одной константы R – универсальной газовой постоянной и выражается через две промежуточные величины – k и T .

Обычно в регрессионном анализе [1–3] предполагается, что вектора \mathbf{x} , \mathbf{f} и \mathbf{a} – это детерминированные величины. Для оценивания регрессионной модели используются *экспериментальные* и *априорные* данные.

Экспериментальные данные получаются в результате экспериментов (измерений, наблюдений), проводимых над исследуемой системой. Набор таких данных включает следующие элементы.

Вектор откликов $\mathbf{y} = (y_1, \dots, y_N)^t$,

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix} \quad (1.8)$$

состоящий из величин сигналов (1.3), измеренных в некоторых условиях. Число измерений – размерность вектора \mathbf{y} – будем обозначать символом N

$$N = \dim(\mathbf{y}) \quad (1.9)$$

Значения независимых переменных (предикторов), характеризующих условия, при которых проводились соответствующие измерения откликов, образуют *матрицу плана* $\mathbf{X} = \{x_{ij}, i=1, \dots, m, j=1, \dots, N\}$, имеющую размерность $(m \times N)$

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{m1} \\ x_{12} & x_{22} & \dots & x_{m2} \\ \dots & \dots & \dots & \dots \\ x_{1N} & x_{2N} & \dots & x_{mN} \end{bmatrix} \quad (1.10)$$

Значения, полученные в ходе эксперимента, отличаются от точных («истинных») значений на величину ошибки. Будем считать, что значения предикторов \mathbf{x} известны точно, либо с ошибкой значительно меньшей, чем измерения откликов. Это обычное предположе-

ние, делаемое в регрессионном анализе [1], которое значительно упрощает оценивание регрессионной модели. При этом будем предполагать, что значения отклика y (1.8) – это случайные переменные, которые отличаются от «истинных» значений сигнала f (1.3) на величины *ошибок* ε .

Традиционно рассматривают ошибки двух типов: абсолютную и относительную. *Абсолютная ошибка* добавляется к «истинным» значениям:

$$y_i = f_i + \varepsilon_i, \quad i = 1, \dots, N \quad (1.11)$$

а *относительная ошибка* умножается на «истинные» значения:

$$y_i = f_i (1 + \varepsilon_i), \quad i = 1, \dots, N \quad (1.12)$$

Относительно вектора случайной ошибки

$$\varepsilon = (\varepsilon_1, \dots, \varepsilon_N)^t \quad (1.13)$$

обычно [1] делают следующие предположения:

Несмещенность. Среднее значение ε равно нулю:

$$E(\varepsilon_i) = 0; \quad i = 1, \dots, N \quad (1.14)$$

Здесь и далее символом $E(\cdot)$ обозначается математическое ожидание случайной величины.

Некоррелированность. Ковариация различных ошибок равна нулю:

$$\text{cov}(\varepsilon_i, \varepsilon_j) = 0; \quad i \neq j \quad (1.15)$$

Гомоскедастичность. Взвешенная дисперсия ошибок постоянна:

$$w_i^2 \text{cov}(\varepsilon_i, \varepsilon_i) = \sigma^2 = \text{const}; \quad i = 1, \dots, N \quad (1.16)$$

Здесь вектор w задает набор *весов*, известных в каждой точке наблюдения

$$w = (w_1, \dots, w_N)^t \quad (1.17)$$

Постоянная величина σ^2 называется *взвешенной дисперсией ошибки*. Обычно она неизвестна и должна оцениваться вместе с параметрами a (1.6). Отметим, что, вообще говоря, переменная σ^2 (1.16) не является дисперсией отклика; последняя может быть определена только для ненулевого веса, как

$$D(y_i) = \frac{\sigma^2}{w_i^2}, \quad w_i \neq 0$$

Здесь и далее символом $D(\cdot)$ обозначается дисперсия случайной величины.

Обычно предполагается, что все весовые коэффициенты равны 1. Если данные не согласуются с предположением о гомоскедастичности (1.16), то нужно установить другие значения весов w_i , обеспечивающие достижение постоянной взвешенной дисперсии в каждой

точке наблюдения. Использование весов является, кроме того, удобным способом регулирования присутствия тех или иных точек в наборе экспериментальных данных. Так для того, чтобы исключить какое-нибудь наблюдение y_i из всего набора, достаточно положить соответствующее значение веса равным нулю, $w_i=0$.

Кроме экспериментальных данных, в регрессионной задаче может присутствовать и *априорная информация* о неизвестных параметрах \mathbf{a} (1.6) и о взвешенной дисперсии σ^2 (1.16). Если такая информация представлена в нормальном *байесовском виде*, то она включает следующие компоненты:

- 1) вектор *априорных значений параметров*:

$$\mathbf{b} = (b_1, \dots, b_p)^t \quad (1.18)$$

- 2) *априорная информационная матрица* размерностью $(p \times p)$:

$$\mathbf{H} = \{h_{\alpha\beta}\}, \quad \alpha, \beta = 1, \dots, p \quad (1.19)$$

- 3) *априорное значение взвешенной дисперсии* ошибки (1.16):

$$s_0^2 \quad (1.20)$$

- 4) *априорное число степеней свободы*:

$$N_0 \quad (1.21)$$

Априорная информация, которая содержит все четыре компонента (1.18) – (1.21), называется *байесовской информацией первого рода*. Иногда априорное значение взвешенной дисперсии (1.20) и ее число степеней свободы (1.21) неизвестны. В этом случае мы имеем так называемую *байесовскую информацию второго рода*.

Подробно роль априорной информации в регрессионном анализе будет разобрана в главе 2. Пока же изложение будет вестись для случая, когда априорная информация отсутствует.

1.2. Метод максимума правдоподобия

Основная задача регрессионного анализа – найти такие значения неизвестных параметров \mathbf{a} (1.6), при которых функция $f(\mathbf{x}, \mathbf{a})$ наилучшим образом приближает значения наблюдаемых откликов y (1.8) во всех точках плана наблюдений \mathbf{X} (1.10). Понятие *наилучшим образом* можно математически формализовать различным способом.

Наиболее популярным является *метод наименьших квадратов* (МНК), в котором мерой близости регрессионной модели к экспериментальным данным является сумма квадратов отклонений

$$S(\mathbf{a}) = \sum_{i=1}^N w_i^2 (y_i - f_i)^2 g_i^2 \quad (1.22)$$

во всех точках плана. Здесь $f_i=f(\mathbf{x}_i, \mathbf{a})$ – это значения модели, а величины w_i – это веса, определенные в формуле (1.16). Для абсолютной ошибки измерения (1.11), $g_i=1$, а для относительной ошибки измерения (1.12), $g_i=1/f_i$.

В методе наименьших квадратов (в нашем случае его следовало бы называть методом *взвешенных* наименьших квадратов) оценки неизвестных параметров

$$\hat{\mathbf{a}} = (\hat{a}_1, \dots, \hat{a}_p)^t, \quad (1.23)$$

ищутся из условия, что они минимизируют сумму $S(\mathbf{a})$, то есть

$$\hat{\mathbf{a}} = \operatorname{argmin} S(\mathbf{a}), \quad (1.24)$$

Помимо МНК, существуют и другие методы оценок параметров регрессии – минимального риска [11], хи-квадрат [12], SIC-метод [15], а также *метод максимума правдоподобия* (ММП), который мы рассмотрим подробно.

Используя векторные обозначения, регрессионную модель с абсолютной ошибкой можно записать в виде

$$\mathbf{y} = \mathbf{f}(\mathbf{X}, \mathbf{a}) + \boldsymbol{\varepsilon} \quad (1.25)$$

В уравнении (1.25) вектор \mathbf{y} и матрица \mathbf{X} известны, поэтому для любых заданных значений параметров \mathbf{a} можно вычислить *вектор остатков*

$$\mathbf{e}(\mathbf{a}) = \mathbf{y} - \mathbf{f}(\mathbf{X}, \mathbf{a}) \quad (1.26)$$

Предположим, нам задана функция совместной плотности распределения ошибок $\boldsymbol{\varepsilon}$

$$p(\boldsymbol{\varepsilon} | \boldsymbol{\psi}), \quad (1.27)$$

которая зависит также от некоторых новых неизвестных параметров $\boldsymbol{\psi}$. Идея ММП состоит в том, что если значение вектора \mathbf{a} будет близко к истинному значению, то значения остатков $\mathbf{e}(\mathbf{a})$ будут близки к ошибкам измерения $\boldsymbol{\varepsilon}$. Если в совместной плотности распределения ошибок (1.27) заменить ошибки остатками, то полученное выражение, являющееся только функцией \mathbf{a} и $\boldsymbol{\psi}$, называется *функцией правдоподобия* выборки [3, 19].

$$L(\mathbf{a}, \boldsymbol{\psi}) = p(\mathbf{e}(\mathbf{a}) | \boldsymbol{\psi}) = p(\mathbf{y} - \mathbf{f}(\mathbf{X}, \mathbf{a}) | \boldsymbol{\psi}), \quad (1.28)$$

Отметим, что величины \mathbf{y} и \mathbf{X} не являются аргументами функции правдоподобия, т.к. они известны.

Предположим, что ошибки распределены нормально, тогда, с учетом предположений (1.14) – (1.16), функция правдоподобия принимает вид

$$L(\mathbf{a}, \sigma^2) = (2\pi)^{-N/2} \sigma^{-N} \exp\left(-\frac{S(\mathbf{a})}{2\sigma^2}\right), \quad (1.29)$$

где величина $\sigma^2 = \psi$ – это взвешенная дисперсия ошибки (1.16), а функция $S(\mathbf{a})$ – это сумма квадратов отклонений, определенная в формуле (1.22). Отметим, что в выражении (1.29) уже учтена возможность относительной ошибки.

Функция правдоподобия $L(\mathbf{a}, \sigma^2)$ зависит от $p+1$ неизвестного параметра

$$\boldsymbol{\varphi} = (\mathbf{a}, \sigma^2).$$

Оценка максимального правдоподобия (ОМП) этого векторного параметра – это значение, при котором функция правдоподобия достигает максимума

$$\hat{\boldsymbol{\varphi}} = \arg \max L(\boldsymbol{\varphi}). \quad (1.30)$$

Известно [3], что ОМП является состоятельной и асимптотически эффективной. Однако, в общем случае, она не будет ни эффективной, ни несмещенной, хотя и будет достаточной. Практика ее применения [16, 17, 18] показала, однако, что ММП дает приемлемые оценки во многих ситуациях. Хотя, в некоторых случаях, имеются методы, дающие лучшие результаты, основным аргументом в пользу использования метода максимума правдоподобия является его общность и простота применения.

Так как логарифм является монотонной возрастающей функцией своего аргумента, то значение $\boldsymbol{\varphi}$, максимизирующее функцию $L(\boldsymbol{\varphi})$ будет минимизировать и выражение $-\ln L(\boldsymbol{\varphi})$. Используя соотношение (1.29) получаем

$$-\ln L(\mathbf{a}, \sigma^2) = \frac{N}{2} \ln 2\pi + \frac{N}{2} \sigma^2 + \frac{S(\mathbf{a})}{2\sigma^2}. \quad (1.31)$$

Отсюда следует, что оценка ММП параметров \mathbf{a} в случае нормального распределения ошибок совпадает с оценкой МНК (см. (1.24)) и, что оценка параметра σ^2 равна

$$s^2 = \frac{S(\hat{\mathbf{a}})}{N}. \quad (1.32)$$

Более сложные случаи ОМП рассмотрены в разделе 1.2.

Анализ «качества» оценок в нелинейной регрессии, т.е. сравнение различных вариантов оценивания весьма затруднительно. Практически можно использовать только асимптотические свойства статистик, проявляющиеся при большом количестве измерений. Тогда известно [19], что оценки ММП будут приближаться по своим свойствам к оценкам, получаемым в линейной регрессии. Это является основным аргументом в пользу сравнения оценок, используемых в нелинейной регрессии с аналогичными оценками применяемым в

линейном случае. Так в частности, легко видеть, что в линейном случае оценка (1.32) является смещенной, и что ее «улучшенный» вариант имеет вид

$$s^2 = \frac{S(\hat{\mathbf{a}})}{N - p}. \quad (1.33)$$

где p – это число неизвестных параметров (1.7).

Метод наименьших квадратов и метод максимума правдоподобия приводят к необходимости решения задачи оптимизации – определения точки минимума некоторой функции $Q(\mathbf{a})$, которая называется *целевой функцией*.

$$\hat{\mathbf{a}} = \operatorname{arg\,min} Q(\mathbf{a}), \quad (1.34)$$

В простейшем случае $Q(\mathbf{a})$ равна $S(\mathbf{a})$ – сумме квадратов отклонений (1.22),

$$Q(\mathbf{a}) = S(\mathbf{a}), \quad (1.35)$$

но возможны и более сложные выражения, которые рассмотрены в разделе 2.1. Решение задачи оптимизации – это самостоятельная, сложная математическая задача и она обсуждается в разделе 4.1.

В дальнейшем изложении будет использоваться термин *обработка данных* (с помощью модели). Он означает процедуру, в ходе которой по экспериментальным данным и модели строится некоторая целевая функция и определяется точка ее минимума – т.е. проводится оценка параметров \mathbf{a} и σ^2 .

1.3. Точность оценивания

Оценки параметров, определяемые с помощью ММП – это случайные величины, реализации которых изменяются вместе с экспериментальными данными. Недостаточно найти точку $\hat{\mathbf{a}}$ минимума целевой функции и назвать ее оценкой параметров \mathbf{a} . Необходимо определить и свойства этой оценки, характеризующие ее точность и надежность. Здесь мы должны уметь отвечать на следующие вопросы. Как могут измениться значения оценок при использовании нового эксперимента? Насколько сильно допустимо изменить значения оценок, чтобы согласие модели с экспериментом оставалось все еще хорошим? Некоторые из этих вопросов рассматриваются в этом разделе, хотя многие проблемы, связанные с проверкой гипотез о регрессии разбираются в разделе 1.4.

Если модель линейна по параметрам, то для решения этих и подобных задач существует хорошо разработанный математический аппарат [1, 20]. В нелинейном регрессионном анализе невозможно предъявить точные формулы. Здесь имеется только две принципиальные возможности. Первая – это использовать линейные стохастические аппроксима-

ции и соответствующие им формулы линейного регрессионного анализа. Вторая – это применять методы стохастического моделирования [21–28] (Монте-Карло, бутстреп и т.п.). В этом разделе будут рассмотрены первый – «квазилинейный» подход к проблеме оценки точности оценивания. Другой – «модельный» подход рассматривается в разделе 3.2.

Известно, что в линейном случае ковариационная матрица оценок $C = \text{cov}(\hat{a}, \hat{a})$ рассчитывается по формуле

$$C = s^2 A^{-1} \quad (1.36)$$

где s^2 оценка взвешенной дисперсии (1.33), а A это информационная матрица, вычисляемая по матрице плана X (1.10) как $A = X'X$. Известно также [3], что если рассматривать нелинейную модель в линейном приближении, то формула (1.36) сохраняется, только матрица A вычисляется иначе. Теперь это – матрица Гессе, т.е. матрица вторых производных по параметрам a от целевой функции в точки минимума

$$A_{\alpha\beta} = \frac{1}{2} \frac{\partial^2 Q(\hat{a})}{\partial a_\alpha \partial a_\beta}, \quad \alpha, \beta = 1, \dots, p \quad (1.37)$$

Информационная матрица интенсивно используется в процедуре поиска (см. раздел 4.1). Известно [3], что проще всего ее вычислять как произведение матриц

$$A = V'V \quad (1.38)$$

где V – это $p \times N$ матрица, чьи элементы – это взвешенные производные от функции модели по параметрам,

$$V_{\alpha i} = w_i \frac{\partial f(x_i, \hat{a})}{\partial a_\alpha}, \quad \alpha = 1, \dots, p; \quad i = 1, \dots, N \quad (1.39)$$

Эти формулы определяют ковариационную матрицу оценок C в простейшем случае, когда отсутствует априорная информация. Уточненные формулы приведены в разделе 2.3.

Как только найдена ковариационная матрица, то легко можно определить и другие связанные с ней характеристики. Среднеквадратичные отклонения (СКО) оценок параметров вычисляются по формуле

$$\text{СКО}(\hat{a}_\alpha) = \sqrt{C_{\alpha\alpha}}, \quad \alpha = 1, \dots, p \quad (1.40)$$

а матрица корреляций по формуле

$$\text{cor}(\hat{a}_\alpha, \hat{a}_\beta) = \frac{C_{\alpha\beta}}{\sqrt{C_{\alpha\alpha}} \sqrt{C_{\beta\beta}}}, \quad \alpha, \beta = 1, \dots, p \quad (1.41)$$

При анализе оценок используется также и F -матрица – матрица, обратная ковариационной матрице C , то есть

$$F = \frac{1}{s^2} A \quad (1.42)$$

Эта матрица подобна байесовской информационной матрице H (1.19). Ее применение исследуется в разделе 2.2.

Распределение оценок параметров в «квазилинейном» приближении естественно считать нормальным с параметрами a и C

$$\hat{a} \sim N(a, C). \quad (1.43)$$

Помимо характеристик качества оценок параметров модели, интересны также и свойства оценки взвешенной дисперсии ошибки (1.33). Ее качество определяется, прежде всего, величиной *числа степеней свободы по регрессии* N_f . Если байесовская информация отсутствует, то N_f рассчитывается по формуле

$$N_f = N_w - p \quad (1.44)$$

где величина N_w – это число измерений, имеющих не нулевые веса w (1.17), т.е.

$$N_w = \sum_{i=1}^N l_i, \quad \text{где } l_i = \begin{cases} 1, & w_i \neq 0 \\ 0, & w_i = 0 \end{cases} \quad (1.45)$$

В «квазилинейном» приближении оценка взвешенной дисперсии распределена по хи-квадрат с N_f степенями свободы

$$s^2 \sim \frac{\sigma^2}{N_f} \chi^2(N_f). \quad (1.46)$$

Все эти значения вычисляются по экспериментальным данным y, X и, следовательно, они сами являются случайными величинами, подверженными выборочной изменчивости. Эти колебания могут быть весьма значительными даже тогда, когда достигнуто удовлетворительное согласие модели с экспериментом. Поэтому, найденные характеристики качества оценок следует рассматривать как некоторую грубую аппроксимацию, правильно отражающую только порядок величин. Кроме того, вся конструкция квазилинейного приближения рухнет, если модель неадекватно описывает исходные данные. Это обстоятельство, впрочем, не должно смущать, т.к. в этом случае необходимо улучшать либо модель, либо исходные данные. Проблема применимости приближенных формул, справедливых для линейных моделей, к задачам нелинейного регрессионного анализа подробно рассмотрена в разделе 3.4.

1.4. Проверка гипотез

Важным инструментом, позволяющим проверить адекватности модели, является проверка *статистических гипотез*. Статистическая гипотеза – это некоторое предположение о распределении выборки экспериментальных данных, в частности о модели, которой описываются эти данные. Решение о справедливости этого предположения должно основываться лишь на имеющихся данных. При этом нужно понимать, что процедура проверки не доказывает истинность или ложность гипотезы. Она только показывает, что имеющиеся экспериментальные данные противоречат или не противоречат ей.

Стандартная процедура (*тест*) состоит в том, что по экспериментальным данным вычисляется некоторое *тестовое значение* ξ (статистика), которое сравнивается с заранее известным *критическим значением* t . Если статистика ξ меньше, чем критическое значение t , то гипотеза принимается, иначе – отвергается. Критическое значение t зависит от выбранного *уровня значимости* α .

Уровень значимости – это мера строгости при проверке гипотез. Чем больше уровень значимости, тем строже проверка. Так, например, гипотеза, принятая для уровня значимости $\alpha = 0.01$, может быть отвергнута для уровня значимости $\alpha = 0.05$.

Величина критического значения определяется распределением статистики ξ и принятым уровнем значимости α как соответствующий *квантиль* этого распределения.

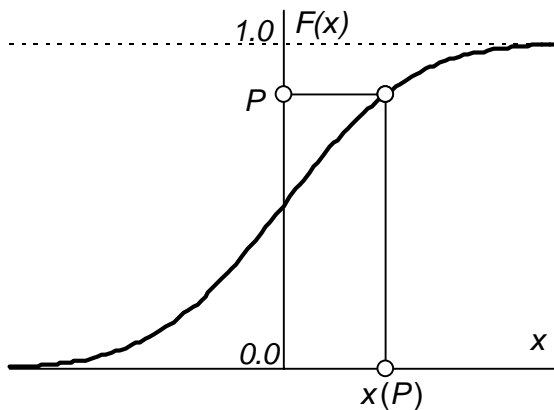


Рис. 1.1 Определение квантиля $x(P)$ для распределения $F(x)$

В статистической литературе имеется некоторая несогласованность в определении понятия квантиля распределения. Используются правосторонние, левосторонние и двусторонние квантили [13, 14, 59]. Поэтому, чтобы избежать путаницы, дадим единое определение квантиля, которое будет использоваться далее повсеместно. График, иллюстрирующий это определение, показан на Рис. 1.1.

Пусть случайная величина ξ имеет (кумулятивную) функцию распределения $F(x)$, т.е.

$$F(x) = \text{Prob} \{ \xi < x \} . \tag{1.47}$$

Квантилем этого распределения будем называть функцию $x(P)$, обратную к (1.47) , т.е.

$$F(x(P)) = P, \quad (1.48)$$

где $0 \leq P \leq 1$

В работе используются квантили следующих распределений, которые обозначаются

стандартное нормальное

$$z(P), \quad (1.49)$$

хи-квадрат с t степенями свободы

$$\chi_t^2(P), \quad (1.50)$$

Стьюдента с t степенями свободы

$$t_m(P), \quad (1.51)$$

F-распределения с t и n степенями свободы

$$F_{m,n}(P). \quad (1.52)$$

Рассмотрим пять различных тестов, которые, по-видимому, наиболее употребительны для проверки адекватности нелинейной регрессии.

Критерий серий

Этот тест используется для проверки гипотезы о некоррелированности ошибок (1.15).

Рассмотрим вектор остатков (1.26), вычисленных для найденных оценок параметров

$$r_i = e_i(\hat{\mathbf{a}}) = f_i(\hat{\mathbf{a}}) - y_i; \quad i = 1, \dots, N,$$

Пусть

P_z – это число неотрицательных остатков, т.е. таких, что $r_i \geq 0$

N_g – это число отрицательных остатков, т.е. таких, что $r_i < 0$ (1.53)

S_e – это число серий остатков, имеющих одинаковый знак.

Тогда *тестовое значение* ξ вычисляется по формуле

$$\xi = |S_e - M - 1| V$$

где $M = \frac{2P_z N_g}{P_z + N_g}, \quad V = \sqrt{\frac{P_z + N_g - 1}{M(M - 1)}}.$

Если данных много (т.е. значений N велико) то *критическое значение* t вычисляется через квантиль нормального распределения (1.49)

$$t = z(1 - \alpha) + 0.5V$$

где α – уровень значимости, а для малых N определяется по таблице [13].

Если статистика ξ больше, чем критическое значение t , то гипотеза отвергается.

Критерий выбросов

Этот тест служит для обнаружения *выбросов*, т.е. таких значений отклика, которые резко отличаются от других данных, например, из-за грубых ошибок. В нем используется широко известный критерий Стьюдента, проверяющий равенство нулю среднего значения в нормальной выборке [14]. Тестовое значение вычисляется по формуле

$$\xi = \max_{1 \leq i \leq N} \left(w_i \frac{|y_i - f_i(\hat{\mathbf{a}})|}{\sqrt{s^2}} \right)$$

а критическое значение – по формуле

$$t = \sqrt{\frac{N_f Z^2}{(N_f - 1) + Z^2}}.$$

Здесь N_f – число степеней свободы (1.44), s^2 – это оценка взвешенной дисперсии ошибки (1.33), а величина Z определяется через квантиль распределения Стьюдента с $N_f - 1$ степенями свободы (1.51) –

$$Z = t_{N_f - 1} \left(1 - \frac{\alpha}{N_f} \right).$$

Если статистика ξ больше, чем критическое значение t , то соответствующий отклик рассматривается как выброс.

Тест альтернативных значений параметров

Этот тест [3, 54] используется для проверки гипотезы о том, что некоторый альтернативный вектор значений параметров

$$\mathbf{a}^* = (a_1^*, \dots, a_p^*)^t$$

не противоречит экспериментальным данным, т.е. что эти значения лежат достаточно «близко» к оцененным оптимальным значениям параметров (1.34)

$$\hat{\mathbf{a}} = (\hat{a}_1, \dots, \hat{a}_p)^t.$$

Тестовое значение вычисляется по формуле

$$\xi = \frac{N_f (Q(\mathbf{a}^*) - Q(\hat{\mathbf{a}}))}{pQ(\hat{\mathbf{a}})} \tag{1.54}$$

где $Q(\mathbf{a})$ – это целевая функция (1.35), N_f – число степеней свободы (1.44), а p – это число неизвестных параметров (1.7).

Критическое значение определяется через квантиль F -распределения (1.52)

$$t = F_{p, N_f} (1 - \alpha) \tag{1.55}$$

Если статистика ξ больше, чем критическое значение t , то гипотеза отвергается.

Все описанные выше критерии могут применяться для любых экспериментальных данных. Два последних теста, которые будут описаны далее, можно проверить только для данных специального типа – имеющих повторные измерения (реплики).

| Data | | | | |
|------|----|------|-------|-------|
| x1 | x2 | Y | F | Var |
| 383 | 48 | 0.63 | 0.653 | 0.005 |
| 383 | 48 | 0.63 | 0.653 | 0.005 |
| 383 | 48 | 0.76 | 0.653 | 0.005 |
| 383 | 72 | 0.76 | 0.844 | 0.001 |
| 383 | 72 | 0.79 | 0.844 | 0.001 |
| 383 | 72 | 0.79 | 0.844 | 0.001 |
| 383 | 72 | 0.84 | 0.844 | 0.001 |
| 383 | 96 | 1.07 | 0.937 | 0.012 |
| 383 | 96 | 0.91 | 0.937 | 0.012 |
| 368 | 24 | 0.13 | 0.106 | 0.000 |
| 368 | 48 | 0.2 | 0.254 | 0.008 |
| 368 | 48 | 0.31 | 0.254 | 0.008 |
| 368 | 48 | 0.13 | 0.254 | 0.008 |

реплика 1, 3 точки

реплика 2, 4 точки

реплика 3, 2 точки

реплика 4, 1 точка

реплика 5, 3 точки

Рис. 1.2 Пример реплик в данных

Репликой называется подвыборка $y^{(j)}$ экспериментальных данных (1.8) имеющих одинаковое значение набора предикторов (1.10) –

$$y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \dots \\ y^{(R)} \end{bmatrix}, \text{ где } y^{(j)} = \begin{bmatrix} y_1^{(j)} \\ y_2^{(j)} \\ \dots \\ y_{n_j}^{(j)} \end{bmatrix} \tag{1.56}$$

Для каждой реплики

$$x_1^{(j)} = x_2^{(j)} = \dots = x_{n_j}^{(j)}$$

На Рис. 1.2 приведен пример данных, имеющих реплики. Видно, что реплика 4 – вырождена, т.к. она состоит только из одной точки. В представленном примере $R=5, n_1=3, n_2=4, n_3=2, n_4=1, n_5=3$.

Если длина реплики больше единицы, т.е. если $n_j \geq 2$, то для этой реплики можно вычислить среднее значение

$$c_j = \frac{\sum_{i=1}^{n_j} w_i^{(j)} y_i^{(j)}}{\sum_{i=1}^{n_j} w_i^{(j)}}$$

и среднеквадратичное отклонение. Последняя вычисляется по разному, в зависимости от типа ошибки:

– для абсолютной ошибки (1.11)

$$d_j^2 = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} [w_i^{(j)} (y_i^{(j)} - c_j)]^2,$$

– для относительной ошибки (1.12)

$$d_j^2 = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \left[w_i^{(j)} \left(\frac{y_i^{(j)}}{c_j} - 1 \right) \right]^2.$$

Если весь вектор откликов разделен на R невырожденных реплик как в (1.56), то можно определить R частных оценок величины σ^2

$$d_1^2, \dots, d_R^2, \tag{1.57}$$

у которых будут соответствующие частные значения степеней свободы

$$v_1 = n_1 - 1, \quad \dots, \quad v_R = n_R - 1 \tag{1.58}$$

Именно значения (1.57) показаны в последнем столбце данных на Рис. 1.2.

Используя эти величины, можно определить объединенную оценку взвешенной дисперсии ошибки по выборке

$$d^2 = \frac{1}{N_s} \sum_{j=1}^R v_j d_j^2, \tag{1.59}$$

где величина

$$N_s = \sum_{j=1}^R v_j = \sum_{j=1}^R m_j - R \tag{1.60}$$

называется числом степеней свободы по выборке.

Теперь, используя введенные понятия, можно описать два оставшихся критерия

Тест адекватности

Этот тест используется для проверки адекватности модели. В нем сравниваются две оценки взвешенной дисперсии ошибки. Первая, s^2 – это оценка по регрессии (1.33), а вторая, d^2 – это оценка по выборке (1.59). Для проверки применяется тест Фишера [54].

Тестовая статистика вычисляется по формуле

$$\xi = \frac{N_f s^2 - N_s d^2}{(R - p) d^2} \quad (1.61)$$

где N_f – число степеней свободы по регрессии (1.44), N_s – число степеней свободы по выборке (1.60), $R > p$ – это число невырожденных реплик (1.56), а p – число неизвестных параметров (1.7).

Критическое значение вычисляется через квантиль F -распределения (1.52)

$$t = F_{R-p, N_s}(1 - \alpha) \quad (1.62)$$

Если статистика ξ больше, чем критическое значение t , то гипотеза отвергается. Проверка гипотезы об адекватности невозможна, если число невырожденных реплик меньше чем число параметров в модели.

Тест дисперсий

В этом тесте проверяется гипотеза о постоянстве взвешенной дисперсии ошибки – условие *гомоскедастичности* (1.16). Для этого частные оценки дисперсий (1.57) сравниваются между собой по критерию Бартлетта [12, 13]. Тестовое значение вычисляется по формуле

$$\xi = \frac{N_s \ln(d^2) - \sum_{j=1}^R v_j \ln(d_j^2)}{I + \frac{\sum_{j=1}^R \frac{1}{v_j} - \frac{1}{N_s}}{3(R-1)}}$$

где $N_s \geq 2$ – число степеней свободы по выборке (1.60), d^2 – это оценка дисперсии по выборке (1.59), (d_1^2, \dots, d_R^2) – частные оценки дисперсии по репликам (1.57), (v_1, \dots, v_R) – частные значения степеней свободы по репликам (1.58) и $R > 2$ – число невырожденных реплик (1.56).

Критическое значение определяется с помощью квантиля распределения хи-квадрат с $R-1$ степенью свободы (1.50)

$$t = \chi_{R-1}^2(1 - \alpha)$$

Если статистика ξ больше, чем критическое значение t , то гипотеза отвергается. Проверка гипотезы о постоянстве дисперсии невозможна, если число невырожденных реплик меньше чем 3.

1.5. Результаты главы 1

В этой главе, являющейся введением в проблематику работы, были введены основные объекты, понятия и методы, используемые в нелинейном регрессионном анализе.

В первом разделе были определены главные объекты регрессионного анализа: *данные, ошибки, веса, модель, параметры и априорная информация*. Во втором разделе рассматривались традиционные подходы, используемые для построения оценок – *метод наименьших квадратов (МНК)* и *метод максимума правдоподобия (ММП)*. В третьем разделе был дан краткий обзор классических способов анализа *точности оценивания*, а четвертом разделе были представлены традиционные методы проверки адекватности – *статистические гипотезы*.

2. Последовательное байесовское оценивание

Идея оценивания параметров регрессии с учетом априорной информации хорошо известна [3], однако она редко применяется на практике. Это связано, прежде всего, с проблемой выбора исходного априорного распределения параметров. Многочисленные дискуссии о правомочности байесовского подхода разделили сообщество статистиков на два клана: байесианцев и анти-байесианцев, каждый из которых приводит многочисленные доводы в защиту своей позиции [29]. Автор принадлежит, скорее, ко второму клану, т.к. полагает, что «волютаристский» выбор априорного распределения может привести к грубым ошибкам. Тем не менее, им был разработан новый метод [30] оценивания параметров регрессии, основанный на использовании байесовского подхода к передаче и использованию информации.

Суть данного подхода [30 – 32] состоит в том, что весь массив исходных данных делится на несколько частей (серий). Поиск параметров ведется последовательно по сериям методом максимума правдоподобия. Результаты оценивания в k -ой серии (параметры, информационная матрица, дисперсия погрешности измерений, число степеней свободы) используются как априорная информация для оценивания в следующей $k+1$ -ой серии. При этом первая серия обрабатывается без априорной информации, поэтому каждое априорное распределение выбирается однозначно и строится последовательность оценок параметров, последняя из которых является окончательной.

Известны, по крайней мере, три типовые ситуации, когда целесообразно применять предлагаемую процедуру. Первая — это тривиальная ситуация, когда размер массива данных слишком велик для его одновременной обработки. Способ разбиения на серии здесь очевиден. Вторая ситуация возникает, когда слишком велик размер вектора неизвестных параметров. Например, если обрабатываются результаты многооткликowego эксперимента, причем регрессионная зависимость для каждого отклика включает в себя как индивидуальные для этого отклика, так и общие для всех откликов параметры. В этом случае размерность информационной матрицы может быть очень велика и процедура ее многократного обращения в ходе поиска [3] оказывается практически невыполнимой. Последовательное оценивание, в котором каждому отклику соответствует своя серия, решает эту проблему. Примеры применения этого метода приведены в главах 6, 7 и 8.

Третий случай связан с возможностью сохранения результатов оценивания параметров регрессионной модели в стандартной компактной форме [33]. При решении практических

задач одни и те же параметры оцениваются в различных моделях. Так, константы скорости химических реакций, оценивающиеся по измерению концентраций реагентов, участвуют и в описании кинетики изменения макроскопических физико-механических свойств [34, 35, 36]. Два этих эксперимента (химический и физический), как правило, разделены во времени, выполняются разными людьми и описываются разными моделями, поэтому информацию об общих параметрах удобно хранить и использовать в стандартной байесовской форме.

В первом разделе этой главы рассматривается метод максимума правдоподобия с учетом априорной информации. Во втором разделе показано, как построить апостериорную информацию и превратить ее в априорную. Метод последовательного байесовского оценивания в общем случае – с учетом общих и частных параметров модели – излагается в третьем разделе. Там же формулируется и доказывается главный результат метода – теорема об эквивалентности оценок в случае линейной регрессии. Четвертый раздел посвящен проблеме обратного байесовского оценивания частных параметров. Простейший модельный пример, иллюстрирующий применение предлагаемого подхода приводится в пятом разделе. Наконец, в последнем, шестом разделе обсуждаются различные практические приложения метода.

2.1. Метод максимума правдоподобия с учетом априорной информации

Рассмотрим задачу оценивания вектора параметров \mathbf{a} входящих в функцию регрессии $f(\mathbf{x}, \mathbf{a})$, в обычной постановке (см. раздел 1.1). В отсутствии априорной информации функция правдоподобия $L(\mathbf{a}, \sigma^2)$ имеет вид (1.29). Обозначим ее теперь $L_0(\mathbf{a}, \sigma^2)$

$$L_0(\mathbf{a}, \sigma^2) = (2\pi)^{-N_w/2} \sigma^{-N_w} \exp\left(-\frac{S(\mathbf{a})}{2\sigma^2}\right), \quad (2.1)$$

где N_w – это число измерений, имеющих не нулевые веса (1.45).

Если имеется априорная информация о значениях параметров \mathbf{a} и взвешенной дисперсии ошибок измерения σ^2 , представленная некоторым распределением $h(\mathbf{a}, \sigma^2)$, то функция правдоподобия меняется на

$$L(\mathbf{a}, \sigma^2) = h(\mathbf{a}, \sigma^2) L_0(\mathbf{a}, \sigma^2) \quad (2.2)$$

Как правило, распределение h неизвестно, а имеющаяся информация ограничивается только данными: (1.18) – (1.21), введенными в разделе 1.1. Там определена априорная информация двух типов: включающая априорное знание о дисперсии σ^2 (*тип 1*) и без нее (*тип 2*). Теперь мы можем конкретизировать практическое использование этих данных.

Будем рассматривать вектор параметров \mathbf{a} как гауссов случайный вектор с математическим ожиданием \mathbf{b} (вектор априорных значений параметров) и матрицей точности $\alpha\mathbf{H}$ (\mathbf{H} – это априорная информационная матрица) [37]

$$\mathbf{a} \sim N(\mathbf{b}, \alpha\mathbf{H}) = \frac{\sqrt{\alpha^p}}{\sqrt{(2\pi)^p}} \sqrt{\det\mathbf{H}} \exp\left[-\frac{\alpha}{2}(\mathbf{a}-\mathbf{b})^t \mathbf{H}(\mathbf{a}-\mathbf{b})\right] \quad (2.3)$$

Величина α зависит от типа априорной информации: $\alpha = \frac{s_0^2}{\sigma^2}$ для априорной информации типа 1 и $\alpha=1$ для типа 2. Здесь s_0^2 – это априорное значение взвешенной дисперсии ошибки.

Для построения распределения величины σ^2 естественно предположить, что случайная величина $N_0 \frac{s_0^2}{\sigma^2}$ распределена по хи-квадрат с N_0 степенями свободы, где N_0 – это априорное число степеней свободы. Поэтому, плотность распределения величины σ^2 равна

$$\sigma^2 \sim \frac{(2N_0 s_0^2)^{\frac{N_0}{2}}}{\Gamma\left(\frac{N_0}{2}\right)} \sigma^{-N_0-2} \exp\left(-N_0 \frac{s_0^2}{2\sigma^2}\right). \quad (2.4)$$

Комбинируя распределения для величин \mathbf{a} и σ^2 , получаем, что для информации по типу 1 априорное распределение $h(\mathbf{a}, \sigma^2)$ имеет вид

$$h(\mathbf{a}, \sigma^2) = \gamma_1 \sigma^{-N_0-2} \exp\left[-\frac{s_0^2}{2\sigma^2}(R(\mathbf{a}) + N_0)\right] \quad (2.5)$$

а для информации по типу 2 вид –

$$h(\mathbf{a}) = \gamma_2 \exp\left[-\frac{1}{2}R(\mathbf{a})\right] \quad (2.6)$$

Здесь $R(\mathbf{a})$ – это квадратичная форма

$$R(\mathbf{a}) = (\mathbf{a}-\mathbf{b})^t \mathbf{H}(\mathbf{a}-\mathbf{b}), \quad (2.7)$$

а величины γ_1 и γ_2 в формулах (2.5) и (2.6) – это нормировочные константы, не зависящие от \mathbf{a} и σ^2 . С учетом соотношения (2.2) функция правдоподобия с априорной информацией по типу 1 примет вид

$$L(\mathbf{a}, \sigma^2) = C_1 \sigma^{-N_w - N_0 - 2} \exp\left[-\frac{1}{2\sigma^2}(S(\mathbf{a}) + s_0^2 R(\mathbf{a}) + N_0 s_0^2)\right], \quad (2.8)$$

а с априорной информацией по типу 2 вид

$$L(\mathbf{a}, \sigma^2) = C_2 \sigma^{-N_w} \exp \left[-\frac{1}{2} \left(\frac{S(\mathbf{a})}{\sigma^2} + R(\mathbf{a}) \right) \right], \quad (2.9)$$

Здесь величина $S(\mathbf{a})$ – это взвешенная сумма квадратов отклонений (1.22), а N_w – это число наблюдений, имеющих не нулевые веса (1.45). Величины C_1 и C_2 – это нормировочные константы, имеющие довольно сложный вид, который, однако, не существенен, т.к. эти величины не зависят от значений искомых параметров \mathbf{a} и σ^2 .

Для определения оценок параметров \mathbf{a} и σ^2 необходимо найти точку, где функция правдоподобия имеет максимум. Легко видеть, что задача оценивания параметров \mathbf{a} с учетом априорной информации, так же, как и без ее учета, сводится к поиску минимума некоторой целевой функции $Q(\mathbf{a})$ (см. формулу (1.34)). Продифференцировав функции (2.8) и (2.9) по параметрам \mathbf{a} и σ^2 , можно установить, что эта функция определяется следующим образом:

$$Q(\mathbf{a}) = S(\mathbf{a}) + B(\mathbf{a}) \quad (2.10)$$

– для байесовской информации первого типа, и

$$Q(\mathbf{a}) = S(\mathbf{a})B(\mathbf{a}) \quad (2.11)$$

– для байесовской информации второго типа.

Вид байесовского члена $B(\mathbf{a})$ зависит от типа априорной информации. Так, для информации типа 1,

$$B(\mathbf{a}) = s_0^2 [N_0 + R(\mathbf{a})] \quad (2.12)$$

а для информации типа 2,

$$B(\mathbf{a}) = \exp \left[\frac{R(\mathbf{a})}{N_w} \right] \quad (2.13)$$

Из ММП следуют и формулы для оценивания параметра σ^2 . Прежде, чем привести их, определим величину N_f – *число степеней свободы с учетом априорной информации*. Для информации 2-го типа она рассчитывается так же, как и без априорной информации (см. формулу (1.44)), а для информации типа 2 эта величина равна

$$N_f = N_w + N_0. \quad (2.14)$$

Оценка взвешенной дисперсии для априорной информации типа 2 вычисляется так же, как и без всякой информации (см. формулу (1.33))

$$s^2 = \frac{S(\hat{\mathbf{a}})}{N_f}. \quad (2.15)$$

Это совершенно естественно, т.к. априорная информация 2-го типа не содержит никаких новых данных о величине дисперсии σ^2 . Для информации 1-го типа, наоборот, априорные данные о величине ошибки позволяют уточнить ее оценку

$$s^2 = \frac{Q(\hat{a})}{N_f} = \frac{S(\hat{a}) + s_0^2 [R(\hat{a}) + N_0]}{N_f}. \quad (2.16)$$

2.2. Апостериорная информация

Основная идея метода последовательного байесовского оценивания (ПБО) состоит в том, что экспериментальные данные разбиваются на части (серии), которые обрабатываются последовательно серия за серией. При этом результаты оценивания каждой серии учитываются при обработке следующей серии как априорная байесовская информация. Рассмотрим, как устроена процедура ПБО и определим правила, по которым осуществляется переход от одной серии данных к другой. Предположим, что все серии имеют общую ошибку измерения, т.е. что здесь применима априорная информация 1-го типа. Тогда в каждой серии для оценки параметров используется одна и та же целевая функция, определенная формулами (2.10) и (2.12)

$$Q(\mathbf{a}) = S(\mathbf{a}) + s_0^2 [N_0 + (\mathbf{a} - \mathbf{b})^t \mathbf{H}(\mathbf{a} - \mathbf{b})], \quad (2.17)$$

хотя величины $S(\mathbf{a}), \mathbf{b}, \mathbf{H}, s_0^2$ и N_0 , разумеется, отличаются на разных шагах. Разложим функцию $Q(\mathbf{a})$ в ряд в точке ее минимума, ограничившись квадратичным приближением

$$Q(\mathbf{a}) \approx Q(\hat{\mathbf{a}}) + (\mathbf{a} - \hat{\mathbf{a}})^t (\mathbf{V}^t \mathbf{V} + s_0^2 \mathbf{H})(\mathbf{a} - \hat{\mathbf{a}}) \quad (2.18)$$

Здесь матрица \mathbf{V} определена уравнением (1.39), т.е. $S(\mathbf{a}) \approx S(\hat{\mathbf{a}}) + (\mathbf{a} - \hat{\mathbf{a}})^t \mathbf{V}^t \mathbf{V}(\mathbf{a} - \hat{\mathbf{a}})$.

Используя формулу (2.16) для оценки дисперсии ошибки, разложение (2.18) можно представить в виде

$$Q(\mathbf{a}) \approx N_f s^2 + (\mathbf{a} - \hat{\mathbf{a}})^t \mathbf{A}(\mathbf{a} - \hat{\mathbf{a}}) \quad (2.19)$$

где $(p \times p)$ матрица \mathbf{A} –

$$\mathbf{A} = \mathbf{V}^t \mathbf{V} + s_0^2 \mathbf{H} \quad (2.20)$$

– это аналог информационной матрицы (1.38), пересчитанный для априорной информации 1-го типа. Аналог F -матрицы определяется по прежней формуле (1.42), причем оценка дисперсии ошибки вычисляется по формуле (2.16).

Итак, после каждого шага ПБО с априорной информацией типа 1 мы получаем следующий набор величин, который естественно назвать *апостериорной информацией 1-го типа*.

1) вектор апостериорных значений параметров:

$$\hat{\mathbf{a}} = (\hat{a}_1, \dots, \hat{a}_p)^t \quad (2.21)$$

2) апостериорная F - матрица (1.42):

$$\mathbf{F} = \{F_{\alpha\beta}\}, \quad \alpha, \beta = 1, \dots, p \quad (2.22)$$

3) апостериорное значение взвешенной дисперсии ошибки (2.16):

$$\sigma^2 \quad (2.23)$$

4) апостериорное число степеней свободы (2.14):

$$N_f \quad (2.24)$$

Этот набор нужно превратить в соответствующий набор *априорной* информации (1.18) – (1.21), который будет использоваться на следующем шаге процедуры, т.е. при обработке новой серии данных. Такой переход осуществляется по следующим простым правилам:

1) вектор *априорных* значений параметров равен вектору *апостериорных* значений:

$$\mathbf{b} = \hat{\mathbf{a}} \quad (2.25)$$

2) *априорная* информационная матрица равна *апостериорной* F -матрице:

$$\mathbf{H} = \mathbf{F} \quad (2.26)$$

3) *априорное* значение взвешенной дисперсии равно *апостериорному* значению

$$s_0^2 = s^2 \quad (2.27)$$

4) *априорное* число степеней свободы равно *апостериорному* числу

$$N_0 = N_f \quad (2.28)$$

При таком построении априорной информации байесовский член $B(\mathbf{a})$ (2.12) на $k+1$ -ом шагу процедуры будет совпадать с целевой функцией $Q(\mathbf{a})$ для k -го шага с точностью до членов третьего порядка в разложении (2.18). Это означает, что оценки параметров, полученные в последовательной процедуре, будут близки к аналогичным оценкам, которые могли бы быть найдены при совместном оценивании.

Итак, мы определили правило превращения апостериорной информации в априорную и, тем самым, построили рекуррентную процедуру, по которой осуществляется переход от k -ой серии данных к $k+1$ -ой. Осталось только определить как «запустить» эту процедуру, т.е. что делать с первой серией, которой ничего не предшествует. Решение очень простое – ее нужно обрабатывать без априорной информации так, как это описано в разделе 1.2. Теперь мы полностью описали метод последовательного байесовского оценивания, который можно представить следующим алгоритмом:

- 1) данные разбиваются несколько серий;
- 2) первая серия обрабатывается с целевой функцией (1.35);
- 3) строится апостериорная информация по формулам (2.21) – (2.24);
- 4) строится априорная информация по формулам (2.25) – (2.28); (2.29)
- 5) следующая серия обрабатывается с целевой функцией (2.17);
- 6) повторяем шаги 3 – 5 до последней серии;
- 7) оценки, полученные для последней серии, являются оценками ПБО.

Аналогичная последовательная процедура может быть построена и для априорной информации 2-го типа. В этом случае информационная матрица должна рассчитываться по формуле

$$A = \exp\left(\frac{R(\hat{\mathbf{a}})}{N_w}\right) \left(\mathbf{V}'\mathbf{V} + \frac{S(\hat{\mathbf{a}})}{N_w} \mathbf{H} \right), \quad (2.30)$$

а в F -матрице должна использоваться оценка дисперсии (2.15). Все остальные соотношения остаются неизменными, но с учетом того, что апостериорная информация типа 2 не содержит данных (2.27) и (2.28) о взвешенной дисперсии ошибки.

2.3. Общие и частные параметры

В предыдущем разделе метод ПБО был определен для случая, когда все экспериментальные данные описываются одной и той же моделью, содержащей одни и те же неизвестные параметры. В такой постановке этот метод полезен, прежде всего, для обработки больших, но однородных массивов данных, где основная проблема – это слишком большая размерность N – число экспериментов. Однако, наиболее интересно использование ПБО для интерпретации разнородных массивов данных, каждый из которых описывается своей регрессионной моделью, содержащей несколько общих для всех этих моделей параметров. Рассмотрим, как строится и используется априорная информация в этом случае.

Пусть имеется M регрессионных моделей

$$f_j(\mathbf{x}, \mathbf{a}_j), \quad j = 1, \dots, M,$$

каждая из которых содержит p_j неизвестных параметров \mathbf{a}_j ,

$$\dim \mathbf{a}_j = p_j,$$

причем в каждом векторе \mathbf{a}_j первые r параметров a_1, \dots, a_r являются общими для всех моделей, а остальные $p_j - r$ параметров a_{r+1}, \dots, a_{p_j} – это частные, присутствующие только в одной модели. Всего, таким образом, всего имеется

$$p = p_I + \dots + p_M - r(M - I) \quad (2.31)$$

неизвестных параметров \mathbf{a} . Пусть также имеются соответствующие этим моделям, M наборов данных $(\mathbf{x}_j, \mathbf{y}_j)$, каждый размерностью N_j , содержащие ошибки, характеризуемые неизвестными взвешенными дисперсиями σ_j^2 . Полное число всех измерений равно –

$$N = N_I + \dots + N_M. \quad (2.32)$$

Цель дальнейшего изложения – определить, как по результатам обработки данных в j -ой серии строится априорная информация для обработки следующей, $j+1$ –ой серии данных. Для j -ой серии данных (далее мы не будем использовать индекс j , чтобы не усложнять обозначения) апостериорная информационная матрица \mathbf{A} (формула (2.20) или (2.30)) имеет блочный вид

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{00} & \mathbf{A}_{01} \\ \mathbf{A}_{01}^t & \mathbf{A}_{11} \end{bmatrix} \quad (2.33)$$

где блок \mathbf{A}_{00} – это $r \times r$ квадратная матрица, которая соответствует первым (общим) r элементам вектора параметров, блок \mathbf{A}_{11} – это $(p_j - r) \times (p_j - r)$ квадратная матрица, которая соответствует последним (частным) $p_j - r$ параметрам, и блок \mathbf{A}_{01} – это прямоугольная матрица размером $r \times (p_j - r)$. Размерность всей матрицы \mathbf{A} есть $p_j \times p_j$. Из этой матрицы нужно построить новую матрицу $\tilde{\mathbf{A}}$, сохранив в ней информацию только об общих параметрах и исключив данные о частных параметрах. Это делается с помощью следующего простого преобразования

$$\tilde{\mathbf{A}} = \mathbf{A}_{00} - \mathbf{A}_{01} \mathbf{A}_{11}^{-1} \mathbf{A}_{01}^t \quad (2.34)$$

Матрица $\tilde{\mathbf{A}}$ имеет размерность $r \times r$. Априорная матрица \mathbf{H} строится из этой матрицы по следующему правилу

$$\mathbf{H} = \frac{1}{s^2} \begin{bmatrix} \tilde{\mathbf{A}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (2.35)$$

где s^2 – это оценка дисперсии ошибки в j -ой модели. Размерность матрицы \mathbf{H} должна быть уже $p_{j+1} \times p_{j+1}$, где p_{j+1} – это число параметров в следующей, $j+1$ –ой модели f_{j+1} . Для этого матрица $\tilde{\mathbf{A}}$ дополняется нулевыми значениями до надлежащей размерности. Аналогично поступим и с априорными значениями параметров для $j+1$ –ой серии –

$$b_\alpha = \begin{cases} \hat{a}_\alpha & , \quad 0 < \alpha \leq r \\ 0 & , \quad r < \alpha \leq p_{j+1} \end{cases} \quad (2.36)$$

Формулы для оценки взвешенной дисперсии не изменятся, вот только число степеней свободы N_f для случая постоянной дисперсии ошибки (априорная информация 1-го типа) теперь нужно вычислять по новой формуле

$$N_f = N_w - N_0 - p_j + r \quad (2.37)$$

где N_w – число наблюдений с ненулевыми весами, а N_0 – априорное число степеней свободы для j -ой серии.

Рассмотрим случай постоянной взвешенной дисперсии ошибки, когда $\sigma_1^2 = \dots = \sigma_M^2 = \sigma^2$.

Теперь мы можем построить алгоритм ПБО, аналогичный приведенному в (2.29).

- 1) данные разбиваются M серий;
- 2) первая серия обрабатывается с целевой функцией (1.35);
- 3) строится апостериорная информация по формулам (2.21) – (2.24);
- 4) строится априорная информация по формулам (2.33)–(2.37);
- 5) следующая серия обрабатывается с целевой функцией (2.17);
- 6) повторяем шаги 3 – 5 до последней серии;
- 7) оценки, полученные для последней серии оценками метода ПБО.

Сравним предлагаемый метод с традиционным методом максимума правдоподобия (ММП), где используется следующая целевая функция

$$S(\mathbf{a}_1, \dots, \mathbf{a}_M) = S_1(\mathbf{a}_1) + \dots + S_M(\mathbf{a}_M). \quad (2.39)$$

Здесь

$$S_j(\mathbf{a}_j) = \sum_{i=1}^{N_j} w_{ji}^2 (y_{ji} - f_j(\mathbf{x}_{ji}, \mathbf{a}_j))^2 g_{ji}^2, \quad j = 1, \dots, M. \quad (2.40)$$

– это частная сумма квадратов отклонений в j -ой серии.

Оценки ММП – это точка минимума этой целевой функции

$$\hat{\mathbf{a}} = (\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_M) = \operatorname{argmin} S(\mathbf{a}_1, \dots, \mathbf{a}_M), \quad (2.41)$$

и оценка дисперсии ошибки

$$s^2 = \frac{S(\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_M)}{N - p}, \quad (2.42)$$

где величины N и p определены в формулах (2.31) и (2.32).

Докажем утверждение, обосновывающее метод ПБО.

Теорема 1. Пусть функции

$$f_j(\mathbf{x}, \mathbf{a}_j), \quad j = 1, \dots, M$$

линейны по параметрам \mathbf{a}_j , а ошибки распределены нормально с взвешенной дисперсией, которая не зависит от j , т.е. $\sigma_1^2 = \dots = \sigma_M^2 = \sigma^2$ тогда совпадают:

- 1) ПБО и ММП оценки общих параметров;
- 2) Матрицы ковариаций общих параметров;
- 3) ПБО и ММП оценки взвешенной дисперсии;

При $M=1$ это утверждение очевидно. Приведем доказательство для $M = 2$, тогда справедливость утверждения для произвольного натурального M будет следовать по индукции. Обозначим

$$\mathbf{a}_1 = \begin{bmatrix} \mathbf{c}_0 \\ \mathbf{c}_1 \end{bmatrix} \quad \mathbf{a}_2 = \begin{bmatrix} \mathbf{c}_0 \\ \mathbf{c}_2 \end{bmatrix},$$

где \mathbf{c}_0 – вектор общих параметров размерностью r , а \mathbf{c}_1 и \mathbf{c}_2 – это вектора частных параметров с размерностями $q_1=p_1-r$ и $q_2=p_2-r$ соответственно. В силу линейности регрессий

$$\mathbf{f}_1 = \mathbf{X}_{01} \mathbf{c}_0 - \mathbf{X}_1 \mathbf{c}_1, \quad \mathbf{f}_2 = \mathbf{X}_{02} \mathbf{c}_0 - \mathbf{X}_2 \mathbf{c}_2,$$

где матрицы плана (1.10) имеют следующие размерности:

$$\mathbf{X}_{01} - r \times N_1, \quad \mathbf{X}_{02} - r \times N_2, \quad \mathbf{X}_1 - q_1 \times N_1, \quad \mathbf{X}_2 - q_2 \times N_1.$$

Тогда целевую функцию ММП (2.39) можно представить в виде

$$S(\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2) = S_1(\mathbf{c}_0, \mathbf{c}_1) + S_2(\mathbf{c}_0, \mathbf{c}_2) = \mathbf{e}_1^t \mathbf{W}_1 \mathbf{e}_1 + \mathbf{e}_2^t \mathbf{W}_2 \mathbf{e}_2 \quad (2.43)$$

где вектора остатков имеют вид

$$\mathbf{e}_1 = \mathbf{y}_1 - \mathbf{X}_{01} \mathbf{c}_0 - \mathbf{X}_1 \mathbf{c}_1, \quad \mathbf{e}_2 = \mathbf{y}_2 - \mathbf{X}_{02} \mathbf{c}_0 - \mathbf{X}_2 \mathbf{c}_2,$$

а \mathbf{W}_1 и \mathbf{W}_2 – это матрицы весов размерностью $N_1 \times N_1$ и $N_2 \times N_2$ соответственно. Из этого соотношения следует, что ММП оценки являются решением следующей системы нормальных уравнений

$$\begin{bmatrix} \mathbf{A}_{001} + \mathbf{A}_{002} & \mathbf{A}_{01} & \mathbf{A}_{02} \\ \mathbf{A}_{01}^t & \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{A}_{02}^t & \mathbf{0} & \mathbf{A}_{22} \end{bmatrix} \times \begin{bmatrix} \mathbf{c}_0 \\ \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{z}_{01} + \mathbf{z}_{02} \\ \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix}, \quad (2.44)$$

включающей матрицы

$$\begin{aligned} \mathbf{A}_{001} &= \mathbf{X}_{01}^t \mathbf{W}_1 \mathbf{X}_{01}, & \mathbf{A}_{002} &= \mathbf{X}_{02}^t \mathbf{W}_2 \mathbf{X}_{02}, \\ \mathbf{A}_{01} &= \mathbf{X}_{01}^t \mathbf{W}_1 \mathbf{X}_1, & \mathbf{A}_{02} &= \mathbf{X}_{02}^t \mathbf{W}_2 \mathbf{X}_2, \\ \mathbf{A}_{11} &= \mathbf{X}_1^t \mathbf{W}_1 \mathbf{X}_1, & \mathbf{A}_{22} &= \mathbf{X}_2^t \mathbf{W}_2 \mathbf{X}_2 \end{aligned}$$

и вектора

$$\begin{aligned} z_{01} &= X_{01}^t W_1 y_1, & z_{02} &= X_{02}^t W_2 y_2, \\ z_1 &= X_1^t W_1 y_1, & z_2 &= X_2^t W_2 y_2 \end{aligned}$$

Исключив неизвестный параметр c_1 из системы (2.44) мы получим следующие уравнения для определения ММП оценок

$$\begin{bmatrix} A_{001} + A_{002} - A_{01} A_{11}^{-1} A_{01}^t & A_{02} \\ A_{02}^t & A_{22} \end{bmatrix} \times \begin{bmatrix} c_0 \\ c_2 \end{bmatrix} = \begin{bmatrix} z_{01} + z_{02} - A_{01} A_{11}^{-1} z_1 \\ z_2 \end{bmatrix} \quad (2.45)$$

Рассмотрим теперь метод ПБО. На первом шагу минимизируется целевая функция

$$Q_1(c_0, c_1) = S_1(c_0, c_1)$$

и определяются оценки \hat{c}_0, \hat{c}_1 как решения системы нормальных уравнений

$$\begin{bmatrix} A_{001} & A_{01} \\ A_{01}^t & A_{11} \end{bmatrix} \times \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} z_{01} \\ z_1 \end{bmatrix}. \quad (2.46)$$

Исключая параметр c_1 из этой системы получаем, что

$$(A_{001} - A_{01} A_{11}^{-1} A_{01}^t) \hat{c}_0 = z_{01} - A_{01} A_{11}^{-1} z_1 \quad (2.47)$$

На втором шаге ПБО целевая функция имеет вид (2.17), т.е.

$$Q_2(c_0, c_2) = S_2(c_0, c_2) + S_1(\hat{c}_0, \hat{c}_1) + (c_0 - \hat{c}_0)^t H (c_0 - \hat{c}_0),$$

Составляя нормальные уравнения для этой функции, получаем

$$\begin{bmatrix} A_{002} + H & A_{02} \\ A_{02}^t & A_{22} \end{bmatrix} \times \begin{bmatrix} c_0 \\ c_2 \end{bmatrix} = \begin{bmatrix} z_{02} + H \hat{c}_0 \\ z_2 \end{bmatrix}. \quad (2.48)$$

где величина \hat{c}_0 определяется из уравнения (2.47)

Сопоставляя систему нормальных уравнений для ММП (2.45) и для ПБО (2.48), видим, что они совпадают, когда матрица H определяется согласно формуле (2.34), т.е.

$$H = A_{001} - A_{01} A_{11}^{-1} A_{01}^t \quad (2.49)$$

Это доказывает первое и второе утверждение приведенной выше теоремы. Третье утверждение доказывается аналогично прямой подстановкой полученных оценок формулы для вычисления оценок дисперсии: ММП (2.42) и ПБО (2.16).

При доказательстве приведенного выше утверждения предполагалась невырожденность матрицы A_{11} в формуле (2.49), что существенно при вычислении априорной информационной матрицы H . Заметим, что вырожденность матрицы общих параметров A_{001} не мешает вычислению матрицы H . Матрица A_{11} отвечает за оценки частных параметров, которые входят только в одну модель. В то же время, если хотя бы одна из матриц A_{001} или A_{11} вырождена, то вырождена и вся матрица в формуле (2.46) (строгая мультиколлинеарность

[54]) и определить оценки параметров классическим ММП нельзя. В этом случае необходимо еще до оценивания модели провести некоторую регуляризацию задачи, что на практике означает фиксирование и выведение из поиска некоторых параметров. Если это частные параметры, то они так и остаются фиксированными, обеспечивая невырожденность A_{II} . Уточнить их далее не удастся, так как в другие модели они не входят. Если же фиксировались некоторые общие параметры, то естественно освободить их (объявить неизвестными) перед вычислением информационной матрицы H . Их значения могут быть уточнены в ходе последующего оценивания.

Итак, показано, что в линейном случае ПБО дает те же оценки, что и обычный ММП. В случае нелинейной параметризации можно говорить только об асимптотических свойствах этих оценок, сравнивая их с ММП оценками. Если оценки ММП обладают некоторыми асимптотическими свойствами, то они будут также и асимптотически нормальными [19]. Заметим, что априорное распределение (2.3) является нормальным. Из доказанной выше теоремы следует что, оценки, построенные с помощью ПБО, будут асимптотически совпадать с ММП оценками и, следовательно, будут обладать теми же свойствами.

Аналогичное утверждение (с заменой целевой функции на функцию (2.11)) справедливо для серии моделей, имеющих разные дисперсии ошибки. При этом последовательно построенная оценка с учетом априорной информации по типу 2 совпадает с оценкой взвешенного МНК. Отметим, что метод ПБО позволяет избежать нестрогой процедуры выбора весов w_{ji} в выражении (2.40).

Для оценки точности байесовского оценивания также применяются «квазилинейное» приближение, основанное на формуле (1.36), где матрица A определяется по формуле (2.30). Однако, можно показать, что *полные* ковариационные матрицы ММП и ПБО оценок будут, вообще говоря, отличаться даже в линейном случае. Это отличие будет наблюдаться в блоках, ответственных за ковариации частных параметров между собой и ковариации с общими параметрами. Причем, очевидно, что первый блок будет иметь большие, а второй – меньшие значения, нежели в ММП. Это связано с тем, что при преобразовании (2.34) информационной матрицы теряются данные о связях между общими и частными параметрами. Кроме того, и сами оценки частных параметров будут, разумеется, отличаться от соответствующих ММП оценок. Это связано с тем, что при их оценивании (формула (2.46)) использовались не окончательные, а промежуточные значения общих параметров. Однако эти оценки можно уточнить – для этого применяется процедура обратного последовательного байесовского оценивания (ОПБО).

2.4. Обратное последовательное байесовское оценивание

Суть метода ОПБО очень проста – после того как проведена процедура ПБО и получены окончательные оценки всех *общих* параметров, еще раз последовательно пересчитываются все серии данных и находятся окончательные оценки *частных* параметров. Для этого в конце ПБО нужно построить «окончательную» апостериорную информацию обо всех общих параметрах и затем использовать ее при пересчете всех частных параметров как априорную по типу 2 (без учета дисперсии ошибки). Кроме того, при поиске оценок всех частных параметров следует исключить из него (зафиксировать) все общие параметры.

Пусть в результате ПБО получены оценки \hat{c}_0 общих параметров, определена апостериорная матрица \tilde{A} (2.34), а по ней построена априорная матрица H (2.35). Тогда целевая функция для каждой j -ой серии данных в ОПБО будет иметь вид

$$Q_j(c_0, c_j) = S_j(c_j) \exp\left[\frac{R(c_0)}{N_{w,j}}\right], \quad j = 1, \dots, M \quad (2.50)$$

где

$$S_j(c_j) = \sum_{i=1}^{N_j} w_{ji}^2 (y_{ji} - f_j(x_{ji}, c_j, \hat{c}_0))^2 g_{ji}^2 \quad (2.51)$$

и

$$R(c_0) = (c_0 - \hat{c}_0)^t H (c_0 - \hat{c}_0) \quad (2.52)$$

Заметим, что сумма квадратов (2.51) зависит только от частных параметров, т.к. величины \hat{c}_0 в ней являются не переменными, а константами. С другой стороны байесовский член (2.52) зависит только от общих параметров c_0 . Поэтому, видно, что минимум целевой функции (2.50) всегда достигается в точке $c_0 = \hat{c}_0$, а точка минимума по частным параметрам c_j в каждой серии определяется независимо, но с учетом значений оценок общих параметров. Все эти тривиальные результаты могли бы быть получены и в том случае, если бы мы просто положили в каждой серии

$$Q_j(c_j) = S_j(c_j), \quad j = 1, \dots, M \quad (2.53)$$

т.е. искали все частные параметры при фиксированных значениях общих параметров, без всякой априорной информации.

Действительно, процедура ОПБО направлена, прежде всего, не на получение оценок частных параметров, а на уточнение их ковариационной матрицы. Суть метода состоит в том,

что целевые функции (2.50) и (2.53) имеют одинаковые точки минимума, но *разные информационные матрицы* –

$$A_j = V_j^t V_j + \frac{S_j(\hat{c}_j)}{N_{w,j}} H \quad (2.54)$$

и

$$A_j = V_j^t V_j \quad (2.55)$$

соответственно. Здесь V_j – это матрица с размерностью $(p_j - r) \times N_j$, которая имеет элементы

$$V_{j,\alpha i} = w_{ji} \frac{\partial f_j(\mathbf{x}_i, \mathbf{c}_j, \hat{c}_0)}{\partial c_{j,\alpha}}, \quad \alpha = 1, \dots, p_j; \quad i = 1, \dots, N_j \quad (2.56)$$

Следует отметить, что в ОПБО применяется априорная информация именно второго типа, т.е. без переноса данных об ошибке измерения. Это делается по тем же соображениям, по которым фиксируются оценки общих параметров в сумме квадратов (2.51) – иначе одни и те же данные будут использоваться два раза для оценки общих параметров c_0 и σ^2 .

Метод обратного последовательного байесовского оценивания, можно представить следующим алгоритмом:

- 1) выполняется процедура ПБО
- 2) строится априорная информация об общих параметрах по формулам (2.33)–(2.37); (2.57)
- 3) каждая серия обрабатывается с целевой функцией (2.50);
- 4) оценки частных параметров являются результатом ОПБО.

2.5. Пример применения ПБО

Проиллюстрируем предлагаемую технику численным примером, моделирующим в простейшем виде типовую ситуацию. Рассмотрим задачу оценивания кинетических параметров химических реакций [38]. Подробности всех расчетов можно найти в [126], где приведен файл *SBE.XLS*, содержащий исходные данные и порядок расчетов.

Часто прямое измерение концентрации $C(t)$ затруднительно и вместо концентрации измеряется величина $D(t)$ (оптическая плотность), связанная с C уравнением

$$D = b + aC, \quad (2.58)$$

где a и b — это параметры метода измерения. При этом нормировочный множитель a не зависит от измеряемого вещества, а значение фона b может меняться. Для оценивания па-

раметров используются два набора данных: калибровочные данные и кинетические данные.

Калибровочные данные представлены на Рис. 2.1. Они содержат значения оптической плотности D (отклик), измеренной при различных значениях концентрации C (независимая переменная). Эти данные используются как источник априорной информации о калибровочных параметрах a и b в модели (2.58).

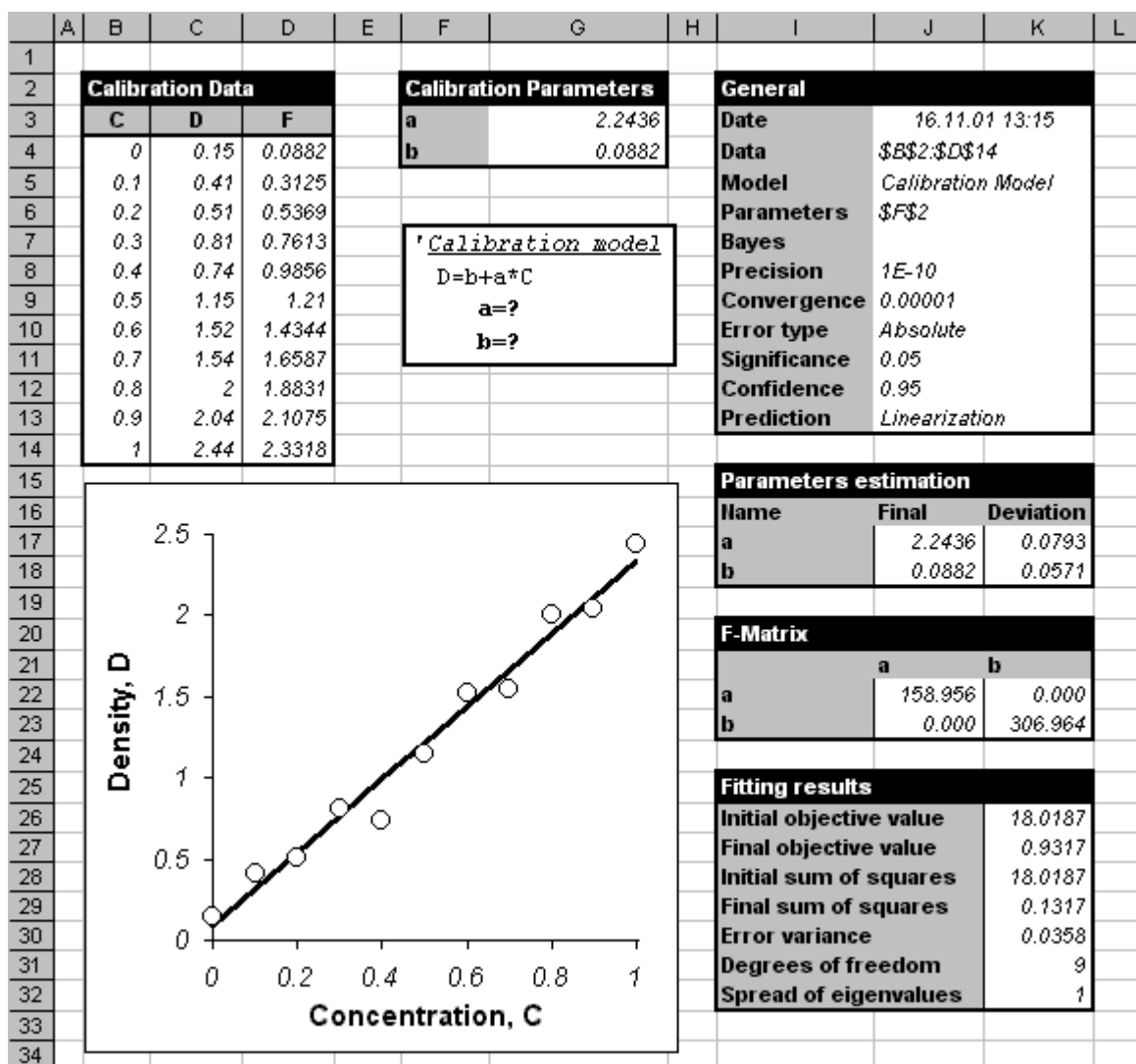


Рис. 2.1 Обработка калибровочных данных системой Fitter

Второй набор данных (Рис. 2.2) содержит результаты кинетического эксперимента, где оптическая плотность D (отклик) измерялась в ходе химической реакции n -го порядка.

$$D=c+aC \quad \frac{dC}{dt} = -kC^n, \quad C(0) = C_0 \quad (2.59)$$

как функция времени t . Задача состоит в нахождении кинетических параметров C_0 , k и n при наличии мешающих параметров a (тот же, что и в калибровочных данных) и c (новое значение).

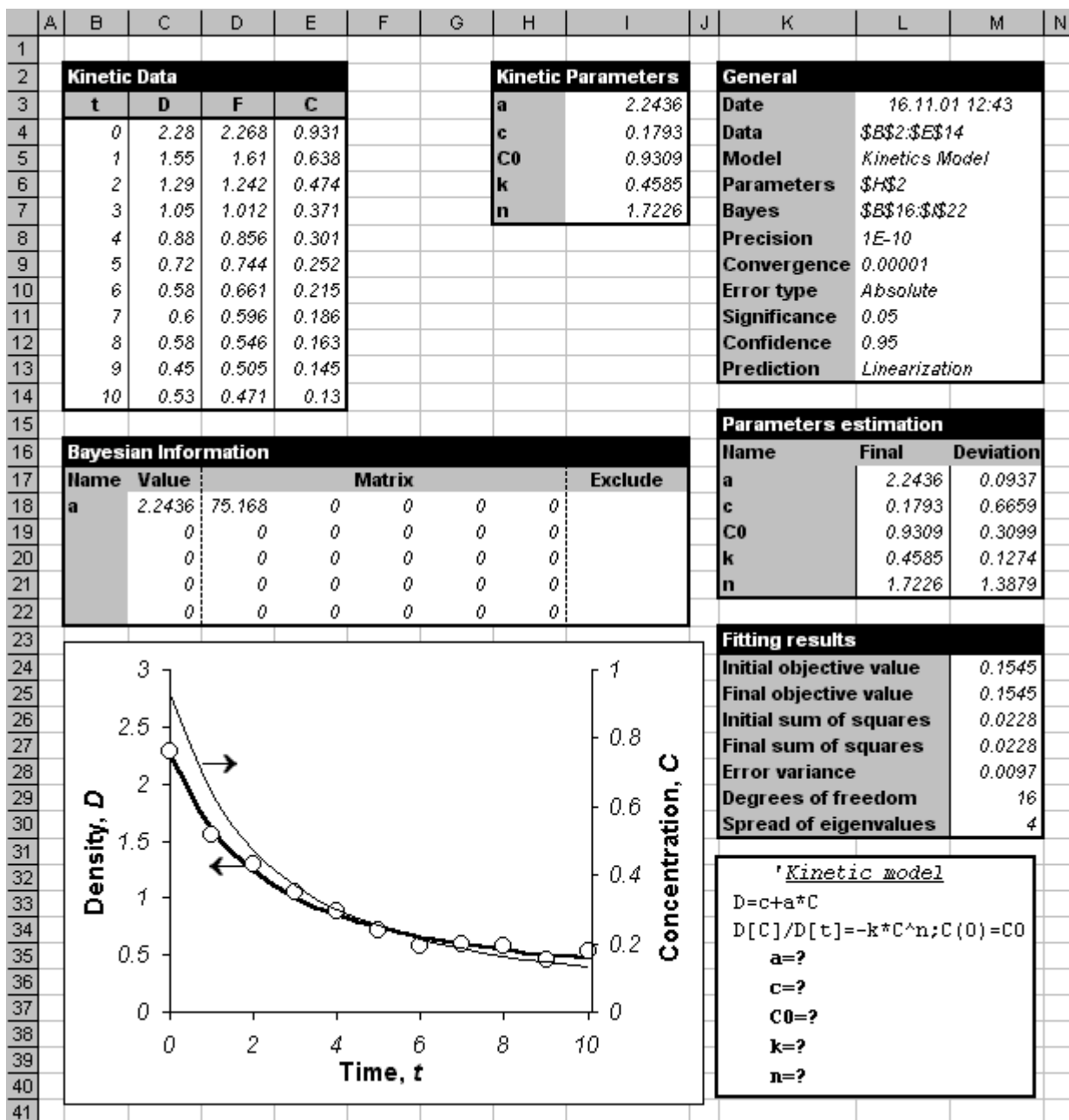


Рис. 2.2 Обработка кинетических данных с учетом априорной информации системой Fitter

Решение задачи состоит из трех шагов: 1) подгонка калибровочных данных, 2) создание байесовской информации и 3) оценивание кинетических данных.

На первом шаге (см. Рис. 2.1) оцениваются калибровочные параметры (1.34) –

$$a=2.2436, b=0.0882,$$

вычисляется F -матрица (1.42) –

$$F = \begin{bmatrix} 263.087 & 375.839 \\ 375.839 & 751.677 \end{bmatrix},$$

взвешенная дисперсия ошибки измерений (1.33) –

$$s^2 = 0.0146,$$

и число степеней свободы (1.44) –

$$N_f = 9.$$

Цель следующего шага – создать байесовскую информацию об общем параметре a , который должен быть перенесен на следующем этапе подгонки в кинетические данные. Априорное число степеней свободы и априорное значение дисперсии устанавливаются равными значениям, найденным на первом этапе. Тем самым, мы предполагаем, что ошибка эксперимента в кинетических и в калибровочных данных одинакова и может быть использована для уточнения дисперсии ошибки измерения. Так как b не является общим параметром, то его надо исключить.

Используя технику, описанную в предыдущем разделе, получаем следующий набор данных для априорной информации типа 1:

1) значения априорных параметров (2.36):

$$a=2.2436 \quad c=0 \quad C_0=0 \quad k=0 \quad n=0$$

2) априорную информационную матрицу, пересчитанную с помощью уравнений (2.34) и (2.35)

$$H = \begin{bmatrix} 75.168 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

3) априорную оценку взвешенной дисперсии

$$s_0^2 = 0.0146$$

4) априорное число степеней свободы

$$N_0 = 9$$

На третьем этапе кинетические данные подгоняются моделью (2.59) с учетом байесовской информации, созданной выше.

Результаты оценивания представлены на Рис. 2.2, и в Табл. 2.1 – в колонке, озаглавленной **ПБО**.

Табл. 2.1 Оценки параметров модели (2.59)

| Параметры | Оценки | | Среднеквадратичные отклонения оценок | | | |
|-----------|--------|---------|--------------------------------------|--------|--------|--------|
| | ПБО | МНК ММП | ПБО | МНК | ММП | ОПБО |
| Общие | a | 2.2436 | 0.0937 | – | 0.0937 | 0.0937 |
| | C_0 | 0.9309 | 0.3099 | 0.1785 | 0.3099 | 0.3099 |
| | k | 0.4585 | 0.1274 | 0.0735 | 0.1274 | 0.1274 |
| | n | 1.7226 | 1.3879 | 0.8058 | 1.3879 | 1.3879 |
| Частные | b | 0.0882 | 0.0682 | 0.0682 | 0.0554 | 0.0531 |
| | c | 0.1793 | 0.6658 | 0.3866 | 0.6659 | 0.6658 |

Конечно, этот пример может быть решен и по-другому. Так кинетические данные могут быть обработаны и без байесовской информации, обычным МНК, но с фиксированным значением параметра $a \equiv 2.2436$. Величины оценок параметров будут те же, что и в ПБО, но их стандартные отклонения изменятся. Колонка таблицы, озаглавленная **МНК**, содержит эти отклонения. Видно, что они значительно ниже тех, которые получены с помощью ПБО, за исключением отклонения параметра b .

Для того чтобы понять, какие результаты «правильнее» (т.е. ближе к эффективным оценкам, обладающим минимальной дисперсией), применим традиционный метод обработки – с помощью совместного ММП оценивания, когда используется целевая функция (2.39), состоящая, в этом примере, из двух частей. Оценки параметров, полученные таким методом, будут опять те же, как и должно быть по доказанному выше основному свойству ПБО. Различие имеется только в величинах среднеквадратичных отклонений оценок параметров (колонка **ММП** в Табл. 2.1). Причем отличие от ПБО заметно только для параметра b . Происхождение этого отличия совершенно понятно – параметр b в методе ПБО определялся только по калибровочным данным, причем совместно с параметром a . В этой процедуре значение общего параметра a определялось без учета кинетических данных, которые, разумеется, улучшили бы его оценку. Действительно, среднеквадратичные отклонения оценок параметра a после учета только калибровочных и после учета еще и кинетических данных равны соответственно: 0.1153 (см. Рис. 2.1) и 0.0937 (см. Рис. 2.2) – получается выигрыш около 20%. С другой стороны, тот же параметр b оценивался в ММП совместно с параметром a по всем данным сразу – и калибровочным и кинетическим. Та-

ким образом, происхождение исследуемого отличия – это то самое взаимовлияние частных и общих параметров, о котором шла речь в предыдущем разделе.

Оценку частного параметра b можно улучшить, если применить процедуру обратного последовательного байесовского оценивания (ОПБО). Результаты применения ОПБО приведены в колонке, озаглавленной **ОПБО**.

Очевидно, что в результате обработки данных важно получить не только оценки неизвестных параметров, но и о правильные характеристики точности оценивания, например, среднеквадратичные отклонения этих оценок. Отвлекаясь пока от проблем, связанных с нелинейностью моделей, которые будут рассмотрены в следующей главе, мы можем сделать следующий вывод – в том случае, когда затруднительно применить ММП, можно использовать метод последовательного байесовского оценивания. Рассмотренный пример показывает, что метод ПБО дает очень близкие результаты для характеристик точности оценивания

2.6. Практическое использование метода ПБО

Метод последовательного байесовского с успехом применяется для решения самых разных задач нелинейного регрессионного анализа. Так в главе 6 подробно разобраны несколько реальных задач по обработке термограмм полимеров. Под термограммой понимается любой сигнал, являющийся функцией от переменной, линейно растущей со временем температуры. В частности, там продемонстрировано, как использовать ПБО для построения доверительных интервалов при прогнозе.

В разделе 4.1 показано, что основная проблема в процедуре минимизации целевой функции – это многократное (псевдо)обращение информационной матрицы A . Размерность этой матрицы определяется числом неизвестных параметров p (2.31). Это число может быть очень велико, как это видно, например, из задачи оценивания кинетических параметров по спектральным данным [43], разобранный в главе 7. Применение ПБО решает эту проблему, разбивая одну большую задачу оценивания на длинную последовательность маленьких задач.

В главе 8 представлен другой пример использования ПБО для обработки многооткликных данных. В задаче [44, 45, 46] исследуются данные по ускоренному старению шинных резин. При трех температурах измерялись кинетические кривые изменения семи показателей свойств материала: разрывного удлинения, разрывной прочности, и т.п. Показатели измерялись различными методами, поэтому дисперсии ошибок были разными, и каждый

из них описывался своей моделью. Эти модели содержали двенадцать неизвестных параметров, из которых только четыре были общими для всех моделей. При использовании ММП понадобилось бы минимизировать целевую функцию (2.39), состоящую из семи частей и одновременно оценивать двенадцать параметров. При этом пришлось бы решать неприятные проблемы [3], связанные с тем, что разные показатели имеют разные ошибки измерения. Применяя метод ПБО, мы успешно оценили все параметры: общие – по всей совокупности данных и частные – только по данным о соответствующем показателе.

Возникает естественный вопрос об однозначности процедуры последовательного оценивания. Неоднозначность может проистекать из разных способов разбиения массива данных на серии и порядка их обработки. Из доказанной в разделе 2.3 теоремы следует, что если регрессия линейна, то неоднозначности нет. Кроме того, можно утверждать, что асимптотически, при большом числе экспериментальных данных различие в оценках будет стремиться к нулю. Это показывает и примеры содержащийся в главах 6 и 7. В нелинейном случае, как мы полагаем, зависимость конечной оценки от способа разбиения на серии будет выборочной. Эта изменчивость является, на самом деле, не недостатком, а преимуществом метода. Во-первых, ее можно устранить, усреднив оценки по всем возможным последовательностям. Хорошо известно, что такой прием, известный сейчас как *bagging* [24-26] значительно улучшает устойчивость и качество оценок. С другой стороны, исследуя изменчивость выборочных оценок по аналогии с известными методиками *jack-knife* и *bootstrap* [23], можно уменьшить смещение и установить доверительные области оценок.

При построении моделей различных процессов большую роль играют результаты, почерпнутые из разных источников. При этом возникает противоречие. С одной стороны, нет возможности хранить в банках данных всю первичную информацию, как из-за больших объемов, так и из-за разнородности [41]. С другой стороны, при хранении только результатов анализа данных, например, оценок кинетических констант, происходит потеря информации. Другой исследователь не имеет возможности корректно использовать полученные ранее оценки, и вынужден повторять эксперименты, уже проведенные его предшественниками. Предлагаемая методика свертки и использования регрессионной информации в компактной байесовской форме позволяет снизить потерю информации, содержащейся в первичном экспериментальном материале, эффективно использовать ее при рассмотрении альтернативных механизмов описания процессов, количественно сравнивать разные описания. Примеры ее применения описаны в работах [42– 53].

2.7. Результаты главы 2

В этой главе был представлен новый метод оценки регрессионных параметров – *последовательное байесовское оценивание* (ПБО). Этот метод разбивает одну большую задачу оценивания на последовательность маленьких задач, связанных между собой *априорной информацией*, передаваемой по цепочке. Разделяя исходные данные на серии, методом ПБО можно решать как задачи с большим числом однородных экспериментальных данных, так и задачи, в которых имеется большое число неизвестных параметров в разнородных данных. Метод ПБО является одним из главных статистических инструментов, используемых в работе, поэтому эта глава посвящена теоретическому исследованию его свойств.

В первом разделе продемонстрировано, как работает *метод максимума правдоподобия* (ММП) в случае, когда имеется априорная информация. Получены соотношения, определяющие вид целевой функции для априорной информации первого и второго типов.

Во втором разделе показано, как, по результатам оценивания, строится *апостериорная информация* и как она превращается в *априорную информацию*. Дано определение простейшего варианта ПБО, пригодного для решения задач, в которых нет частных параметров.

Метод ПБО в общем случае – с учетом *общих и частных* параметров модели – был изложен в третьем разделе. Доказано главное свойство этого метода – *теорема об эквивалентности ПБО–оценок и ММП–оценок*. Доказано, что для линейных моделей они совпадают, и, что для нелинейных моделей ПБО–оценки асимптотически сходятся к ММП–оценкам. Этот результат имеет большое теоретическое значение. Из него следует, что ПБО–оценки обладают теми же асимптотическими свойствами, что и ММП–оценки: эффективностью, состоятельностью, несмещенностью и т.п. Кроме того, из этой теоремы вытекает, что результат ПБО асимптотически не зависит от порядка, в котором обрабатываются данные. Тем самым был подведен теоретический базис, необходимый для широкого практического применения этой процедуры.

Четвертый раздел посвящен проблеме *обратного последовательного байесовского оценивания* (ОПБО) частных параметров. Этот метод является простым обобщением ПБО. Он применяется после того, как проведена процедура ПБО и получены окончательные оценки всех *общих* параметров. С помощью ОПБО еще раз последовательно пересчитываются все серии данных и находятся ПБО–оценки *частных* параметров.

Метод ПБО имеет большое практическое значение. На модельном примере было показано, как он применяется и проведено его сравнение с другими подходами: ММП, МНК.

Основное преимущество ПБО перед другими методами можно проиллюстрировать простейшей механической аналогией. Разбивая одну большую работу на цепочку небольших, но связанных между собой действий, ПБО действует подобно рычагу, который проигрывает в пути, но зато выигрывает в силе. Поэтому с его помощью можно последовательно анализировать очень большие, в принципе, неограниченные, объемы экспериментальных данных. Соответствующие примеры рассмотрены в третьей части работы.

3. Учет нелинейности регрессии

Конечная цель регрессионного анализа – это предсказание поведения системы за пределами области наблюдений. Только содержательные модели, базирующиеся на научных представлениях о сути происходящих процессов, позволяют получить корректный результат [38, 39]. При этом важно определить не только само значение прогноза, но и оценить его точность, которая зависит, прежде всего, от дальности экстраполяции, а не от ошибки измерения. Точность оценивания нелинейных регрессионных моделей – это задача, в которой наиболее остро проявляется нелинейность модели. Действительно, поиск оценок параметров модели – нелинейная оптимизация – исследован настолько хорошо, что его практическое применение уже ни сколько не сложнее, чем для линейных моделей (см. раздел 4.1). В то же время задача определения точности изучена пока очень плохо – слишком значительные аналитические проблемы возникают перед исследователем, пытающимся предложить простые алгоритмы для ее решения.

Наиболее удобно представлять результат прогноза в виде *доверительного интервала*. Существует два подхода к нахождению таких интервалов – аналитический (или параметрический) и выборочный (или непараметрический). Первый подход использует некоторую достаточную статистику с известным законом распределения. Его можно применять только в сравнительно простых случаях, например, для линейной регрессии. Второй способ требует наличия большого количества повторных измерений, по которым требуемый интервал и строится как соответствующий процентиль. К сожалению, для нелинейной регрессии ни один из этих подходов не годится, т.к. первый слишком неточен, а второй слишком медленен.

В этой главе предлагаются новые подходы к проблеме точности прогнозирования. В первом разделе рассматриваются различные традиционные способы построения доверительных интервалов. Во втором разделе предлагается новый способ доверительного оценивания, названный “*связанное моделирование*”, который является разновидностью метода Монте-Карло, понимаемого в широком смысле [27, 56]. Все эти методы различаются как по сложности выполнения, так и по точности, т.е. по близости результата к «истинному» доверительному интервалу. Для иллюстрации этих алгоритмов и для проверки их точности используется модельный пример, приведенный в третьем разделе. Различия в доверительных интервалах, построенных для нелинейной модели различными методами, может быть очень велико, а может быть и пренебрежимо мало с «инженерной» точки зрения. В по-

следнем, четвертом разделе этой главы дается новое определение *коэффициента нелинейности*, который предлагается использовать для принятия решения о том, какой метод можно использовать в данной задаче.

3.1. Традиционные методы построения доверительных интервалов

Рассмотрим нелинейную регрессию (1.11)

$$y=f(\mathbf{x},\mathbf{a})+\boldsymbol{\varepsilon}, \quad (3.1)$$

где \mathbf{a} – вектор неизвестных параметров (1.6) размерностью p , \mathbf{x} – вектор факторов (1.2), $\boldsymbol{\varepsilon}$ – это случайные ошибки (1.13), удовлетворяющие условиям (1.14)-(1.16). Для поиска оценок параметров мы будем использовать метод максимума правдоподобия (ММП), описанный в разделах 1.2 и 2.1

При исследовании доверительных интервалов одна и та же величина часто выступает то, как случайная переменная, то, как детерминированное значение, являющееся реализацией этой случайной величины. Поэтому, для удобства изложения будем использовать следующие обозначения. Целевую функцию $Q(\mathbf{a})$, определенную соотношениями (1.34) или (2.10) или (2.11), будем теперь обозначать

$$Q(\mathbf{a}|\mathbf{y}). \quad (3.2)$$

Символ $|\mathbf{y}$ здесь означает, что целевая функция – это случайная величина, зависящая от реализации экспериментальных данных. Аналогично будем обозначать значения оценок параметров

$$\boldsymbol{\alpha}(\mathbf{y}|\mathbf{a},\sigma^2)=\hat{\mathbf{a}} \quad (3.3)$$

найденных как точка минимума целевой функции

$$\boldsymbol{\alpha}(\mathbf{y}|\mathbf{a},\sigma^2)=\arg \min Q(\mathbf{a}|\mathbf{y}). \quad (3.4)$$

и оценки взвешенной дисперсии ошибки (2.15) или (2.16)

$$s^2(\mathbf{y}|\mathbf{a},\sigma^2)=\frac{Q(\boldsymbol{\alpha}|\mathbf{y})}{N_f}. \quad (3.5)$$

Такие обозначения подчеркивают то, что все эти оценки – случайные функции вектора откликов \mathbf{y} , которые, однако, зависят от «истинных» значений параметров модели \mathbf{a} и σ^2 .

Пусть $g(\mathbf{a})$ – это дифференцируемая скалярная функция вектора параметров \mathbf{a} . Тогда статистика

$$\gamma = g(\boldsymbol{\alpha}(\mathbf{y}|\mathbf{a})) \quad (3.6)$$

является точечной оценкой величины $g(\mathbf{a})$. Случайная величина $\gamma_+(P)$ называется верхним пределом одностороннего доверительного интервала для величины g на уровне достоверности P , если

$$\text{Prob}\{\gamma_+(P) > g\} \geq P \quad (3.7)$$

Для простоты мы будем рассматривать только верхний предел, имея в виду, что нижний предел – γ_- легко определяется по верхнему как

$$\gamma_-(P) = \gamma_+(1-P).$$

В качестве $g(\mathbf{a})$ будем рассматривать, прежде всего, значение самой регрессионной функции (3.1) в некоторой точке \mathbf{x}_p лежащей за пределами области наблюдения

$$g(\mathbf{a}) = f(\mathbf{x}_p, \mathbf{a}), \quad \mathbf{x}_p \neq \mathbf{x}_i \quad i = 1, \dots, N \quad (3.8)$$

Пусть $F(\mathbf{z}, u | \mathbf{a}, \sigma^2)$ – это совместная функция распределения статистик (3.4) и (3.5) т.е.

$$F(\mathbf{z}, u | \mathbf{a}, \sigma^2) = \text{Prob}\{(\mathbf{a} < \mathbf{z}) \wedge (s^2 < u)\} \quad (3.9)$$

тогда распределение $G(w | \mathbf{a}, \sigma^2)$ случайной величины (3.6) определяется как

$$G(w | \mathbf{a}, \sigma^2) = \text{Prob}\{\gamma < w\} = \int_{g(\mathbf{a}) < w} dz_1 \dots \int dz_p \int_0^\infty du \frac{\partial^p F(\mathbf{z}, u | \mathbf{a}, \sigma^2)}{\partial z_1 \dots \partial z_p} \quad (3.10)$$

Рассмотрим некоторые конкретные экспериментальные данные \mathbf{y}_θ , по которым найдены соответствующие значения оценок α_θ и s_θ^2 – реализации случайных величин (3.4) и (3.5).

Тогда доверительный интервал (3.7) может быть построен с помощью статистики (3.6) как

$$\gamma_+(P) = G^{-1}(P | \alpha_\theta, s_\theta^2) \quad (3.11)$$

где G^{-1} - функция обратная к (3.10) т.е. квантиль (1.48) распределения

$$G(G^{-1}(P | \dots) | \dots) = P$$

Конечно, распределение (3.9) в общем случае не известно, поэтому для построения доверительных интервалов используются распределения \hat{F} и \hat{G} , с той или иной точностью приближающие распределения (3.9) и (3.10).

Рассмотрим теперь основные (традиционные) методы построения доверительных интервалов.

Стохастическая аппроксимация (S-метод)

Известно [3], что для *линейной* нормальной регрессии распределение (3.9) является гауссовым. В нелинейном случае это утверждение верно лишь асимптотически, при большом

числе наблюдений. Тем не менее, можно построить доверительный интервал, опираясь на это предположение. Положим

$$\hat{F} = \Phi(\mathbf{a}, \sigma^2 \mathbf{A}^{-1}) \quad (3.12)$$

где Φ – функция p -мерного нормального распределения со средним значением \mathbf{a} и матрицей ковариаций $\sigma^2 \mathbf{A}^{-1}$. Матрица \mathbf{A} – это информационная матрица, определенная в формулах (1.38) или (2.20) или (2.30). Распределение \hat{G} в этом приближении также будет нормальным с параметрами:

$$E(\gamma) = g_0, \quad D(\gamma) = s_0^2 \mathbf{v}^t \mathbf{A}^{-1} \mathbf{v}$$

где $g_0 = g(\mathbf{a}_0)$, а $\mathbf{v} = \nabla g(\mathbf{a}_0)$. Из (3.11) получаем, что

$$\gamma_+(P) = g_0 + z(P) s_\gamma \quad (3.13)$$

где $z(P)$ – это P -квантиль нормального распределения (1.49), $s_\gamma = \sqrt{D(\gamma)}$.

В некоторых случаях (например, когда $g(\mathbf{a}) \geq 0$) доверительный интервал получается гораздо точнее, если в качестве распределения \hat{G} брать не нормальное, а логнормальное распределение. Легко видеть, что в этом случае выражение для доверительного интервала имеет вид

$$\gamma_+(P) = g_0 e^{z(P) u_\gamma} \quad (3.14)$$

где

$$u_\gamma^2 = \ln \left[0.5 \left(1 + \sqrt{1 + 4 \frac{s_\gamma^2}{g_0^2}} \right) \right]$$

Доверительный интервал в форме (3.13) или (3.14) построить легко, тем более что матрица \mathbf{A} все равно вычисляется в ходе поиска параметров регрессии. Однако его точность в нелинейном случае не удовлетворительна.

Линеаризация (L-метод)

В некоторых случаях можно подобрать преобразования, превращающие регрессионную модель в линейную. Пусть существуют такие отображения

$$\tilde{y} = h(y), \quad \tilde{\mathbf{a}} = \mathbf{b}(\mathbf{a}), \quad \tilde{\mathbf{x}} = \mathbf{w}(\mathbf{x}),$$

что одновременно

$$h(f(\mathbf{w}(\mathbf{x}), \mathbf{b}(\mathbf{a}))) = \tilde{\mathbf{f}}^t(\tilde{\mathbf{x}}) \tilde{\mathbf{a}} \quad \text{и} \quad h(g(\mathbf{b}(\mathbf{a}))) = \tilde{\mathbf{g}}^t \tilde{\mathbf{a}}.$$

Оценивание новых параметров $\tilde{\mathbf{a}}$ сводится тогда к поиску минимума квадратичной функции

$$\tilde{Q}(\tilde{\mathbf{a}}|\tilde{\mathbf{y}}) = (\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\tilde{\mathbf{a}})^t (\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\tilde{\mathbf{a}})$$

где. $\tilde{\mathbf{y}} = h(\mathbf{y})$, $\tilde{\mathbf{X}} = (\tilde{f}(\tilde{\mathbf{x}}_1), \dots, \tilde{f}(\tilde{\mathbf{x}}_N))^t$

В этом случае доверительный интервал (3.11) имеет вид

$$\gamma_+(P) = h^{-1}(\tilde{\mathbf{g}}_0^t \tilde{\mathbf{a}}_0 + z(P)\tilde{s}_\gamma) \quad (3.15)$$

где

$$\tilde{\mathbf{a}}_0 = (\tilde{\mathbf{X}}^t \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^t \tilde{\mathbf{y}}_0, \quad \tilde{s}_\gamma^2 = \frac{\tilde{Q}(\tilde{\mathbf{a}}_0|\tilde{\mathbf{y}}_0)}{N-p} \tilde{\mathbf{g}}^t (\tilde{\mathbf{X}}^t \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{g}},$$

а $z(P)$ – это P -квантиль нормального распределения (1.49).

Такой способ построение интервала очень распространен [1, 64]. Однако, хорошо известно, что преобразование координат в исходной регрессии может приводить к серьезным ошибкам в оценивании. Ниже мы увидим это на примере.

Имитационное моделирование (М-метод)

Методы, изложенные выше, опираются на довольно грубые приближения вида распределения оценок параметров \mathbf{a} и прогнозируемой функции $g(\mathbf{a})$. Избежать этих недостатков, можно используя методы имитационного моделирования [27]. Основная идея этих методов состоит в том, чтобы смоделировать с помощью метода Монте-Карло распределение оценок параметров (3.4) и (3.5) и взять выборочный процентиль распределения (3.6) в качестве доверительного интервала. Рассмотрим этот подход подробнее.

Пусть \mathbf{a}_0 и s_0^2 – это реализации оценок (3.4) и (3.5), полученные в эксперименте, т.е. при $\mathbf{y}=\mathbf{y}_0$. Рассмотрим случайные величины \mathbf{a}^*

$$\mathbf{a}^*(\mathbf{y} | \mathbf{a}_0, s_0^2) = \arg \min Q(\mathbf{a} | \mathbf{y}^*). \quad (3.16)$$

и γ^*

$$\gamma^* = g(\mathbf{a}^*). \quad (3.17)$$

где

$$\mathbf{y}^* = f(\mathbf{x}, \mathbf{a}_0) + \boldsymbol{\varepsilon}^* \quad (3.18)$$

Случайные величины $\boldsymbol{\varepsilon}^*$ (псевдоошибки) распределены нормально с дисперсией s_0^2 , т.е.

$$\boldsymbol{\varepsilon}^* \sim N(\mathbf{0}, s_0^2 \mathbf{I}). \quad (3.19)$$

Из этих формул следует, что случайные величины α^* строятся так, что их «истинными» значениями являются a_0 – реализации оценок, а ошибки имеют «истинное» значение дисперсии s_0^2 . По-видимому, впервые такой подход был предложен в [55]. Главное допущение этого метода состоит в том, что величины (3.16) и (3.17) “подобны”, соответственно, (3.4) и (3.5). Точнее, предполагается, что существует монотонное отображение $r(\cdot)$ (сам вид его не важен) такое, что $r(\gamma) - r(g)$ и $r(\gamma^*) - r(g_0)$ имеют одинаковое распределение, симметричное относительно нуля [22]. Если это верно, то доверительный интервал для g может быть получен как выборочный процентиль распределения случайной величины γ^* . Практически это означает следующий алгоритм:

- 1) по исходным данным y_0 определяются величины a_0 и s_0^2
- 2) строится нормальный независимый вектор ошибок (3.19)
- 3) создаются модельные данные (3.18) (3.20)
- 4) оцениваются α^* и γ^* по (3.16) и (3.17)
- 5) независимо повторяя шаги 2)-4) получаем выборку значений $\gamma_1^*, \dots, \gamma_M^*$

Тогда

$$\gamma_+(P) = P\text{-процентиль} \{ \gamma_1^*, \dots, \gamma_M^* \} \quad (3.21)$$

Опыт показывает, что M -метод дает очень точные значения для доверительных интервалов. Сомнения вызывает только применение нормального распределения для генерирования ошибок на втором шагу алгоритма. Как альтернативу можно использовать *bootstrap* метод.

Bootstrap (B-метод)

Этот метод был предложен и развит в работах [21-23, 56]. Его основная идея состоит в том, чтобы заменить псевдослучайные нормальные ошибки (3.19) на случайные повторные выборки из вектора остатков

$$e = y_0 - f(x, a_0). \quad (3.22)$$

т.е.

$$e^* = (e_{i_1}, \dots, e_{i_N})^t \quad (3.23)$$

Результат получается неплохим, однако главным недостатком этого и предыдущего методов является большие затраты времени на реализацию. Очевидно, что главные неприятности связаны с шагом 4) в алгоритме (3.20). Если оценивается сложная модель с большим числом нелинейных параметров, то время, затрачиваемое на одну минимизацию, достига-

ет 10 секунд. Для уверенного прогноза необходимо выполнить не менее $M=1000$ повторений, что дает 10000 сек или более 3 часов работы компьютера.

3.2. Новые методы построения доверительных интервалов

В работе [57] были предложены два новых метода для построения доверительных интервалов в нелинейной регрессии. Рассмотрим их подробно.

Свободное моделирование (*F-метод*)

Ошибки при построении доверительных интервалов (3.13), (3.14) и (3.15) могут возникать, как из-за ненормальности распределения (3.9), так и из-за нелинейности функции (3.6). Уточним доверительный интервал (3.13), приняв во внимание нелинейную зависимость $g(\alpha)$. Рассмотрим нормальную векторную случайную величину α^*

$$\alpha^* \sim N(a_0, s_0^2 A^{-1}), \quad (3.24)$$

и порожденную ей случайную величину

$$\gamma^* = g(\alpha^*).$$

Принимая то же допущение об эквивалентности γ^* и γ , что и в *S*-методе, можно построить доверительный интервал как процентиль модельной выборки $\gamma_1^*, \dots, \gamma_M^*$. Для ее создания необходимо многократно получать реализации нормальных случайных величин α^* из распределения (3.24). Проще всего это сделать с помощью разложения Холецкого [58, 65] для матрицы $A^{-1} = B^2$, где B – это треугольная положительно определенная матрица. Тогда

$$\alpha^* = a_0 + s_0^2 B e, \quad \text{где } e \sim N(0, I).$$

Т.к. стадия поиска оценок отсутствует, то время, необходимое на реализацию этого алгоритма, значительно меньше, чем у *M*-метода. Однако *F*-метод не дает достаточно точных величин, что связано, на наш взгляд, с необъективностью использования нормального распределения для моделирования оценок параметров.

Связанное моделирование (*A-метод*)

Этот метод предлагается как компромисс между точностью *M*- и быстротой *F*-методов. Его идея состоит в том, чтобы, сохраняя принцип моделирования параметров, а не данных, уточнить вид распределения (3.24), приблизив его к реальному.

Основным предположением этого метода является то, что распределение случайной величины

$$c(\alpha | y) = \frac{Q(\alpha | y) - Q(a_0 | y)}{s_0^2} \quad (3.25)$$

подчиняется закону хи-квадрат с p степенями свободы.

$$c(\alpha | y) \sim \chi_p^2 \quad (3.26)$$

Здесь Q – это целевая функция (3.2), a_0 и s_0^2 – это реализации оценок (3.4) и (3.5), полученные в эксперименте, т.е. при $y=y_0$. Для линейной нормальной регрессии утверждение (3.26) верно в точности т.к.

$$c(\alpha) = (\alpha - a_0)^t A^{-1} (\alpha - a_0)$$

Для нелинейной модели его справедливость основана на законе больших чисел [59].

На этом эмпирическом факте базируется метод связанного моделирования, уточняющий метод свободного моделирования. Его основная идея состоит в том, чтобы использовать при моделировании оценок параметров распределение хи-квадрат с p степенями свободы вместо нормального распределения. Для того чтобы построить выборку параметров, удовлетворяющих условиям (3.25) и (3.26), воспользуемся стандартным приемом «приема-отклонения» [27].

Разобьем отрезок $[0, 1]$ на m не перекрывающихся интервалов точками P_k (см. Рис. 3.1)

$$0 = P_0 < P_1 < \dots < P_m = 1.$$

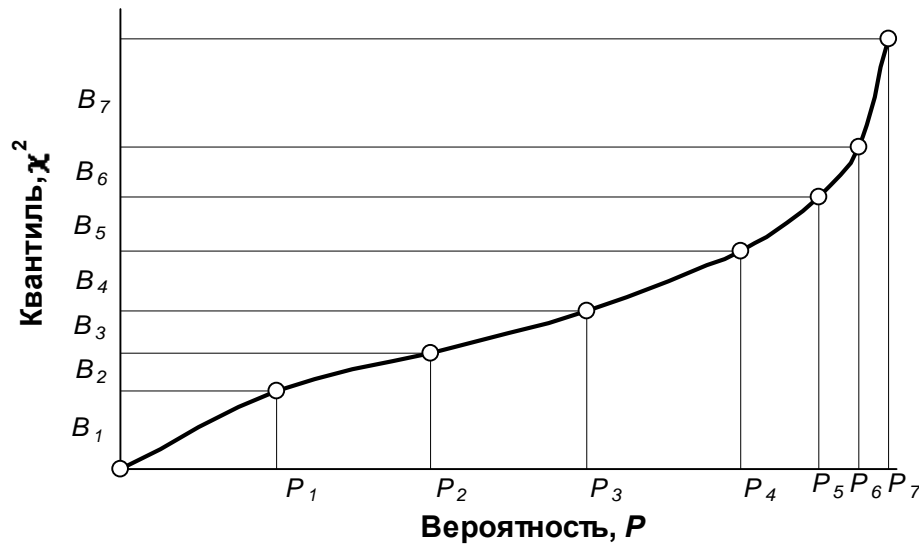


Рис. 3.1 Разбиение на «корзины» в A-методе

С помощью функции, обратной к распределению хи-квадрат с p степенями свободы (квантиль (1.50)), отобразим это разбиение на полуось

$$[0, +\infty] = \bigcup_1^m B_k$$

так, что

$$\text{Prob}\{\chi^2 \in B_k\} = P_k - P_{k-1}.$$

Набор «корзин» B_1, \dots, B_m мы будем использовать при построении алгоритма связанного моделирования. Очевидно, что при моделировании с M повторениями, в каждую из них в среднем должно попасть

$$M_k = M(P_k - P_{k-1}) \quad (3.27)$$

случайных реализаций.

Алгоритм процедуры связанного моделирования (А-метод) можно представить теперь следующим образом.

- 1) моделируем случайную величину α^* с распределением (3.24);
- 2) вычисляем величину $s(\alpha^* | y_0)$ по (3.25) и определяем в какую «корзину» B_k она попала;
- 3) если корзина уже заполнена, то возвращаемся к 1) иначе вычисляем γ^* ;
- 4) независимо повторяя шаги 1)-3) получаем выборку значений $\gamma_1^*, \dots, \gamma_M^*$;

$$(3.28)$$

Доверительный интервал определяется в соответствии с соотношением (3.21).

Численный эксперимент, приведенный в разделе 3.3, показывает, что А-метод дает весьма удовлетворительный интервал, причем временные затраты на его реализацию увеличиваются не намного, по сравнению с F-методом.

Смысл предлагаемого подхода становится прозрачен, если учесть, что выражение (3.25) для $s(\alpha | y_0)$ при фиксированном $y=y_0$ определяет область безразличия [3] оценок параметров

$$I(C) = \{a : Q(a | y_0) - Q(a_0 | y_0) < C\} \quad (3.29)$$

для функции правдоподобия (2.8) или (2.9)

Из определения (3.29) следует, что $I(C)$ – это область в пространстве неизвестных параметров a , ограниченная контуром, образованным срезом целевой функции на уровне C . Подробнее этот взгляд на область безразличия будет исследован в разделе 3.4.

Из предположения (3.26) следует, что

$$C = s_0^2 \chi_p^2(P),$$

где $\chi_p^2(P)$ – это P -квантиль распределения хи-квадрат с p степенями свободы (1.50).

Иными словами, утверждение –

$$\text{Prob}\{\alpha : I(P)\} \geq P, \quad (3.30)$$

выполняется с достаточной степенью точности для нелинейных моделей даже при небольших степенях свободы.

3.3. Модельный пример построения интервалов

Для проверки всех описанных выше методов использовался модельный пример, что позволило сравнить величины доверительных интервалов, построенных различными способами, с «истинными», известными заранее значениями модели. Истоки этого примера лежат в задачах, исследованных при прогнозировании срока службы резиновых изделий [46]. В качестве уравнения регрессии была выбрана следующая модель

$$f(\mathbf{x}, \mathbf{a}) = 1 - \exp[-(kx_2)^{a_1}] \quad (3.31)$$

$$\text{где } k = \exp\left(a_2 - \frac{1000a_3}{x_1}\right)$$

Здесь предикторы: x_1 – это температура (К), x_2 – это время (час). Элементы вектора параметров \mathbf{a} имеют следующий смысл: a_1 – масштабный, a_2 – это модифицированная предекспонента, a_3 – это модифицированная энергия активации. Характер и смысл модификации параметров a_2 и a_3 подробно объясняются в разделе 4.4.

Исходные данные (см. Рис. 3.2) представляют три серии измерений, проведенные при значениях фактора x_1 равным соответственно 383, 368 и 353, что отвечает типичным условиям ускоренных испытаний (см. главу 8). Значения фактора x_2 изменялись с шагом 24 час.

Значения откликов y_0 были созданы с помощью генератора псевдослучайных нормальных величин [27] при следующих «истинных» значениях параметров функции (3.31)

$$a_1=1.5, \quad a_2=17.0, \quad a_3=8.0, \quad \sigma^2=0.005 \quad (3.32)$$

Оценки \mathbf{a}_0 , полученные по этим данным, составили

$$\alpha_1=1.3865, \quad \alpha_2=18.8578, \quad \alpha_3=8.6897, \quad s^2=0.0067 \quad (3.33)$$

В качестве функции g (3.8) был взят результат экстраполяции модели на условия

$$x_1=293 \text{ К и } x_2=8760 \text{ час.}$$

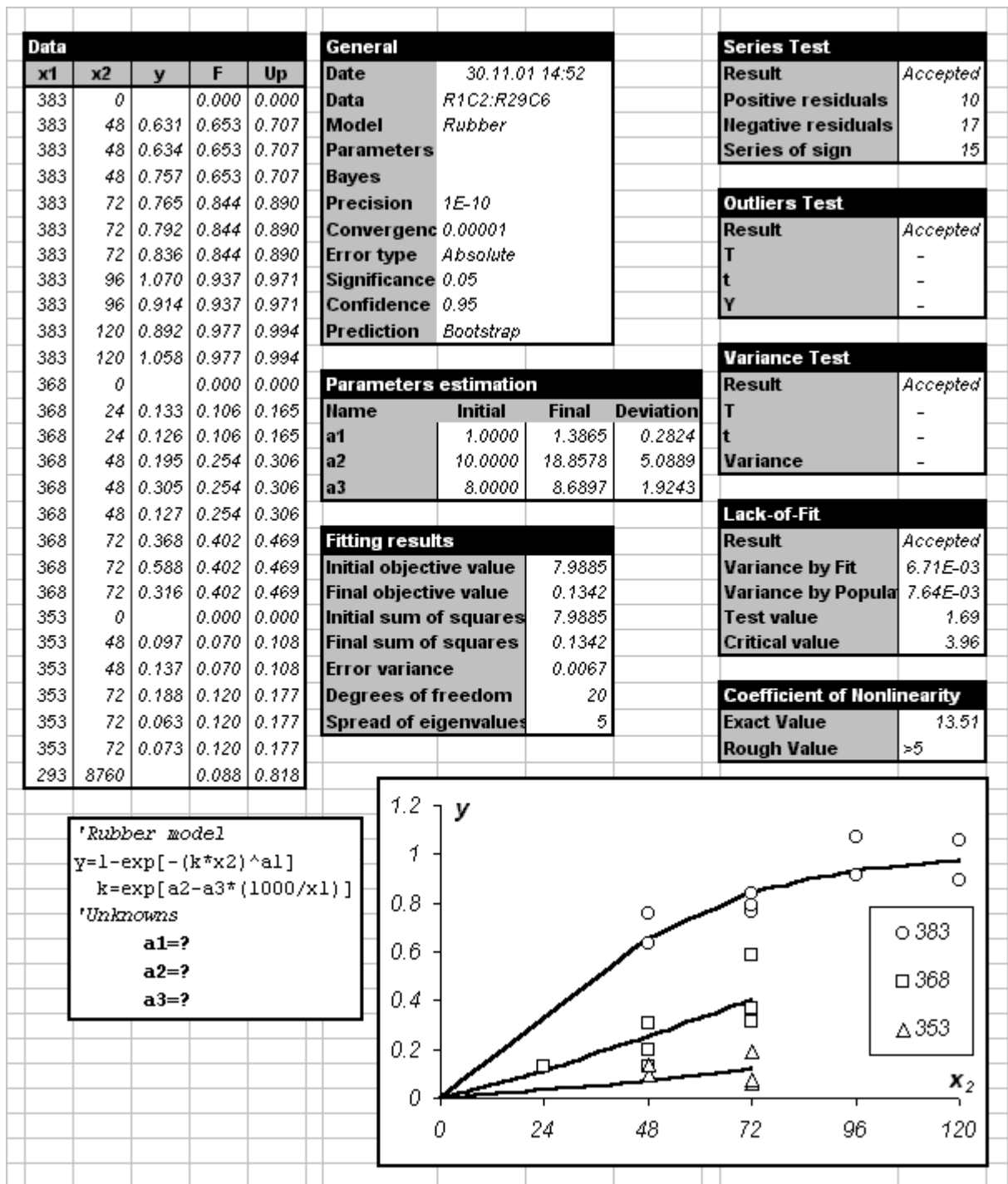


Рис. 3.2 Обработка модельных данных системой Fitter

Тем самым оценивался доверительный интервал для прогноза отклика на 1 год при комнатной температуре

$$g(\alpha) = f(8760, 293, \alpha).$$

“Истинный прогноз”, вычисленный при значениях (3.32), составил

$$g(1.5, 17, 8) = 0.1472, \tag{3.34}$$

а его оценка, вычисленная при значениях (3.33),

$$g_0 = g(1.3865, 18.8578, 8.6897) = 0.0877. \quad (3.35)$$

Матрица ковариаций получилась следующая

$$s^2 A^{-1} = \begin{bmatrix} 0.0797 & -1.2812 & -0.4851 \\ -1.2812 & 25.8972 & 9.7921 \\ -0.4851 & 9.7921 & 3.7029 \end{bmatrix}.$$

Методами, описанными выше, были построены односторонние (верхние) доверительные интервалы $\gamma_+(P)$ для величины g , при доверительной вероятности P , изменяющейся от 0.001 до 0.999 с шагом 0.001. Наибольшие различия наблюдались в области $P > 0.8$, которая и приведена на Рис. 3.3.

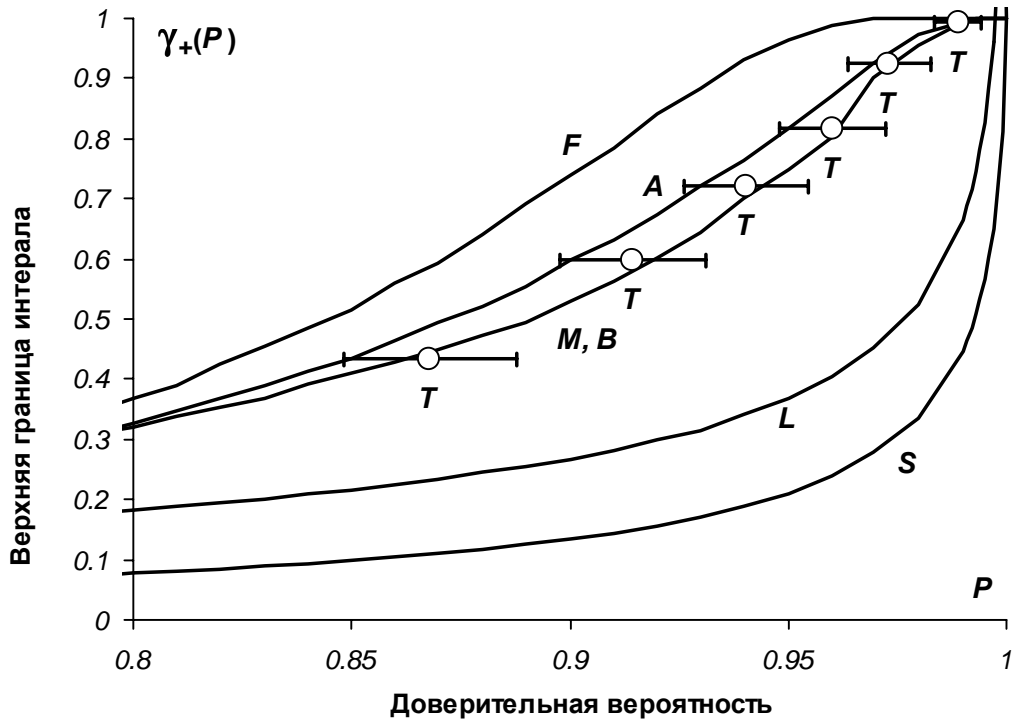


Рис. 3.3 Зависимость верхней границы интервала γ_+ от доверительной вероятности P для различных методов построения: F – F -метод, A – A -метод, M – M -метод, B – B -метод, L – L -метод, S – S -метод, T (⊕) – «точные» значения доверительного интервала.

На этом рисунке показано, как изменяется верхняя граница разных интервалов (кривые F, A, B, M, L, S) в зависимости от величины доверительной вероятности P . Кроме того, способом, описанным ниже, были определены «точные» значения доверительного интервала при $P=0.8678, 0.9143, 0.9402, 0.9599, 0.9728, 0.9887$, (кружки T с диапазоном возможных ошибок). В дальнейшем, говоря о доверительном интервале, мы будем иметь в виду всю совокупность точек (т.е. кривую) $\gamma_+(P)$, при всех допустимых значениях достоверности P .

Перейдем к описанию того, как строились эти доверительные интервалы

Функция (3.31) неотрицательна, поэтому доверительный интервал методом *стохастической аппроксимации* (*S*-метод) вычислялся по формуле (3.14). Его значения приведены на Рис. 3.3 (кривая **S**). Видно, что он располагается значительно ниже “точного” значения и является худшим из всех приведенных вариантов.

Простыми преобразованиями

$$\tilde{y} = \ln[-\ln(1-y)], \quad \tilde{x}_1 = 1/x_1, \quad \tilde{x}_2 = \ln x_2, \quad \tilde{a}_1 = a_1, \quad \tilde{a}_2 = -a_1 a_3, \quad \tilde{a}_3 = a_1 a_2,$$

можно *линеаризовать* модель и привести ее к виду

$$\tilde{f}(\tilde{\mathbf{x}}, \tilde{\mathbf{a}}) = \tilde{a}_1 \tilde{x}_1 + \tilde{a}_2 \tilde{x}_2 + \tilde{a}_3.$$

При этом, однако, приходится пожертвовать двумя экспериментальными точками

$$y(383, 96) = 1.070 \quad \text{и} \quad y(383, 120) = 1.058,$$

лежащими выше уровня $y=1$ (см. Рис. 3.2). Доверительный интервал, построенный по формуле (3.15), (*L*-метод) лежит уже выше (кривая **L** на Рис. 3.3), но он все еще сильно занижен относительно точек **T**.

Для проведения процедуры *имитационного моделирования* (*M*-метод) мы взяли в качестве исходных значений \mathbf{a}_0 и s_0^2 величины оценок (3.33) и провели $M=5000$ повторений по алгоритму (3.20). В результате был получен доверительный интервал (кривая **M**), который замечательно согласуется с “точными” значениями. Однако, затраты времени на реализацию этого алгоритма составили около 2 часов машинного времени (расчет велся на компьютере с процессором Pentium-100), что представляется неприемлемым для такой простой регрессионной задачи.

B-метод (*bootstrap*) дал очень близкие результаты (кривая **B** совпадает с кривой **M**) с такими же затратами времени на вычисления.

Метод *свободного моделирования*, проведенный с тем же числом $M=5000$ повторений, дал завышенные по сравнению с “точными” результаты (кривая **F** на Рис. 3.3). Это связано, по-видимому, с тем, что распределение (3.24) слишком широко, по сравнению с истинным. Затраты времени на реализацию этого алгоритма – около 20 сек.

Для реализации метода *связанного моделирования* по алгоритму (3.28) были выбраны следующие значения. Число повторений $M=5000$ и число “корзин” $t=28$. Способ разбиения на «корзины» приведен в Табл. 3.1.

Табл. 3.1 Разбиение на «корзины» в *A*-методе

| <i>k</i> | <i>P_k</i> | <i>M_k</i> |
|--------------|----------------------|----------------------|
| 0 | 0.000 | |
| 1 | 0.100 | 500 |
| 2 | 0.200 | 500 |
| ... | ... | ... |
| 8 | 0.800 | 500 |
| 9 | 0.900 | 500 |
| 10 | 0.910 | 50 |
| 11 | 0.920 | 50 |
| ... | ... | ... |
| 18 | 0.980 | 50 |
| 19 | 0.990 | 50 |
| 20 | 0.991 | 5 |
| 21 | 0.992 | 5 |
| ... | ... | ... |
| 27 | 0.998 | 5 |
| 28 | 0.999 | 5 |
| Всего | | 5000 |

Первые 9 интервалов $P_k - P_{k-1}$ для построения «корзин» взяты размером 0.1 (от 0 до 0.9 - объем $M_k=500$); следующие 9 интервалов - размером 0.01 (от 0.9 до 0.99 - объем 50); последние 10 интервалов - размером 0.001 (от 0.99 до 1.0 - объем 5). Результат вычисления доверительных интервалов *A*-методом представлен кривой *A* на Рис. 3.3. Эта кривая несколько отличается от «точных» значений, но это не кажется существенным, т.к. различие не велико и имеет правильное направление (интервал завышен). Затраты времени составили около 25 сек.

На Рис. 3.3 (круги *T*) представлены также «точные» границы доверительного интервала и указано, насколько точны эти «точные» значения (горизонтальные линии).

Эти величины были получены с использованием следующего алгоритма.

- 1) создаем выборку $y=f(x,a)+\varepsilon$, где $\varepsilon \sim N(0, \sigma^2 I)$ при *a* и σ^2 равным истинным значениям (3.32);
- 2) оцениваем параметры α и s^2 ;
- 3) определяем доверительные интервалы $\gamma_+(P)$, тем или иным методом; (3.36)
- 4) сравниваем полученные величины с истинным значением *g* (3.34);
- 5) независимо повторяя *M* раз шаги 1-4, получаем число промахов (случаев, когда $\gamma_+(P) < g$) для каждого значения *P*.

Этот алгоритм был реализован для методов стохастической аппроксимации и связанного моделирования при $M=3093$. Результаты представлены в Табл. 3.2.

В первом столбце находятся значения доверительной вероятности *P*, для которых проводились расчеты. Во втором и третьем столбцах приведены числа промахов при построении интервалов *S*- и *A*- методами. В четвертом столбце находится минимальное, в пятом – среднее, а в шестом максимальное теоретические значения промахов в соответствии с биномиальным законом распределения [12] при уровне значимости 0.05. В седьмом столбце приведена оценка реальной доверительной вероятности P_p , рассчитанной по формуле

$$P_l(P) = 1 - \frac{N_l(P)}{M}$$

где $N_j(P)$ – это число промахов при A -методе (третий столбец Табл. 3.2).

Табл. 3.2 Число промахов при проверке доверительных интервалов, построенных S -методом: и A -методом. Общее число попыток 3093.

| P | Метод | | Должно быть | | | P_1 | $\gamma_+^0(P_1)$ |
|------|-------|-----|-------------|---------|-----|-------|-------------------|
| | S | A | Min | Среднее | Max | | |
| 0.99 | 182 | 35 | 21 | 31 | 40 | 0.989 | 0.995 |
| 0.97 | 279 | 84 | 77 | 93 | 108 | 0.973 | 0.926 |
| 0.95 | 344 | 124 | 134 | 155 | 174 | 0.960 | 0.817 |
| 0.93 | 410 | 185 | 193 | 217 | 239 | 0.940 | 0.720 |
| 0.90 | 470 | 265 | 281 | 309 | 336 | 0.914 | 0.598 |
| 0.85 | 569 | 409 | 431 | 464 | 496 | 0.868 | 0.434 |

По значениям этих вероятностей можно определить «точные» величины границ доверительных интервалов при этих вероятностях (последний столбец в Табл. 3.2). Для этого достаточно вычислить

$$\gamma_+^0(P_1) = \gamma_+(P(P_1)), \quad (3.37)$$

где $\gamma_+(\cdot)$ – это соответствующая граница доверительного интервала, полученного A -методом. Именно эти значения изображены на Рис. 3.3 точками T .

На этом же примере была проведена и численная проверка исходного предположения A -метода о том, что случайная величина $c(\alpha | y)$ (3.25) распределена по закону χ^2 с $p=3$ степенями свободы. Для этого в независимой серии испытаний, совмещенной с процедурой имитационного моделирования (3.20), вычислялось отклонение

$$\delta_1 = \hat{c}(P) - \chi_3^2(P) \quad (3.38)$$

где $\chi_3^2(P)$ - это P -квантиль распределения χ^2 с тремя степенями свободы (1.50), а $\hat{c}(P)$ - выборочный процентиль величины (3.25) в M -методе.

Также была проведена проверка того, хорошо ли техника «приема-отклонения» моделирует распределение χ_3^2 при реализации алгоритма (3.28) A -метода. Для этого вычислялось другое отклонение

$$\delta_2 = \tilde{c}(P) - \chi_3^2(P) \quad (3.39)$$

где $\tilde{c}(P)$ – это выборочный процентиль величины (3.25) при A -методе.

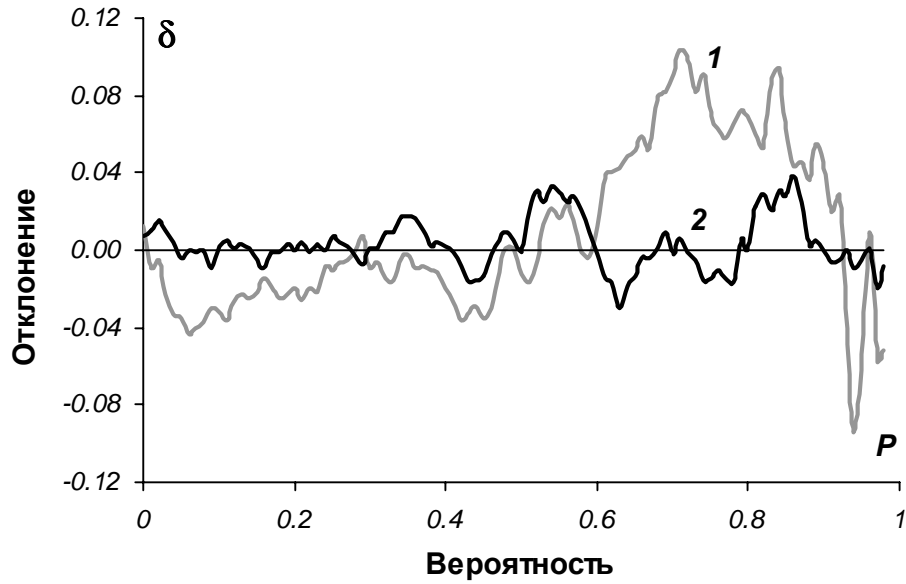


Рис. 3.4 Отклонения выборочных процентилей величины $c(\alpha|y)$ от квантилей распределения хи-квадрат при: 1 - M-методе, 2 -A-методе.

Малость величины δ_1 характеризует правильность гипотезы (3.26), тогда как малость отклонения δ_2 характеризует качество воспроизведения распределения χ^2 при связанном моделировании. На Рис. 3.4 представлены эти отклонения в зависимости от величины доверительной вероятности P .

Из этого рисунка видно, что при $P > 0.75$ размах отклонений δ_1 начинает увеличиваться. С другой стороны, отклонения δ_2 малы, поэтому моделируемое в A-методе распределение близко к χ^2_3 . В тоже время, выборочная оценка дисперсии этого распределения, равна 6.3705, тогда как теоретически должна быть 6. Такое отличие ($M=5000$) существенно на уровне значимости 0.001 и свидетельствует о том, что при реализации идеи «корзин» в A-алгоритме, мы строим более широкое распределение, чем требуется. Этот вывод согласуется с результатами, представленными на Рис. 3.3, где кривая A проходит несколько выше, чем требуется.

В заключение исследуем вопрос о том, как зависит величина доверительного интервала в A-методе (3.28) от числа повторений M .

На Рис. 3.5 представлены результаты 10 попыток построения верхних границ доверительных интервалов в рассматриваемом примере. Левый график а) соответствует $M=5000$, а правый б) $M=1000$. Кроме того, на обоих рисунка жирной кривой показано среднее значение

ние, соответствующее кривой *A* на Рис. 3.3. Видно, что с ростом числа повторений *M* изменчивость интервала уменьшается.

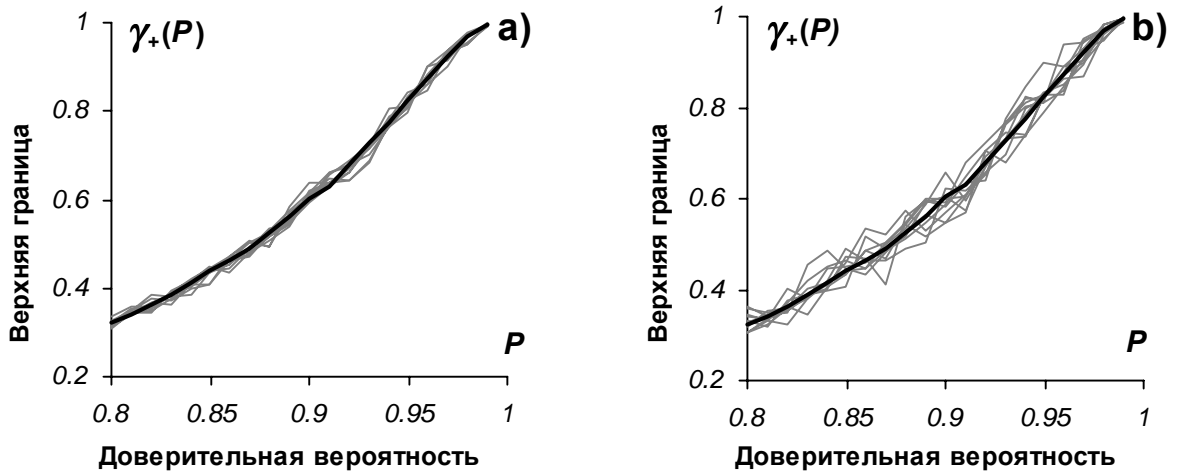


Рис. 3.5 Изменчивость верхних границ доверительных интервалов в *A*-методе при разном числе повторений *M*: а) *M*=5000 и б) *M*=1000.

Эту изменчивость можно охарактеризовать величиной

$$v(M) = \int_0^1 \frac{\sqrt{D[\gamma_+(P)]}}{E[\gamma_+(P)]} dP$$

являющейся средней вариацией величины границы интервала. Она показывает, насколько сильно может измениться найденная граница при реализации другой попытки вычисления той же величины (выборочная изменчивость).

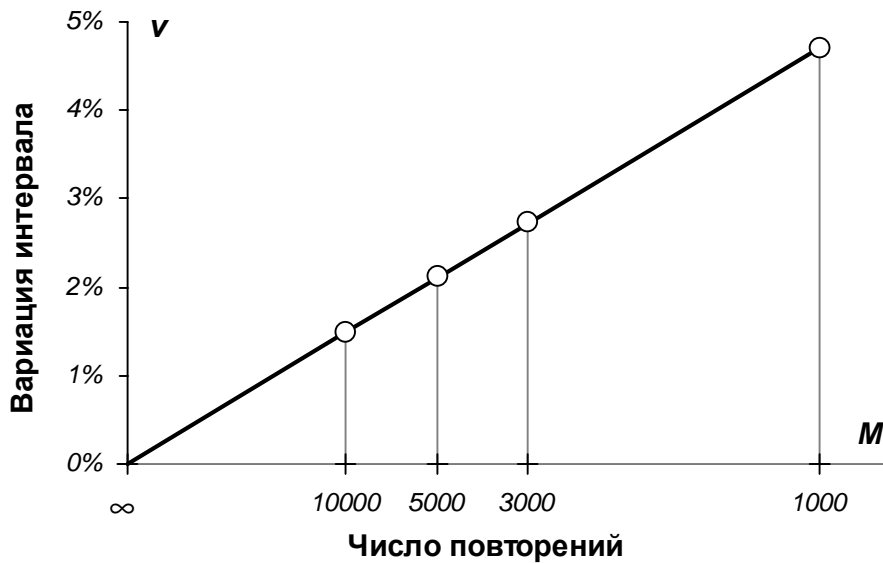


Рис. 3.6 Вариация доверительного интервала в *A*-методе в зависимости от числа повторений *M*.

На Рис. 3.6 приведены значения этой вариации (выраженной в процентах) для рассматриваемого примера при $M=1000, 3000, 5000$ и 10000 . Видно, что эта значения прекрасно описывается стандартной зависимостью [60] от числа повторений

$$v(M) = \frac{Const}{\sqrt{M}}$$

для метода Монте-Карло.

Таким образом, видно, что с «инженерной» точки зрения число повторений $M=1000$ удовлетворительно вычисляет требуемый интервал.

3.4. Коэффициент нелинейности

В примере, рассмотренном в предыдущем разделе, различие в доверительных интервалах, построенных различными методами, было очень значимо. Можно, однако, привести и другие многочисленные примеры, где это различие пренебрежимо мало с «инженерной» точки зрения. В чем же разница между линейной и нелинейной регрессией? Почему «сложная» модель иногда больше похожа на линейную модель, чем «простая» модель? Как можно оценить «степень нелинейности» модели? Ответы на эти вопросы можно дать, используя коэффициент нелинейности [32].

Для определения коэффициента нелинейности вернемся опять к процедуре связанного моделирования, названной в разделе 3.2 A -методом (3.28). На первом шагу этого алгоритма генерируется псевдослучайный нормальный вектор α^* . Его распределение выбрано так, чтобы для линейной регрессионной модели получить правильное распределение оценок параметров – нормальное со средним значением равным a_0 и ковариационной матрицей равной $C = s_0^2 A^{-1}$ (3.24). Иными словами, можно ожидать, что для линейной регрессии P -ая доля сгенерированной выборки $\alpha_1^*, \dots, \alpha_M^*$ будет лежать внутри эллипса, заданного уравнением

$$(\alpha - a_0)^t A^{-1} (\alpha - a_0) = s_0^2 \chi_p^2(P) \quad (3.40)$$

где $\chi_p^2(P)$ это P -квантиль распределения хи-квадрат с p степенями свободы (1.50), а A – это информационная матрица, определенная формулами (1.38) или (2.20) или (2.30) (кривые **2A** и **2B** на Рис. 3.8).

В нелинейном случае этот эллипс заменяется контуром целевой функции

$$Q(\alpha | y_0) - Q(a_0 | y_0) = s_0^2 \chi_p^2(P) \quad (3.41)$$

названный ранее «областью безразличия» целевой функции (3.29) (кривые *IA* и *IB* на Рис. 3.8). Если регрессионная модель линейна, то, очевидно, что эти кривые совпадут.

В методе связанного моделирования необходимо получить выборку оценок параметров, распределенную в соответствии с гипотезой (3.26). Это значит, что для всех значений вероятности P ($0 < P < 1$), доля смоделированных значений α^* , которые попали в область (3.41), ограниченную контуром целевой функции, должна быть равна P .

Для достижения этого выбирается $m+1$ значение вероятности $0 < P_1 < P_2 < \dots < P_m < 1$, которые делят пространство параметров на m отдельных областей («корзин») B_k , и рассчитывается ожидаемое число попаданий M_k (3.27) в каждую из этих корзин. При этом

$$M_1 + M_2 + \dots + M_m = M. \quad (3.42)$$

На каждом шаге моделирования определяется, в какой области B_k оказалась текущая реализация α^* . Если в эту область уже попало ожидаемое число точек M_k , то эта реализация отвергается и делается следующая случайная попытка. Ясно, что чем больше модель отличается от линейной, тем больше реализаций будет отвергнуто, и тем больше попыток в методе Монте-Карло нужно сделать, чтобы достичь требуемой величины удачных реализаций M . Исходя из этого соображения, можно построить критерий, который оценивает нелинейность модели как расстояние между аппроксимирующим эллипсом (3.40) и контуром целевой функции (3.41).

Можно дать точное определение этой величины, используя известные расстояния [14] между двумя распределениями $f(x)$ и $g(x)$

– Кульбака-Лейблера

$$\rho_1(f, g) = \int_{R^n} \ln \frac{f(x)}{g(x)} f(x) dx$$

– Хеллингера

$$\rho_2(f, g) = \int_{R^n} \left(\sqrt{f(x)} - \sqrt{g(x)} \right)^2 dx.$$

Однако, по-видимому, лучше всего использовать расстояние χ^2

$$\rho_3(f, g) = \int_{R^n} \frac{(f(x) - g(x))^2}{g(x)} dx$$

Эту величину можно получить, исследуя число неудачных попыток в A -методе.

Пусть $M_1 + \dots + M_m = M$ – это ожидаемые значения, а $M'_1 + \dots + M'_m = M'$ – это реальные числа попаданий в «корзины» B_k , полученные в A -методе. Очевидно, что $M_k \leq M'_k$. Рассмотрим величину

$$\chi^2 = \sum_{i=1}^m \frac{(M'_k M - M_k M')^2}{M_k M M'} \quad (3.43)$$

которая является стандартной статистикой для сгруппированных наблюдений [12]. Если регрессионная модель линейна, то эта величина имеет хи-квадрат распределение с m степенями свободы.

Коэффициент нелинейности Γ может быть определен как следующим образом

$$\Gamma = \frac{\chi^2}{\chi_m^2(0.95)} \quad (3.44)$$

В этом соотношении числитель – это статистика (3.43), а знаменатель – это 0.95-квантиль хи-квадрат распределения с m степенями свободы (1.50). Для линейной (или похожей на линейную) модели коэффициент Γ должен быть меньше чем 1. Чем дальше этот коэффициент от единицы, тем модель более нелинейная.

Табл. 3.3 Пересчет «точного» в «грубый» коэффициент нелинейности.

| «Точный» коэффициент нелинейности, Γ | «Грубый» коэффициент нелинейности, γ |
|---|---|
| от 0.0 до 1.2 | 1 |
| от 1.2 до 1.5 | Γ |
| от 1.5 до 2.5 | 2 |
| от 2.5 до 3.5 | 3 |
| от 3.5 до 4.5 | 4 |
| более 4.5 | >5 |

Для получения точного значения коэффициента Γ нужно выполнить много попыток в методе Монте-Карло ($M > 1000$). Однако, точное значение коэффициента нелинейности мало интересно. Он должен показывать лишь тенденции в модели, поэтому предлагается заменить «точное» значение Γ на «грубое» значение γ «квантованное» так, как показано в Табл. 3.3. Этот грубый коэффициент γ может быть рассчитан с помощью довольно небольшого числа попыток ($M \approx 500$).

Известны много работ посвященных исследованию коэффициента нелинейности [61, 62, 63, 64], начиная со знаменитой статьи Била [61]. Все они опираются на геометрические

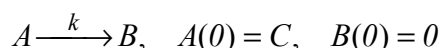
свойства регрессионной кривой и рассчитывают коэффициент нелинейности N_ϕ как меру кривизны этой кривой – т.е. ее отличие от линейной аппроксимации. Они были разработаны для того, чтобы «подправлять» доверительную область оценок параметров

$$Q(\alpha | y_0) - Q(a_0 | y_0) = s_0^2 \left(1 + \frac{p+2}{p} N_\phi \right) F_{p, N_f}(P) \quad (3.45)$$

где $F_{p, N_f}(P)$ – это P -квантиль F -распределения со степенями свободы p и N_f (1.52).

Предлагаемый коэффициент нелинейности (3.44) – это случайная величина, реализации которой могут несколько отличаться, что и компенсируется за счет округления в коэффициенте γ . Такой коэффициент имеет совершенно другой смысл – он используется для того, чтобы определить, можно ли применять упрощенный подход стохастической аппроксимации (S -метод (3.13)) для построения доверительных интервалов, или же необходимы более точные, но медленные методы, такие как связанное моделирование (A -метод (3.28)).

Рассмотрим простой, но поучительный пример использования коэффициента нелинейности. Пусть имеется простейшая реакция первого порядка с кинетической схемой



Соответствующие уравнения для концентраций реагентов имеют следующий вид. Расход A описывается как –

$$A(t) = C e^{-kt} \quad (3.46)$$

а накопление B как –

$$B(t) = C(1 - e^{-kt}). \quad (3.47)$$

На этом примере мы покажем, что с точки зрения регрессионного анализа, первое уравнение является практически линейным относительно неизвестных параметров C и k , тогда как второе – существенно нелинейное.

Рассмотрим модельные данные, представленный на Рис. 3.7.

Каждая кинетическая кривая (A и B) была вычислена при значениях параметров $C=1.0$, $k=0.3$. К полученным значениям концентраций были добавлены случайные нормальные ошибки. Чтобы исключить влияние случайностей на оценки, величины этих ошибок были одинаковыми для каждой кривой. Так получились модельные экспериментальные данные, отмеченные на Рис. 3.7 квадратиками и кружками.

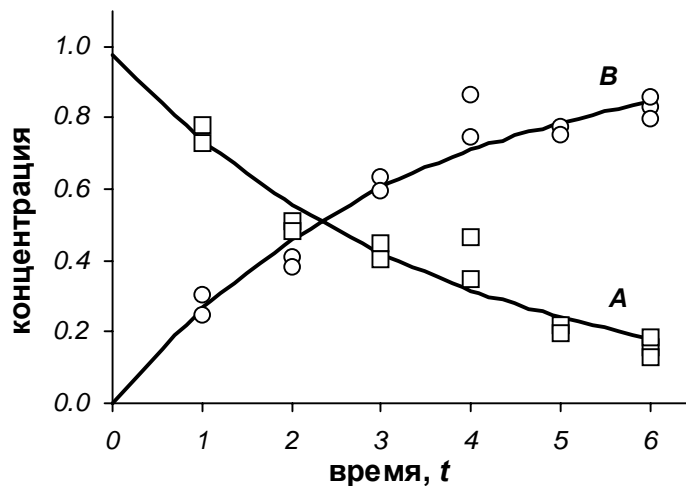


Рис. 3.7 Модельные данные реакции $A \rightarrow B$.

Каждая серия этих данных была обработана независимо и получены оценки параметров, а также коэффициентов нелинейности. Оказалось, что для модели A (измерения расхода реагента A) коэффициенты нелинейности равны $\Gamma=0.77$, $\gamma=1$, тогда как для модели B (измерения накопления реагента B) они составили $\Gamma=3.42$, $\gamma=3$.

Полученные результаты объясняет Рис. 3.8, где изображены контуры целевых функций и аппроксимирующие их эллипсы для моделей A и B. Квадратом ($0A$) и кружком ($0B$) показаны оценки по моделям A и B. Треугольником (0) отмечены точные значения параметров. Видно, что в первом случае кривые $1A$ и $2A$ практически совпадают, тогда как во втором случае соответствующие кривые $1B$ и $2B$ сильно отличаются.

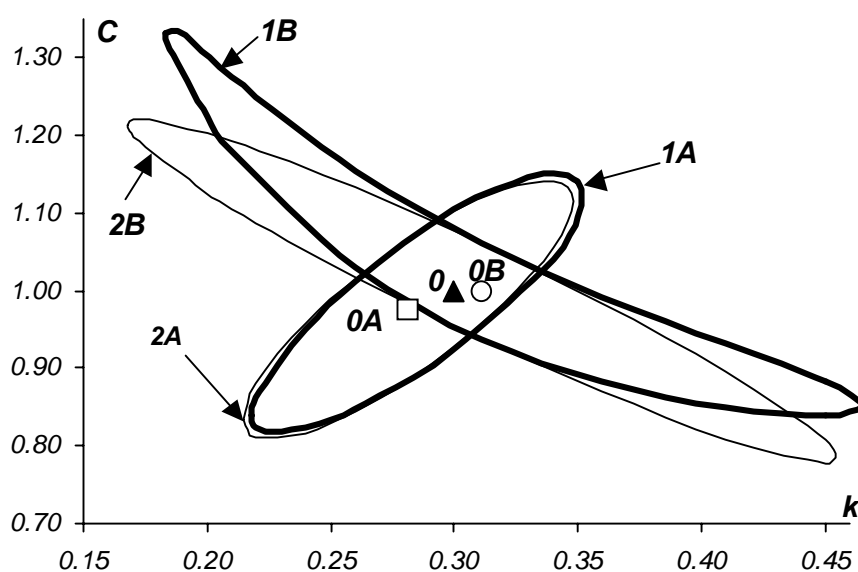


Рис. 3.8 Контуры целевой функции ($1A$ и $1B$) и аппроксимирующие эллипсы ($2A$ и $2B$).

Таким образом, мы видим, что, несмотря на схожесть внешнего вида моделей (3.46) и (3.47), они имеют совершенно разные «степени нелинейности». При этом план измерений был одинаковым в обоих случаях.

Разумеется, имеется различие и в доверительных интервалах для прогноза концентраций.

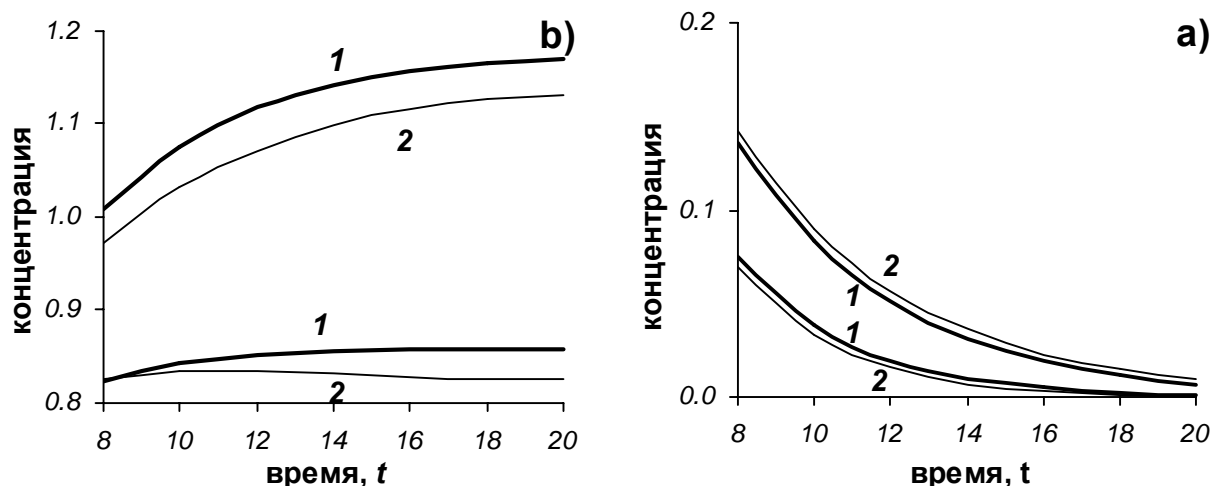


Рис. 3.9 Доверительные коридоры прогноза концентраций по модели А (а) и по модели В (б)

На Рис. 3.9 показаны доверительные коридоры для моделей А (а) и В (б). Толстыми линиями (1) изображены границы интервалов, полученных с помощью связанного моделирования (А-метод), а тонкими линиями (2) показаны границы интервалов, полученных с помощью стохастической аппроксимации (S-метод). Из рисунка видно, что в первом случае различие между кривыми 1 и 2 значительно меньше, чем во втором. Это говорит о том, что первая модель очень близка к линейной и поэтому S-метод хорошо работает. Вторая модель значительно более нелинейна и поэтому S-метод дает плохие результаты.

В рассмотренном примере мы видели, что при одинаковом плане эксперимента коэффициент нелинейности может зависеть от вида модели. Однако он также сильно может зависеть и от плана эксперимента. Следующий пример, продолжающий исследование, начатое в разделе 3.3, иллюстрирует эту идею.

На Рис. 3.10 (графики а и с) представлены два вида данных, смоделированные с помощью одной и той же модели (3.31) с одними и теми же значениями параметров (3.33). Они отличаются только планом эксперимента. Соседние графики б и д показывают соответствующие результаты доверительного прогнозирования с помощью S- и А-методов на одни и те же условия $T=293\text{K}$.

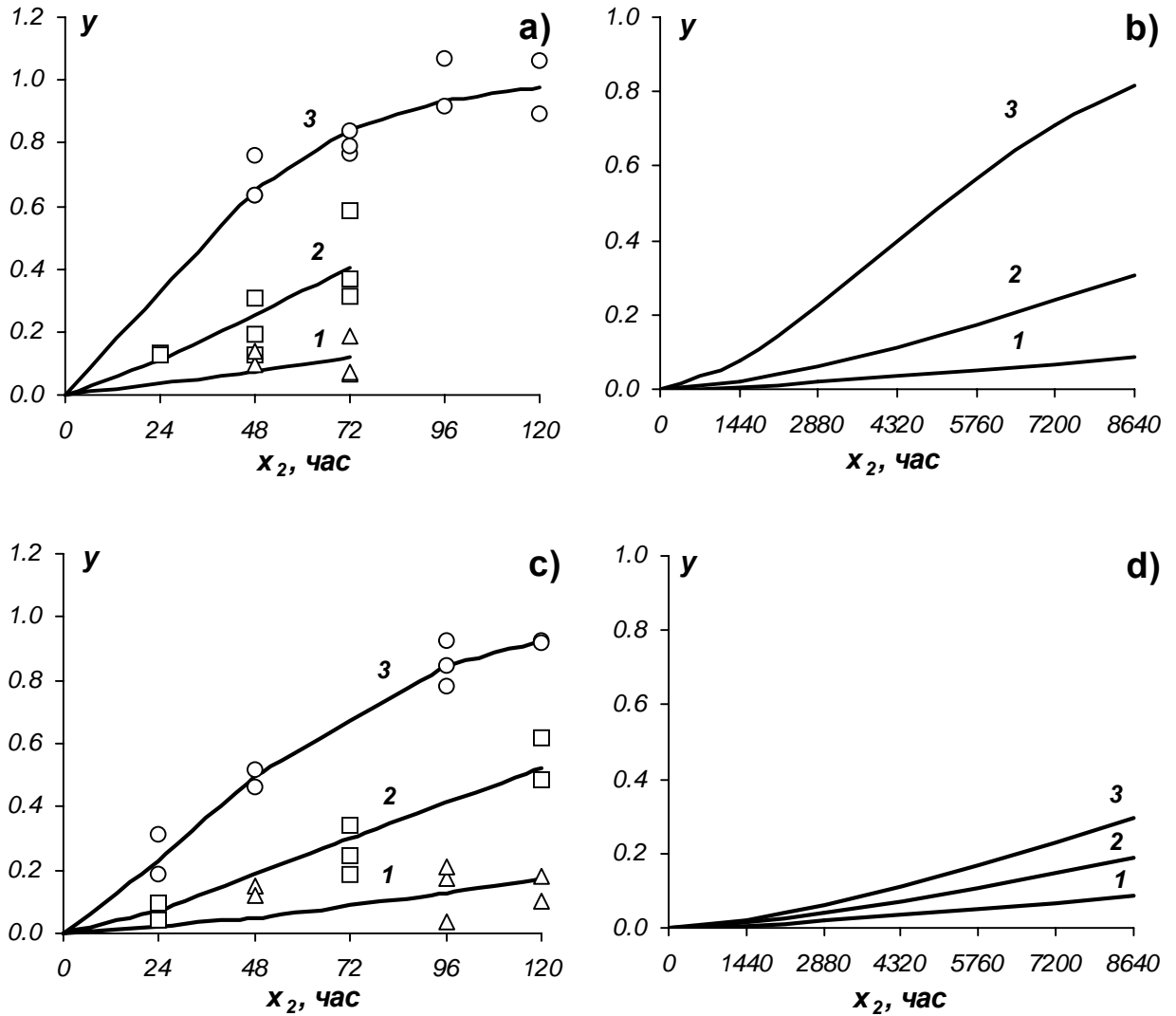


Рис. 3.10 Данные и модель (a, c) 1 (Δ), $T=353\text{K}$; 2 (\square), $T=368\text{K}$; 3 (\circ), $T=383\text{K}$ и соответствующие результаты прогноза (b, d): среднее значение (1) и доверительные границы по S-методу (2) и A-методу (3).

Сравнивая отличия в кривых 2 и 3 на этих рисунках, можно увидеть, что нелинейность этих задач, по-видимому, разная. Вычисленные коэффициенты нелинейности, подтверждает эту идею. Для первого плана эксперимента (графики a и b) рассчитанные значения составили $\Gamma=12.8$, $\gamma>5$. Для второго плана эксперимента (графики c и d) эти величины равны $\Gamma=0.69$, $\gamma=1$.

Коэффициент нелинейности находит широкое применение при решении практических задач. С его помощью можно определять, какой метод доверительного прогнозирования нужно использовать для достижения приемлемых результатов. В этой работе он использовался при анализе стандартной базы данных NIST (раздел 4.3) и в задачах, разобранных в главах 6 и 7.

3.5. Результаты главы 3

В этой главе рассматривались некоторые важные особенности, которые отличают нелинейное и линейное моделирование. Основное назначение нелинейной, содержательной регрессии – это прогнозирование, поэтому главной проблемой, является построение правильных доверительных интервалов при экстраполяции значений отклика. Эта задача изучена пока очень плохо – слишком значительные аналитические проблемы возникают перед исследователем, пытающимся предложить простые алгоритмы для ее решения. Применение элементарных формул, справедливых для линейных моделей часто приводит к большим ошибкам. Другие подходы, основанные на методах прямого имитационного моделирования, требуют значительных затрат времени.

В этой главе был представлен новый способ доверительного оценивания, названный “*связанное моделирование*”, который позволяет быстро строить правильные доверительные интервалы для нелинейных задач. Этот подход относится к классу методов имитационного моделирования (Монте-Карло, bootstrap, jack-knife). Его отличие состоит в том, что моделируются не исходные данные, а оценки параметров. При этом достигается та же точность, что и в традиционном статистическом моделировании, но примерно в 1000 раз быстрее.

Для иллюстрации этого алгоритма и для проверки его точности был рассмотрен модельный пример, приведенный в третьем разделе. Показано, что предлагаемый метод отличается высокой точностью даже для сильно нелинейных моделей. Кроме того, продемонстрировано, что простейшие приемы, основанные на стохастической аппроксимации, приводят к грубым ошибкам.

Практика показывает, что различие в доверительных интервалах, построенных для нелинейной модели различными методами, может быть очень велико, а может быть и пренебрежимо мало с «инженерной» точки зрения. В последнем, четвертом разделе этой главы было дано определение *коэффициента нелинейности* γ , который предлагается использовать для принятия решения о том, какой метод можно использовать в данной задаче.

Различные коэффициенты нелинейности уже давно используются в статистической литературе. Все они опираются на геометрические свойства регрессионной функции и рассчитываются как мера кривизны этой кривой. Предлагаемый коэффициент нелинейности γ – это случайная величина. Такой коэффициент имеет совершенно другой смысл – он используется для того, чтобы определить, можно ли применять упрощенный подход стохас-

тической аппроксимации для построения доверительных интервалов, или же необходимы более точные, но медленные методы, такие как имитационное моделирование.

Показано, что коэффициент нелинейности γ зависит как от вида модели, так и от плана эксперимента. Тем самым доказано, что «нелинейность» является свойством не модели, а всей регрессионной задачи в целом. Поэтому степень нелинейности можно изменить, подбирая соответствующий план эксперимента.

Вычислительные аспекты нелинейного регрессионного анализа

Для того чтобы теоретические идеи нашли практическое воплощение при решении реальных задач, нужно разработать все необходимые вычислительные алгоритмы и, на их базе, создать программу – инструмент для исследователя. В этой части работы рассматриваются вычислительные и программные аспекты нелинейного регрессионного анализа.

Важно отметить, что некоторые алгоритмы и программы, используемые в этой части, были разработаны другими авторами. В частности, алгоритм минимизации, основанный на обращении матричной экспоненты, был предложен Б.В. Павловым и А.Я. Повзнером [74]. Кроме того, большой вклад в разработку некоторых алгоритмов и написание программ внесла О.Е. Родионова [32, 128], являющаяся полноправным соавтором программного обеспечения – системы Fitter.

В этой части работы всего две больших главы. Глава 4 посвящена алгоритмам, используемым в регрессионном анализе. Это, прежде всего, методы минимизации целевой функции. Они достаточно хорошо известны [3], однако в применении их к физико-химическим данным имеется одна особенность. Дело в том, что физико-химический эксперимент – это вещь сложная и дорогостоящая, и, подчас, имеются принципиальные ограничения, не позволяющие провести измерения в той области, которая была бы желательной с точки зрения планирования эксперимента. Поэтому ключевым моментом, определяющим успех обработки данных, является точность вычисления модели и особенно ее производных. В этой главе излагаются оригинальные методы, позволяющие решить одновременно две важные задачи – упростить пользователю задание моделей и добиться высокой точности вычислений. Кроме того, в этой главе производится сравнение разработанных алгоритмов с другими, известными программами нелинейного регрессионного оценивания. В заключение, излагаются методы борьбы с главной проблемой в анализе данных – мультиколлинеарностью.

Следующая глава 5 посвящена описанию компьютерной системы Fitter – основного инструмента для практического анализа данных. Подробно рассказывается о возможностях, особенностях и ограничениях этой программы. По сути, эта глава является развернутой инструкцией для пользователя этого программного продукта.

4. Алгоритмы

Теоретически обосновать возможность применения того или иного метода для решения нелинейных регрессионных задач – это лишь половина, а может быть и меньшая часть работы. Не менее важно разработать алгоритмы для практического воплощения этих методов, и, в конце концов, создать удобный инструмент, с помощью которого даже мало искушенный в математических тонкостях пользователь смог бы легко решить регрессионные задачи. В этой главе рассматриваются основные алгоритмы, применяемые в нелинейном регрессионном анализе.

По-видимому, наиболее важной задачей является выбор алгоритма для минимизации целевой функции. Эта проблема исследуется уже давно и достигнуты большие успехи в ее решении [3]. Предпочтение отдается градиентным методам поиска, которые хорошо зарекомендовали себя при решении плохо обусловленных задач, которые часто встречаются при анализе физико-химических данных, а точнее методу матричной экспоненты [74], разработанному учеными ИХФ РАН.

Минимизация тесно связана с проблемой вычисления модели и ее производных по параметрам. Для ее решения был разработан оригинальный алгоритм, позволяющий интерпретировать модели, задаваемые пользователем в естественной математической форме записи – как систему явных, неявных и дифференциальных уравнений. Другой оригинальный алгоритм, описанный ниже, находит и вычисляет все необходимые производные в аналитической форме. Это позволяет использовать очень сложные модели и находить их производные с высокой точностью, необходимой при поиске оценок.

Качество используемых алгоритмов, их эффективность можно сравнить, только сопоставляя их с другими аналогичными программами. При этом нужно использовать стандартный, общепризнанный набор тестовых данных, например базу NIST [81]. Результаты такого тестирования представленные в этой главе. Они подтверждают правильность использованных подходов.

Для задач анализа физико-химических данных характерно явление мультиколлинеарности. Подходы к решению этой проблемы исследуются в конце данной главы.

4.1. Минимизация целевой функции

Ранее мы видели (см. раздел 1.2), что поиск параметров \mathbf{a} сводится к проблеме минимизации некоторой целевой функции $Q(\mathbf{a})$ –

$$\hat{\mathbf{a}} = \operatorname{argmin} Q(\mathbf{a}), \quad (4.1)$$

где функция $Q(\mathbf{a})$ может быть и суммой квадратов отклонений (1.22) и логарифмом функции правдоподобия (1.28) и байесовской функцией в форме (2.10) или (2.11). Для решения задачи (1.34) разработано много различных методов [1, 3, 66], хотя до сих пор не найден метод, который оказался бы лучшим во всех отношениях. В нелинейном регрессионном анализе задача минимизации характеризуется двумя особенностями. Во-первых, целевая функция является сильно нелинейной, хотя и гладкой. Во-вторых, число неизвестных параметров относительно невелико – не более 20. Эти обстоятельства приводят к выбору *итерационной процедуры* для поиска минимума целевой функции.

В этой процедуре строится последовательность точек

$$\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots \rightarrow \hat{\mathbf{a}}, \quad (4.2)$$

которая сходится к точек минимума функции $Q(\mathbf{a})$. Первая точка в этой последовательности называется *начальным значением* и должна задаваться. Вычисление каждой следующей точки \mathbf{a}_{i+1} в последовательности (4.2) называется *i-ой итерацией*, а разность

$$\mathbf{d}_i = \mathbf{a}_{i+1} - \mathbf{a}_i, \quad (4.3)$$

– *i-м шагом*. Итерация называется *допустимой*, если она приводит к уменьшению целевой функции, т.е.

$$Q_{i+1} < Q_i, \quad (4.4)$$

где

$$Q_i = Q(\mathbf{a}_i), \quad i=1,2,\dots \quad (4.5)$$

Каждая итерация выполняется по формуле

$$\mathbf{a}_{i+1} = \mathbf{a}_i + h_i \mathbf{n}_i, \quad (4.6)$$

где вектор \mathbf{n}_i – называется *направлением*, а скаляр h_i – *размером* шага. Различные методы отличаются способами выбора этих величин. В частности, существует класс методов, называемых *градиентными*, в которых итерации осуществляется по формуле

$$\mathbf{a}_{i+1} = \mathbf{a}_i - h_i \mathbf{R}_i \mathbf{v}_i, \quad (4.7)$$

где \mathbf{R}_i – *положительно определенная* матрица, а \mathbf{v}_i – это *градиент* целевой функции, т.е. вектор, у которого компонента с номером α равна

$$v_{i,\alpha} = \frac{\partial Q(\mathbf{a}_i)}{\partial a_\alpha}, \quad \alpha = 1, \dots, p, \quad (4.8)$$

Разные градиентные методы отличаются способом выбора матрицы \mathbf{R}_i и скаляра h_i .

Метод *наискорейшего спуска* использует $\mathbf{R}_i = \mathbf{I}$. Этот метод слишком медленный и не эффективен для сильно нелинейных задач.

В *методе Ньютона* используются $h_i = 1$ и $\mathbf{R}_i = \mathbf{A}_i^{-1}$, где \mathbf{A}_i – это матрица Гессе целевой функции (1.37).

$$A_{\alpha\beta} = \frac{1}{2} \frac{\partial^2 Q(\mathbf{a})}{\partial a_\alpha \partial a_\beta}, \quad \alpha, \beta = 1, \dots, p \quad (4.9)$$

Известно [3], что если целевая функция квадратичная (т.е. регрессия линейна по параметрам), то матрица \mathbf{A}_i является положительно определенной и метод Ньютона сходится за один шаг. В общем случае (нелинейная регрессия) этот метод неэффективен, т.к. вдали от точки минимума матрица Гессе может оказаться не положительно определенной и метод разойдется. Кроме того, расчет вторых производных является очень трудоемким процессом.

Избежать вычисления вторых производных можно в *методе Ньютона–Гаусса*, где вторые производные от уравнения модели просто опускаются. При этом матрица \mathbf{A} вычисляется по формулам (1.38) или (2.20) или (2.30), в которых используются только первые производные от модели. Это метод можно рассматривать как решение последовательности линейных регрессионных задач, получаемых из исходной модели линейной аппроксимацией в точке \mathbf{a}_i

$$f(\mathbf{x}, \mathbf{a}) = f(\mathbf{x}, \mathbf{a}_i) + \mathbf{g}(\mathbf{a}_i)(\mathbf{a} - \mathbf{a}_i). \quad (4.10)$$

где вектор $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_p(\mathbf{x}))$ производных от модели по параметрам состоит из элементов

$$g_\alpha = \frac{\partial f(\mathbf{x}, \mathbf{a}_i)}{\partial a_\alpha}, \quad \alpha = 1, \dots, p. \quad (4.11)$$

К достоинствам метода Ньютона–Гаусса относится не только простота вычислений, но и то, что матрица \mathbf{A} обладает лучшими свойствами – во многих случаях она оказывается положительно определенной там, где матрица Гессе не удовлетворяет этому условию.

В методе Ньютона–Гаусса на каждом шагу процедуры нужно решать линейные уравнения

$$\mathbf{A}_i \mathbf{d}_i = -\mathbf{v}_i, \quad (4.12)$$

из которых определяется шаг \mathbf{d}_i . Для решения системы (4.12) необходимо обратить матрицу \mathbf{A}_i , которая вдали от минимума целевой функции часто является плохо обусловленной. Тогда вместо матрицы

$$\mathbf{R}_i = \mathbf{A}_i^{-1}, \quad (4.13)$$

которая не существует или сильно не устойчива, используется псевдообратная матрица \mathbf{A}_i^+ . Примером такого подхода является *метод Марквардта* [67, 68, 69]. Его основная идея состоит в том, что матрица \mathbf{R}_i выбирается в виде

$$\mathbf{R}_i = (\mathbf{A}_i + \lambda \mathbf{P})^{-1} \quad (4.14)$$

где \mathbf{P} – это некоторая положительно определенная матрица, а λ – достаточно большое число. В этом случае матрица $\mathbf{A}_i + \lambda \mathbf{P}$ всегда будет положительно определенной, не зависимо от того, какова матрица \mathbf{A}_i . Существует много способов выбора матрицы \mathbf{P} и регуляризатора λ в методе Марквардта. Например, часто выбирают матрицу \mathbf{P} диагональной с элементами, равными диагональным элементам матрицы \mathbf{A}_i . [3, 19]. Существуют и другие методы псевдообращения – метод главных компонент [19, 108, 123, 127], шаговая регрессия [70], Хаусхолдера [71, 72], Грамма-Шмитда [73].

Мы отдаем предпочтение алгоритму *Павлова-Повзнера* [74], основанному на вычислении матричной экспоненты. Он обеспечивает высокую устойчивость и, кроме того, позволяет определить разброс собственных значений матрицы, а также завершенность поиска. Главная идея этого подхода следующая. Для обращения матрицы \mathbf{A}_i на каждой итерации (индекс i опущен для простоты) используется следующая рекуррентная формула

$$\mathbf{B}(2\tau) = \mathbf{B}(\tau)[2\mathbf{I} - \mathbf{A}\mathbf{B}(\tau)] \quad (4.15)$$

Легко видеть, что матрица $\mathbf{B}(\tau)$ удовлетворяет следующему матричному дифференциальному уравнению

$$\frac{d\mathbf{B}(\tau)}{d\tau} = \mathbf{I} - \mathbf{A}\mathbf{B}(\tau); \quad \mathbf{B}(0) = \mathbf{0} \quad (4.16)$$

с решением

$$\mathbf{B}(\tau) = \int_0^\tau \exp(-\mathbf{A}s) ds \quad (4.17)$$

В соответствии с формулой (4.17), матрица $\mathbf{B}(\tau) \rightarrow \mathbf{A}^{-1}$ при $\tau \rightarrow \infty$. Если \mathbf{A} – это вырожденная матрица, то матрица $\mathbf{B}(\tau)$ не теряет смысла и представляет псевдообратную матрицу \mathbf{A}^+ .

При использовании этого алгоритма нужно выбрать достаточно большое число “удвоенный” в уравнении (4.15), а также малое значение β для исходной матрицы $\mathbf{B}_1 = \mathbf{B}(\tau) = \beta \mathbf{I}$.

Тогда каждая следующая матрица $\mathbf{B}_{n+1} = \mathbf{B}(2^n \tau)$ рассчитывается по предыдущей $\mathbf{B}_n = \mathbf{B}(2^{n-1} \tau)$ как

$$\mathbf{B}_{n+1} = \mathbf{B}_n [2\mathbf{I} - \mathbf{A}\mathbf{B}_n] \quad (4.18)$$

Выбор параметра β осуществляется по формуле

$$\beta = 0.01 \frac{\|\mathbf{A}\|}{p \operatorname{tr}(\mathbf{A})}$$

где p – это число параметров, т.е. размерность матрицы,

$$\|\mathbf{A}\| = \sum_{\alpha=1}^p \sum_{\beta=1}^p |A_{\alpha\beta}|$$

– это норма матрицы \mathbf{A} , а

$$\operatorname{tr}(\mathbf{A}) = \sum_{\alpha=1}^p A_{\alpha\alpha}$$

– это след матрицы \mathbf{A} .

Здесь число удвоений n играет ту же роль, что и регуляризатор λ в методе Марквардта. Чем больше n , тем ближе \mathbf{B}_n к обратной матрице \mathbf{A}^{-1} .

В ходе рекуррентной процедуры (4.15) легко контролировать величину

$$sp = \operatorname{tr}(\mathbf{I} - \mathbf{A}\mathbf{B}_n) \quad (4.19)$$

являющуюся следом матрицы $\mathbf{I} - \mathbf{A}\mathbf{B}_n$. Это значение меняется от p , при $n=0$, до нуля при $n \rightarrow \infty$ (см. Рис. 1.1). Такое поведение величины sp связано с природой *собственных значений* матрицы \mathbf{A} . Алгоритм (4.18) обращения матрицы ведет себя подобно методу главных компонент [19, 108, 123, 127]. Вычисления начинаются с наибольшего собственного значения, затем рассчитывается следующее по величине значение, затем следующее и, так продолжается до тех пор, пока все собственные значения не будут учтены или процедура остановлена. Когда поиск еще далек от точки минимума, лучше ограничить число собственных значений, учитываемых в процедуре обращения. Это делается посредством ограничения числа удвоений. Если величина sp незначительно меняется в течение n удвоений, это означает, что отношение двух соседних собственных значений больше чем 2^n , так что разумная граница числа допустимых удвоений – это, например, 20 . Чем ближе поиск к завершению, тем меньше величина (4.19). Исходя из этого, можно построить удобный критерий

$$cp = \frac{p - sp}{p} \times 100\% \quad (4.20)$$

который естественно интерпретировать как *завершенность поиска*. Для начальных итераций его значение близко к 0%, а для завершающих около 100%. (См. Рис. 5.8) Это же значение используется как один из признаков сходимости поиска (см. ниже).

Разумеется, метод Павлова-Повзнера можно использовать не только в минимизации, но и для обращения матрицы. Более того, с его помощью можно определить такую полезную характеристику матрицы, как *разброс собственных значений*, называемую также числом обусловленности. По определению эта величина равна

$$N(A) = \log_{10} \frac{\lambda_{\max}}{\lambda_{\min}} \quad (4.21)$$

где λ_{\max} и λ_{\min} – это максимальное и минимальное собственные числа матрицы A . Чем больше величина $N(A)$, тем хуже обусловлена матрица, и тем труднее ее обратить точно. Рассмотрим величину sp (4.19) в зависимости от числа удвоений n . По определению $sp(1)=p$. Пусть $sp(n_1) = p-1$ и $sp(n_2) = 0$. Можно доказать, что

$$N(A) \approx (n_2 - n_1) \log_{10} 2 \quad (4.22)$$

Рассмотрим пример, в котором матрица A имеет вид

$$A = \begin{bmatrix} 70.000 & 28.750 & -13419.204 & -16.970 \\ 28.750 & 19.111 & -4725.558 & -5.620 \\ -13419.204 & -4725.558 & 321406.729 & 4058.444 \\ -16.970 & -5.620 & 4058.444 & 5.637 \end{bmatrix}$$

Ее собственные значения равны

$$\lambda_1 = 0.471 \quad \lambda_2 = 3.983 \quad \lambda_3 = 21.925 \quad \lambda_4 = 3201475.0989,$$

поэтому

$$N(A) = \log_{10} \frac{3201475.0989}{0.471} \approx 7.$$

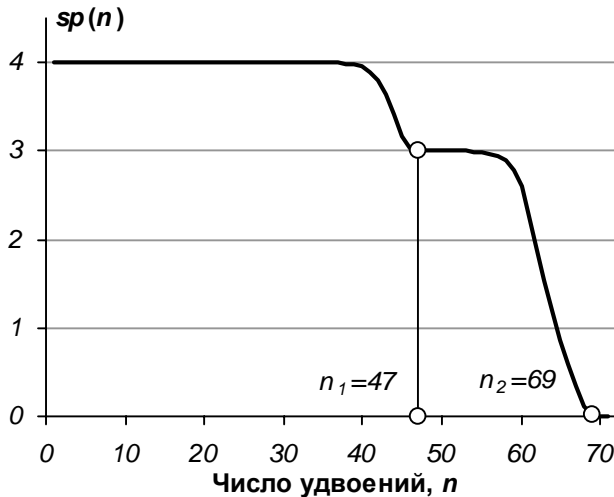


Рис. 4.1 Изменение величины sp в зависимости от числа удвоений n

На Рис. 4.1 показано как при применении метода (4.18) к этой матрице изменялась величина (4.19) в зависимости от числа удвоений n .

Видно, что

$$sp(47) = p-1=3 \text{ и } sp(69) = 0$$

поэтому

$$N(A) = (69 - 47) \log_{10} 2 \approx 7.$$

Таким образом, применяемый алгоритм минимизации состоит из двух вложенных рекуррентных процедур: внешняя

– это градиентный поиск методом Ньютона-Гаусса, а внутренняя – это псевдообращение матрицы методом матричной экспоненты. Условия завершения внутренней процедуры были установлены выше. Теперь сформулируем признаки завершения внешней (градиентной) процедуры. Их два и они сравнивают значения, полученные на двух последних итерациях. Во-первых, значения целевой функции должны отличаться мало, т.е.

$$Q_i - Q_{i+1} < \delta Q_i. \tag{4.23}$$

Во-вторых, нормы матрицы A также должны быть близки

$$| \|A_i\| - \|A_{i+1}\| | < \delta \|A_i\| \tag{4.24}$$

Величина δ , которая участвует в условиях останова (4.23) и (4.24) называется *точностью подгонки (Convergence of Fitting)*. Она задается заранее (см. Рис. 5.6) в пределах от 0.1 до 10^{-10} .

4.2. Вычисление модели и ее производных

Вычисление значений модели, изначально представленной в виде текста, составляет важнейшую часть процедуры нелинейного регрессионного анализа. Именно то, что модель можно вводить в естественной математической форме и является одним из главных преимуществ разработанной программы. В этом разделе рассматривается алгоритм, который превращает текст модели в процедуру ее вычисления и, вслед за тем, в численный результат.

Для иллюстрации этого алгоритма будет использоваться простая задача, представленная на Рис. 4.2.

| Elongation at break | | | | Parameters | |
|---------------------|----------|------------|----------------|------------|----------|
| T | t | ELB | fit | a | 0.971968 |
| 140 | 0 | 4.92 | 5.16027 | b | 4.188305 |
| 140 | 0 | 4.8 | 5.16027 | k | 0.005057 |
| 140 | 0 | 5.2 | 5.16027 | E | 12.53334 |
| 140 | 0 | 5.28 | 5.16027 | | |
| 140 | 0 | 5.32 | 5.16027 | | |
| 140 | 2 | 2.6 | 2.85453 | | |
| 140 | 2 | 2.52 | 2.85453 | | |
| 140 | 2 | 2.6 | 2.85453 | | |
| 140 | 2 | 2.6 | 2.85453 | | |

```
'ELB model
ELB=a+b*exp(-K*t)
K=k*exp(-E*X)
X=1000/(T+273)-Xm
Xm=2.77
a=?
b=?
k=?
E=?
```

Рис. 4.2 Модель относительного удлинения

Она представляет часть примера, разобранный в главе 8. В ней моделируется изменение отклика – относительного удлинения при разрыве (ELB) в зависимости от двух предикторов: температуры (T) и времени (t). Для этого применяется модель (**ELB model**), которая зависит от четырех неизвестных параметров (a, b, k и E). При ее записи использовались три промежуточные переменные (Xm, X и K) и три константы (1000, 273 и 2.77). Подробности синтаксиса моделей приведены в разделе 5.3. На рисунке также приведена таблица значений параметров (**Parameters**) и часть таблицы данных (**Elongation at break**). В последней таблице имеется столбец, с заголовком **Fit**. Он содержит вычисленные значения модели, которые сравниваются с экспериментальными значениями (столбец **ELB**). Далее будет показано, как программа вычисляет эти величины для заданных значений параметров и предикторов.

Анализ модели начинается с выделения в ней идентификаторов, констант и названий стандартных функций (см. Табл. 5.3). В нашем примере есть одна такая функция - экспонента – exp. Затем составляется *матрица связей*, подобная показанной в Табл. 4.1.

В этой таблице по горизонтали перечислены все идентификаторы: сначала неизвестные параметры, затем промежуточные переменные и последним – отклик. По вертикали отложены только промежуточные переменные и отклик. Отличить параметр от любой другой переменной, участвующей в модели, можно по характерному признаку – все параметры входят в модель в виде уравнения с вопросительным знаком в правой части (см. Рис. 4.2). На пересечении строк и столбцов матрицы связей стоят шестнадцатеричные числа, характеризующие явную зависимость величины в строке от величины в столбце – если число не равно нулю, то зависимость есть. Так, например, видно, что величина Xm ни от чего не зависит, а величина K зависит от k, E и X. Смысл отличных от нуля значений будет объяснен ниже.

Табл. 4.1 Матрица связей

| | | параметры | | | промежуточные | | | отклик | |
|---------------|-----|-----------|-----------|-----------|---------------|-----------|-----------|-----------|-----|
| | | a | b | k | E | Xm | X | K | ELB |
| промежуточные | Xm | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | X | 0 | 0 | 0 | 0 | 1b | 0 | 0 | 0 |
| | K | 0 | 0 | 1c | 1d | 0 | 1e | 0 | 0 |
| отклик | ELB | 1f | 20 | 0 | 0 | 0 | 0 | 21 | 0 |

Простой анализ матрицы связей позволяет, в частности, определить какой из идентификаторов является откликом – сумма чисел в его столбце должна быть равно нулю. В нашем примере это, разумеется, ELB. Более тонкий анализ матрицы дает возможность ранжировать все промежуточные переменные, т.е. установить в каком порядке они должны вычисляться. В рассматриваемом примере этот порядок такой. Сначала вычисляется Xm, затем X, за ней K и последним – отклик ELB. Найти правильный порядок вычислений (*иерархию*) – это очень важно, т.к. в сложной модели (см. например, Рис. 9.15) он далеко не очевиден и нужно освободить пользователя от решения этой сложной проблемы. В Табл. 4.1 приведен окончательный вид матрицы связей. Прежде, чем получить его, нужно проделать несколько операций над исходной – более сложной матрицей. Мы, однако, не будем вдаваться в тонкости этого процесса.

После анализа матрицы связей можно создавать *буфер вычислений* – одномерный массив переменной длины, подобный тому, который показан в Табл. 4.2. Этот массив содержит как исходные значения величин, участвующих в модели, так и все промежуточные результаты вычислений. При составлении и использовании буфера необходимо контролировать относительный адрес всех значений, которые в нем содержатся. В таблице этот адрес приведен в шестнадцатеричном формате в столбце, озаглавленном **Адрес**. Рядом указаны численные значения (столбец **Значения**), которые находились в массиве при расчете рассматриваемого примера. Справа и слева расположены комментарии, позволяющие понять устройство этой таблицы.

Табл. 4.2 Структура буфера вычислений

| Солежание | Адрес | Значение | Величина |
|--|-------|-----------|--------------------|
| Стандартные константы | 0000 | 0.0 | ноль |
| | 0001 | 1.0 | единица |
| | 0002 | -1.0 | минус единица |
| | 0003 | 3.1415927 | число π |
| | 0004 | 0.5 | |
| | 0005 | -0.5 | |
| | | | резервные значения |
| | 000f | | |
| Результат | 0010 | -3.76512 | |
| Предикторы | 0011 | 140.0 | T |
| | 0012 | 2.0 | t |
| Параметры | 0013 | 0.971968 | a |
| | 0014 | 4.188305 | b |
| | 0015 | 0.005057 | k |
| | 0016 | 12.53333 | E |
| | 0017 | 2.77 | Xm |
| Промежуточные величины | 0018 | -0.348692 | X |
| | 0019 | 0.399831 | K |
| | 001a | 2.85453 | ELB |
| Частные производные от всех правых частей по всем промежуточным величинам и параметрам | 001b | -1.0 | $d[F(X)]/d[Xm]$ |
| | 001c | 79.0658 | $d[F(K)]/d[k]$ |
| | 001d | 0.139418 | $d[F(K)]/d[E]$ |
| | 001e | -5.01122 | $d[F(K)]/d[X]$ |
| | 001f | 1.0 | $d[F(ELB)]/d[a]$ |
| | 0020 | 0.449481 | $d[F(ELB)]/d[b]$ |
| | 0021 | -3.76512 | $d[F(ELB)]/d[K]$ |
| | 0022 | 0.0 | $D[Xm]/D[a]$ |
| | 0023 | 0.0 | $D[Xm]/D[b]$ |
| | 0024 | 0.0 | $D[Xm]/D[k]$ |
| | 0025 | 0.0 | $D[Xm]/D[E]$ |
| Полные производные от промежуточных величин по параметрам | 0026 | 0.0 | $D[X]/D[a]$ |
| | 0027 | 0.0 | $D[X]/D[b]$ |
| | 0028 | 0.0 | $D[X]/D[k]$ |
| | 0029 | 0.0 | $D[X]/D[E]$ |
| | 002a | 0.0 | $D[K]/D[a]$ |
| | 002b | 0.0 | $D[K]/D[b]$ |
| | 002c | 79.0658 | $D[K]/D[k]$ |
| | 002d | 0.139418 | $D[K]/D[E]$ |
| Полные производные от отклика по параметрам | 002e | 1.0 | $D[ELB]/D[a]$ |
| | 002f | 0.449481 | $D[ELB]/D[b]$ |
| | 0030 | -297.692 | $D[ELB]/D[k]$ |
| Частные константы | 0031 | -0.524927 | $D[ELB]/D[E]$ |
| | 0032 | 1000.0 | |
| | 0033 | 273.0 | |
| | 0034 | 2.77 | |
| | 0035 | -0.799663 | |
| | 0036 | 0.799663 | |
| | 0037 | -0.799663 | |
| Начало вычислений | 0038 | -2.0 | |
| | 0039 | 2.0 | |
| | 003a | 2.0 | |
| | 003b | 2.0 | |
| | 003c | 2.0 | |
| | 003d | -2.0 | |

Первые шестнадцать величин, занимающих ячейки с адресами от 0000 до 000f, это *стандартные константы*: 0, 1, -1 и т.д., которые часто встречаются в моделях. Следующая ячейка с адресом 0010 предназначена для временного хранения каждого *результата вычислений* – отсюда эта величина забирается и помещается по нужному адресу. Эти первые семнадцать элементов буфера вычислений всегда фиксированы, т.е. они одни и те же для любой модели, тогда как последующие элементы уже зависят от ее конкретных особенностей.

Элементы, начинающиеся с адреса 0011, содержат текущие значения *предикторов*. Количество этих элементов зависит от числа предикторов. В рассматриваемом примере их два: t и τ . Значения 140.0 и 2.0 соответствуют приведенным в таблице данных на Рис. 4.2 (выделенная строка). Они помещаются по адресам 0011 и 0012 при составлении буфера. Разумеется, в другой модели может быть большее (или меньшее) количество предикторов. Тогда изменятся и их адреса в буфере.

Сразу же за предикторами размещаются *параметры* модели. В примере их четыре, поэтому они занимают места с адресами от 0013 до 0016. Их значения соответствуют величинам, приведенным в таблице параметров Рис. 4.2. Здесь также может быть иное число параметров, поэтому адреса в буфере зависят от вида модели.

После параметров следуют *промежуточные величины* и *отклик*. Их значения неизвестны заранее, а вычисляются программой, так, как будет объяснено ниже. В примере эти величины занимают в буфере адреса с 0017 по 001a. Порядок следования этих величин определяется иерархией, определенной ранее при анализе матрицы связей.

Затем в буфере идут *частные производные* от правых частей всех промежуточных переменных и отклика по всем промежуточным переменным и параметрам, от которых они зависят. Эта сложная фраза содержит очень простую идею. В процедуре поиска параметров используются производные (4.11) от модели по параметрам. Для их вычисления сначала нужно найти частные производные от всего по всему. То, как это делается, будет объяснено ниже, а пока необходимо просто зарезервировать в буфере нужное число элементов. Очевидно, что частных производных будет ровно столько, сколько имеется ненулевых элементов в матрице связей (см. Табл. 4.1). В нашем примере их семь – начиная от $\partial X / \partial X_t$ и заканчивая $\partial ELB / \partial K$ – и они занимают ячейки с адресами от 001b до 0021. Кстати, именно величины этих адресов и стоят в соответствующих клетках матрицы связей. Здесь есть одна тонкость. Если уравнение переменной имеет вид неявной функции

или дифференциального уравнения, то частные производные нужно брать от правой части соответствующего уравнения, а не от его решения.

За частными производными размещаются *полные производные* по параметрам модели. Сначала идут производные от всех промежуточных величин, а затем от отклика. Число этих ячеек равно –

$$(\text{Число промежуточных величин} + 1) \times \text{Число параметров.}$$

В рассматриваемом примере это $(3+1) \times 4 = 16$ элементов с адресами от 0022 до 0031. Их порядок также иерархический.

Потом в буфер помещаются *частные константы*, т.е. те константы, которые участвуют в модели и не являются *стандартными*. В примере их три – 1000, 273, 2.77 – и они занимают ячейки с адресами от 0032 до 0034. Этим адресом закрывается *постоянная часть* буфера – при вычислении значений *этой* модели в точках *этой* таблицы данных будет меняться не длина этой части буфера, а только содержание его ячеек.

Ячейка стоящая сразу после частных констант является началом *переменной части* буфера, длина которой будет изменяться по ходу вычислений. В примере адрес этой ячейки 0035.

Создав буфер вычислений необходимой длины, мы можем вернуться к анализу модели. Теперь ее нужно переписать, используя вместо содержательных идентификаторов новые *адресные* идентификаторы, построенные из соответствующих адресов в буфере вычислений (См. Рис. 4.3)

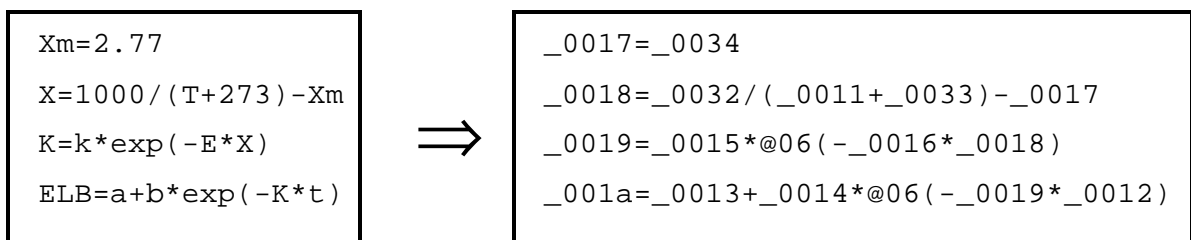


Рис. 4.3 Преобразование модели

Например, идентификатор X_m имеет адрес 0017, поэтому его новый адресный идентификатор – это $_0017$. Одновременно строки в модели выстраиваются в правильном иерархическом порядке. Для новых идентификаторов стандартных функций используется иная схема. Все стандартные функции (См. Табл. 5.3) имеют собственный номер, поэтому новый идентификатор, например, экспоненты – это @06. Символ подчеркивания « $_$ » является признаком переменной, а символ « $@$ » – признаком функции.

Теперь модель подготовлена к *компиляции*, т.е. к созданию кода, вычисляющего значения каждой из строк модели. Рассмотрим, как делается на примере.

Первая строка модели на Рис. 4.3 – это простейшее уравнение, пересылающее значение, содержащееся в буфере по адресу 0034 (константа 2.77), в ячейку буфера с адресом 0017. Оно компилируется в две строки. Сначала содержимое ячейки 0000 (это ноль) складывается с содержимым ячейки 0034 и результат помещается в первую рабочую ячейку 0035. Затем содержимое ячейки 0000 складывается с содержимым рабочей ячейки 0035 и результат помещается по адресу всех окончательных результатов 0010. Для записи этих операций используется кодировка трехадресной машины, причем, номер операции обозначается двухзначным шестнадцатеричным числом, соответствующим номеру в Табл. 5.2 и Табл. 5.3. Так можно получить *предварительный код*, показанный в левой части Табл. 4.3

Табл. 4.3 Предварительный код уравнения $x_m=2.77$

| Оп | Опр. 1 | Опр. 2 | Рез. | Оп. | Операнд 1 | Операнд 2 | Результат |
|----|--------|--------|------|-----|-----------|-----------|-----------|
| 05 | 0000 | 0034 | 0035 | + | 0 | 2.77 | 0+2.77 |
| 05 | 0000 | 0035 | 0010 | + | 0 | 2.77 | 0+2.77 |

Если записать этот предварительный код в одну строку, разбив ее на группы по два символа

05 00 00 00 34 00 35 05 00 00 00 35 00 10,

то окончательный код можно представить в виде символьной строки, состоящей из ASCII символов

⚡ 4 5⚡ 5 ▶

Эта строка может быть расшифрована и опять превращена в предварительный код, с помощью которого и проводятся все необходимые вычисления. В конце вычислений остается только забрать полученный результат из ячейки с адресом 0010 и поместить его в нужный адрес буфера – 0017. Так вычисляется первое уравнение модели.

Вторая строка уже сложнее

$$_0018=_0032/(_0011+_0033)-_0017$$

т.к. в ней имеются скобки и несколько бинарных операций. Она компилируется рекурсивно, превращаясь сначала в две строки –

$$_0035=_0011+_0033$$

$$_0018=_0032/_0035-_0017$$

потом в четыре строки –

$$_0036=_0011+_0033$$

_0035=_0036
 _0037=_0032/_0035
 _0018=_0037-_0018

затем в пять строчек –

_0036=_0011+_0033
 _0035=_0036
 _0037=_0032/_0035
 _0038=_0037-_0017
 _0018=_0038

а уже затем в код, показанный в Табл. 4.4.

Табл. 4.4 Предварительный код для уравнения $x=1000/(T+273)-2.77$

| Оп | Опр. 1 | Опр. 2 | Рез. | Оп. | Операнд 1 | Операнд 2 | Результат |
|----|--------|--------|------|-----|--------------|-------------------|---------------------|
| 05 | 0011 | 0033 | 0036 | + | T | 273 | T+273 |
| 05 | 0000 | 0036 | 0010 | + | 0 | T+273 | 0+T+273 |
| 05 | 0000 | 0036 | 0035 | + | 0 | T+273 | 0+T+273 |
| 03 | 0032 | 0035 | 0037 | / | 1000 | T+273 | 1000/(T+273) |
| 04 | 0037 | 0017 | 0038 | - | 1000/(T+273) | 2.77 | 1000/(T+273)-2.77 |
| 05 | 0000 | 0038 | 0010 | + | 0 | 1000/(T+273)-2.77 | 0+1000/(T+273)-2.77 |

Окончательный код имеет вид

⌘ ◀ 3 6⌘ 6 ▶⌘ 6 5 2 5 7⌘ 7 ⌘ 8⌘ 8 ▶

Аналогично создается и код для вычисления двух оставшихся строк в модели. Он приведен в Табл. 4.5 и Табл. 4.6. В этих уравнениях участвует стандартная функция (exp), которая кодируется операцией с номером 06.

Табл. 4.5 Предварительный код для уравнения $K=k*\exp(-E*X)$

| Оп | Опр. 1 | Опр. 2 | Рез. | Оп. | Операнд 1 | Операнд 2 | Результат |
|----|--------|--------|------|-----|-----------|-------------|---------------|
| 02 | 0016 | 0018 | 0036 | * | E | X | E*X |
| 04 | 0000 | 0036 | 0037 | - | 0 | E*X | 0-E*X |
| 05 | 0000 | 0037 | 0010 | + | 0 | -E*X | 0-E*X |
| 05 | 0000 | 0037 | 0035 | + | 0 | -E*X | 0-E*X |
| 06 | 0035 | 0000 | 0038 | exp | -E*X | 0 | exp(-E*X) |
| 02 | 0015 | 0038 | 0039 | * | k | exp(-E*X) | k*exp(-E*X) |
| 05 | 0000 | 0039 | 0010 | + | 0 | k*exp(-E*X) | 0+k*exp(-E*X) |

Табл. 4.6 Предварительный код для уравнения $ELB=a+b*\exp(-K*t)$

| Оп | Опр. 1 | Опр. 2 | Рез. | Оп. | Операнд 1 | Операнд 2 | Результат |
|----|--------|--------|------|-----|-----------|--------------|-----------------|
| 02 | 0019 | 0012 | 0036 | * | K | t | K*t |
| 04 | 0000 | 0036 | 0037 | - | 0 | K*t | 0-K*t |
| 05 | 0000 | 0037 | 0010 | + | 0 | -K*t | 0-k*T |
| 05 | 0000 | 0037 | 0035 | + | 0 | -K*t | 0-k*T |
| 06 | 0035 | 0000 | 0038 | exp | -K*t | 0 | exp(-K*t) |
| 02 | 0014 | 0038 | 0039 | * | b | exp(-K*t) | b*exp(-K*t) |
| 05 | 0013 | 0039 | 003a | + | a | b*exp(K*t) | a+b*exp(-K*t) |
| 05 | 0000 | 003a | 0010 | + | 0 | a+b*exp(K*t) | 0+a+b*exp(-K*t) |

Если строка модели имеет вид *неявного* или *дифференциального* уравнения, то сначала из нее извлекается часть, заключенная между знаком равенства « \Leftarrow » и знаком точка с запятой «;», которая служит разделителем между формулой и граничным (начальным) условием (см. раздел 5.3). После этого компилируется только эта выделенная часть. Тем самым создается код, который используется для вычисления правой части уравнения. Для вычисления же самой переменной применяются традиционные методы решения таких уравнений. Так неявное уравнение решается методом половинного деления [75], а дифференциальное – методом Рунге-Кутты [76]. При этом точность численного решения определяется величиной *точности вычислений ε* , задаваемой пользователем (см. раздел 5.7).

Регрессионный анализ нелинейных моделей требует высокой точности расчетов. Причем точность важна не сама по себе (практические вычисления не требуют более четырех значащих цифр), но она необходима как оружие в борьбе с мультиколлинеарностью (см. подробности в разделе 4.4). Например, для обращения матрицы с разбросом собственных значений (4.21) равным 8-ми десятичным порядкам (от 10^{-4} до 10^{+4}), необходимо осуществлять промежуточные вычисления с точностью около 10^{-10} . Если же для обработки необходимо использовать неявное или дифференциальное уравнения, промежуточные вычисления должны быть сделаны еще более точно для уменьшения накопления ошибок.

Точность вычислений зависит, прежде всего, от используемого метода расчета производных от модели (4.11) по параметрам. Существует три возможных решения.

Метод Пауэла–Флэтчера [77, 78, 79], который вообще не требует расчета производных, оказался довольно слабым и ненадежным [3] подходом. Он не может быть применен для сложных моделей и, кроме того, требует дополнительных объемистых вычислений [37] по нахождению статистических характеристик, описывающих качество оценок (см. раздел 1.3).

Второй подход использует численное вычисление производных [80], применяя для этого конечно-разностные выражения типа

$$\frac{\partial f}{\partial a} = \frac{f(a+h) - f(a)}{h}.$$

Такое решение является самым популярным. Во многих широко-известных программах для анализа нелинейной регрессии используется подобный подход (см. раздел 4.3). Однако разностные методы сильно снижают точность вычислений, и, как правило, не дают больше, чем 6 или 7 правильных знаков в результате. Поэтому становится невозможным

использование дифференциальных или неявных уравнений в нелинейном регрессионном анализе из-за полной потери точности.

И, наконец, последнее решение – это использование прямых аналитических вычислений производных. Некоторые программы предлагают пользователям самим вводить аналитические формулы для вычисления производных [90] – что является, по-видимому, совершенно неправильным. Для вычислений имеется компьютер и это его черная работа. Поэтому был разработан специальный алгоритм для нахождения аналитических производных. Когда пользователь вводит формулу, например $y = \exp(-a \cdot t)$, то программа сама находит нужные производные, т.е. определяет, что $dy/da = -t \cdot \exp(-a \cdot t)$. В сочетании с компилятором, который создает объектный код для расчета значений модели по ее аналитической формуле, это позволяет провести расчеты с высокой точностью даже для систем дифференциальных уравнений.

Все необходимые производные определяются в соответствии со схемой «от всего по всем», описанной выше, причем это делается для уравнений выраженных в адресных идентификаторах (см. Рис. 4.3). Например, для рассматриваемого примера по этому алгоритму были определены производные представленные в Табл. 4.7. Результаты вычислений этих выражений должны помещаться по адресам от 001b до 0021 в буфере вычислений (см. Табл. 4.2). Их можно получить, откомпилировав выражения для производных так же, как это было сделано для исходной модели.

Табл. 4.7 Вычисление частных производных «от всех по всему»

| производная | выражение |
|-----------------------------|--|
| $\partial X / \partial X_m$ | _001b = -(_0001) |
| $\partial K / \partial k$ | _001c = (@06(-_0016*_0018)) |
| $\partial K / \partial E$ | _001d = (_0015*@06(-_0016*_0018)*(-((_0018)))) |
| $\partial K / \partial X$ | _001e = (_0015*@06(-_0016*_0018)*(-((_0016)))) |
| $\partial ELB / \partial a$ | _001f = _0001 |
| $\partial ELB / \partial b$ | _0020 = (@06(-_0019*_0012)) |
| $\partial ELB / \partial K$ | _0021 = (_0014*@06(-_0019*_0012)*(-((_0012)))) |

Теперь можно вычислить и полные производные от всех промежуточных переменных и отклика по параметрам. Для этого используется матрица связей (Табл. 4.1) и правило дифференцирования сложных функций

$$\frac{dI_i}{da_\alpha} = \frac{\partial I_i}{\partial a_\alpha} + \sum_{k < i} \frac{\partial I_i}{\partial I_k} \frac{dI_k}{da_\alpha} \quad (4.25)$$

Здесь a_α – это параметры, а I_i – это переменные модели (промежуточные или отклик). Так как все расчеты по формуле (4.25) проводятся в иерархическом порядке, то все нужные полные производные оказываются уже вычисленными. В нашем примере это выглядит так

Табл. 4.8 Вычисление полных производных по параметрам

| производная | выражение |
|--|-------------------|
| $dX_m/da = 0$ | _0022=_0000 |
| $dX_m/db = 0$ | _0023=_0000 |
| $dX_m/dk = 0$ | _0024=_0000 |
| $dX_m/dE = 0$ | _0025=_0000 |
| $dX/da = 0$ | _0026=_0000 |
| $dX/db = 0$ | _0027=_0000 |
| $dX/dk = 0$ | _0028=_0000 |
| $dX/dE = 0$ | _0029=_0000 |
| $dK/da = 0$ | _002a=_0000 |
| $dK/db = 0$ | _002b=_0000 |
| $dK/dk = \partial K/\partial k$ | _002c=_001c |
| $dK/dE = \partial K/\partial E$ | _002d=_001d |
| $dELB/da = \partial ELB/\partial a$ | _002e=_001f |
| $dELB/db = \partial ELB/\partial b$ | _002f=_0020 |
| $dELB/dk = (\partial ELB/\partial K)(dK/dk)$ | _0030=_0021*_002c |
| $dELB/dE = (\partial ELB/\partial K)(dK/dE)$ | _0031=_0021*_002d |

Вычисление производных от переменных, заданных неявными функциями, проводится аналогично, но с использованием формулы

$$\frac{dI_i}{da_\alpha} = - \left(\frac{\partial F_i}{\partial a_\alpha} + \sum_{k < i} \frac{\partial F_i}{\partial I_k} \frac{dI_k}{da_\alpha} \right) / \frac{\partial F_i}{\partial I_i} \quad (4.26)$$

где F_i – это функция, стоящая в правой части неявного уравнения для переменной I_i (см. раздел 5.3), т.е. в уравнении

$$0 = F_i(\mathbf{x}, \mathbf{a}, \mathbf{I}); L < I_i < R.$$

Дифференцирование по параметрам дифференциального уравнения

$$\frac{dI_i}{dt} = f(y, t, \mathbf{x}, \mathbf{a}, \mathbf{I}); I_i(t_0) = y_0$$

приводит к появлению нового дифференциального уравнения

$$\frac{dg_{i\alpha}}{dt} = \frac{\partial f}{\partial a_\alpha} + \sum_{k < i} \frac{\partial f}{\partial I_k} \frac{dI_k}{da_\alpha}; g_{i\alpha}(t_0) = \frac{dy_0(t_0)}{da_\alpha}$$

где

$$g_{i\alpha} = \frac{dI_i}{da_\alpha}$$

– это производная от решения по параметру a_α . Особые уравнения возникают при дифференцировании по параметру, от которого зависит начальная точка – t_0

$$\frac{dg_{i\alpha}}{dt} = \frac{\partial f}{\partial a_\alpha} + \sum_{k < i} \frac{\partial f}{\partial I_k} \frac{dI_k}{da_\alpha}; g_{i\alpha}(t_0) = -f(t_0) \frac{dy_0(t_0)}{da_\alpha}.$$

Таким образом, получаются процедуры, вычисляющие и заполняющие буфер вычислений значениями всех промежуточных и окончательных переменных модели, а также их производными по параметрам. Теперь можно менять те значения элементов буфера, которые отвечают за разные точки плана экспериментальных данных (в примере – это адреса 0011 и 0012) и получать в том же буфере соответствующие им значения модели (адрес 001a) и ее производных (адреса с 002e по 0031). Кроме того, можно менять и значения параметров (адреса с 0013 по 0016) и проводить подгонку модели к имеющимся данным.

4.3. Тестирование программ

Окончательный результат работы программы зависит не только от удачного выбора алгоритма, но и от того, как хорошо он реализован практически. Для проверки устойчивости, надежности и точности программы обычно используются стандартные тестовые задачи, для которых правильный ответ известен заранее. Подобный набор данных [81] был разработан Национальным Институтом Стандартов и Технологий (NIST, США). В этом наборе имеются 27 модельных и реальных нелинейных регрессионных задач различного уровня сложности (см. Табл. 4.9). Подробности всех расчетов можно найти в [126], где приведены файлы *NIST1.XLS–NIST3.XLS*, содержащий исходные данные и порядок расчетов.

В базе данных для каждой модели приведены исходные данные вместе с сертифицированными значениями оценок параметров и их стандартных отклонений. Сертифицированные значения – это наилучшие из доступных решений, проверенные NIST. Они были получены на 128-битовой машине и были подтверждены, по крайней мере, двумя различными алгоритмами и пакетами программ, использующими аналитическое вычисление производных. Все сертифицированные результаты (оценки параметров, среднеквадратичные отклонения параметров, остаточная сумма квадратов, среднеквадратичная ошибка приближения) даны с точностью 11 десятичных разрядов для каждого набора данных.

В Табл. 4.9 представлены следующие величины.

Табл. 4.9 Тестовые задачи NIST

| Название задачи | Сложность | | Тип модели | Число параметров | γ | N(A) | Верные цифры | |
|-----------------|-----------|-----|------------------|------------------|----------|-------------------|--------------|-----|
| | NIST | Наш | | | | | параметры | СКО |
| Misra1a | 1 | 1 | экспоненциальная | 2 | 1 | 13 | 11 | 11 |
| Chwirut2 | 1 | 1 | экспоненциальная | 3 | 1 | 5 | 11 | 11 |
| Chwirut1 | 1 | 2 | экспоненциальная | 3 | 1 | 5 | 11 | 11 |
| Lanczos3 | 1 | 1 | экспоненциальная | 6 | >5 | 8 | 11 | 9 |
| Gauss1 | 1 | 1 | экспоненциальная | 8 | 1 | 8 | 11 | 11 |
| Gauss2 | 1 | 1 | экспоненциальная | 8 | 1 | 8 | 11 | 11 |
| DanielWood | 1 | 1 | смешанная | 2 | 1 | 3 | 11 | 11 |
| Misra1b | 1 | 1 | смешанная | 2 | 1 | 14 | 11 | 11 |
| Kirby2 | 2 | 1 | рациональная | 5 | 1 | 14 | 11 | 11 |
| Hahn1 | 2 | 3 | рациональная | 7 | 1 | 18 [*]) | 11 | 11 |
| Nelson | 2 | 2 | экспоненциальная | 3 | >5 | 17 [*]) | 11 | 11 |
| MGH17 | 2 | 1 | экспоненциальная | 5 | >5 | 9 | 11 | 11 |
| Lanczos1 | 2 | 1 | экспоненциальная | 6 | 1 | 9 | 11 | 3 |
| Lanczos2 | 2 | 1 | экспоненциальная | 6 | >5 | 9 | 11 | 9 |
| Gauss3 | 2 | 1 | экспоненциальная | 8 | 1 | 9 | 11 | 11 |
| Misra1c | 2 | 1 | смешанная | 2 | 1 | 15 | 11 | 11 |
| Misra1d | 2 | 1 | смешанная | 2 | 1 | 14 | 11 | 11 |
| Rozzman1 | 2 | 2 | смешанная | 5 | 1 | 16 [*]) | 11 | 11 |
| ENSO | 2 | 2 | смешанная | 9 | 1.4 | 2 | 10 | 10 |
| MGH09 | 3 | 1 | рациональная | 4 | >5 | 7 | 11 | 10 |
| Thurber | 3 | 2 | рациональная | 7 | >5 | 10 [*]) | 11 | 11 |
| VoxBOD | 3 | 1 | экспоненциальная | 2 | 1 | 4 | 11 | 10 |
| Ratkowsky2 | 3 | 1 | экспоненциальная | 3 | 1 | 7 | 11 | 11 |
| MGH10 | 3 | 3 | экспоненциальная | 3 | >5 | 16 [*]) | 11 | 10 |
| Eckerle4 | 3 | 2 | экспоненциальная | 3 | 1 | 1 | 11 | 11 |
| Ratkowsky3 | 3 | 1 | экспоненциальная | 4 | >5 | 6 | 11 | 11 |
| Bennett5 | 3 | 3 | смешанная | 3 | >5 | 18 [*]) | 11 | 7 |

В первом столбце – названия примеров. Во втором и третьем столбцах приведены уровни сложности по NIST и по нашим результатам. Здесь 1 означает простую, а 3 – сложную задачу. В следующих столбцах указан тип модели и число неизвестных параметров в ней. Столбец, озаглавленный γ , содержит значения коэффициента нелинейности (см. раздел 3.4), а озаглавленный N(A) – величины разброса собственных значений (4.22), причем значения, отмеченные знаком ^{*}), могут быть существенно уменьшены с помощью мас-

штабирования (см. раздел 4.4). В последних двух столбцах приведено количество верных знаков в оценках параметров и стандартных отклонений, полученных при расчете с помощью программы Fitter.

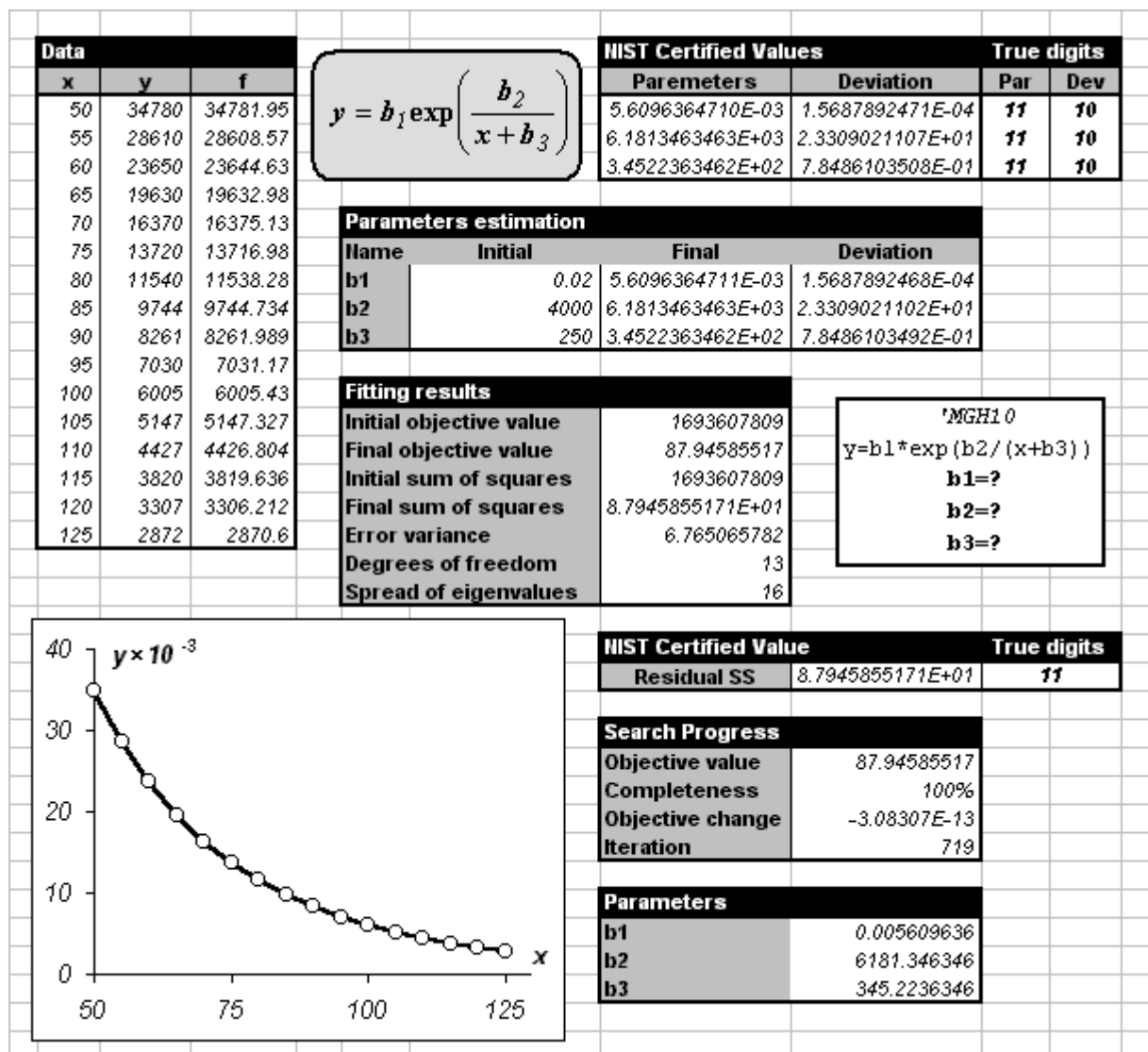


Рис. 4.4 Задача MGH10

Для большинства примеров эта программа дает 10 или 11 правильных знаков в оценках параметров, что является хорошим результатом, учитывая, что расчеты проводились с двойной точностью лишь на 32-битовой машине. В некоторых случаях оценки среднеквадратичных отклонений отличались от стандартных. Причины этого разные. Например, в задаче *Lanczos1* значения среднеквадратичных отклонений очень малы – в среднем около 10^{-10} . Это связано с тем, что и ошибка в данных очень мала – всего 10^{-13} . Поэтому, в этом примере мы не можем обратить F -матрицу с нужной точностью. В результате сами оценки находятся очень точно, а их отклонения только с тремя значащими цифрами. В задаче *Bennett5* получаются только 7 правильных знаков в стандартных отклонениях. Это

связано с большим разбросом собственных значений (величина $N(A)$). В принципе, разброс можно уменьшить, произведя нормировку и шкалирование, однако в этом случае не удалось подобрать подходящего преобразования.

Один из примеров NIST (задача *MGH10*) представлен на Рис. 4.4. Здесь y является зависимой переменной, а x – независимой переменной. Эти смоделированные данные подробно описаны в [82]. В работе [83] говорится о том, что эта задача представляет большие трудности даже для очень хороших алгоритмов. Из таблицы Табл. 4.9 видно, что в этом примере высокая степень нелинейности ($\gamma > 5$) сочетается с большим разбросом собственных значений $N(A)=16$. Для решения подобных задач нужно устанавливать высокую точность вычислений $\varepsilon = 10^{-15}$ и высокую точность подгонки (4.23) $\delta = 10^{-10}$. Это приводит к тому, что для обработки данных приходится выполнять много итераций – в данном случае – 719.

Этот же пример использовался для сопоставления Fitter с другими программами, использующими нелинейные регрессионные методы для анализа данных. Каждая из этих программ имеет свои сильные и слабые стороны. Для того чтобы определить, какая из программ лучше нужно сначала сформулировать критерии, по которым производится подобное сравнение. В первую очередь необходимо определить конечного пользователя данного программного обеспечения. В нашем случае это исследователь, преимущественно химик или биолог, которому необходимо математическое обеспечение для обработки и анализа данных. Поэтому программа должна быть проста в обращении и не требовать долгого времени для ее освоения. Она также должна иметь дополнительные статистические функции для анализа качества входных данных, анализа регрессионной модели и качества полученных результатов. С учетом всех этих соображений и были сформулированы следующие критерии:

- 1) надежность оптимизационного алгоритма и точность вычислений (Est);
- 2) точность статистических характеристик оценок параметров (Stat);
- 3) возможность ввода сложных функций, определяемых пользователем: явные и неявные функции, ОДУ (UDF);
- 4) размерность, т.е. сколько переменных (уравнений, откликов, предикторов) позволяет вводить программа (Dim);
- 5) скорость сходимости алгоритма, т.е. его эффективность (Eff);
- 6) наличие дополнительных статистических возможностей: проверка гипотез, статистическое моделирование (Tools)

7) возможность модификации целевой функции, т.е. взвешивание, байесовское оценивание, использование абсолютной/относительной ошибки (OFM);

8) совместимость со стандартным математическим обеспечением, таким как Microsoft Office (Comp).

Краткие обозначения, приведенные в скобках, например UDF, используются ниже на Рис. 4.5. Среди введенных нами критериев есть как объективные, например, точность вычислений, так и субъективные, например, простота использования, для определения которых использовались экспертные оценки.

Эти критерии были применены к программам, перечисленным в Табл. 4.10.

Табл. 4.10 Программы для нелинейного регрессионного анализа

| № | Программа | Поставщик |
|---|----------------|---|
| 1 | DataFit | Oakdale Engineering, Version 6.1 (Trial) © 1995-1999 [84] |
| 2 | Fitter | Polycert Ltd., Version 1.0 [5] |
| 3 | GraphPad Prism | GraphPad Software Inc. Version 2.01(Trial) [85] |
| 4 | NLREG | Phillip H. Sherrod, Shareware version [86] |
| 5 | Origin | Microcal Software, Inc. Version 5.0 (Demo) [87] |
| 6 | SAAM II | Saam Institute, University of Washington, Version 1.1 (Demo) [88] |
| 7 | Sigmaplot | SPSS Inc., Version 4.00 (Demo) [89] |
| 8 | Solver | Microsoft Excel |
| 9 | XLSTAT | Thierry Fahmy, Shareware version [90] |

Здесь собраны представители только одного класса прикладных пакетов – это небольшие, простые в использовании программы. Такие пакеты как Mathematica или MatLab находятся в другой категории, где от пользователя уже требуются навыки программиста. Все перечисленные выше программы оценивались только с одной точки зрения – их пригодности для решения задач нелинейного регрессионного анализа.

Все эти программы широко известны и часто применяются. Вот их общие свойства:

- работают под Windows, интерфейс стандартен для Windows –приложений;
- позволяют экспортировать и импортировать файлы в ASCII формате;
- имеют библиотеки моделей;
- позволяют задавать собственные модели, но только в виде явных функций;
- имеют широкие возможности для построения графиков.

DataFit - это инструмент для научных и инженерных расчетов, позволяющий строить графики входных данных, искать регрессионные функции и проводить статистический анализ. В DataFit возможно построение нелинейных многошаговых моделей, имеющих до 20 независимых переменных и до 100 неизвестных параметров. Надо отметить высокую точность оценки параметров, которая достигается применением экстраполяционного метода Ричардсона для численного вычисления производных по параметрам. Как таблицы данных, так и результаты, могут быть экспортированы в Excel и представлены в формате HTML файлов. Еще одна интересная особенность данного пакета, это возможность автоматически генерировать программный код для Бейсика или Си, с тем, чтобы пользователь мог его вставлять в свои программы.

GraphPad Prism позволяет легко преобразовывать и анализировать исходные данные с помощью нелинейной регрессии. Программа позволяет экспортировать и импортировать графики в виде картинок различных форматов (WMF, TIF, GIF, PCX или BMP файлов). Однако программа имеет существенное ограничение для уравнений модели, вводимой пользователем – нельзя использовать в уравнении более одной переменной, не допускает применения более 16 промежуточных переменных.

NLREG – это специальная программа для регрессионного анализа. Она применяется для обработки моделей множественной, линейной, полиномиальной, экспоненциальной и логистической регрессии, а также для нелинейной регрессионной модели общего вида. Эта программа показала высокую точность при обработке данных и высокую стабильность оптимизационного алгоритма [91].

Origin предназначен для анализа данных, построения регрессионных кривых, вычисление различных статистик и построение графиков. Совместим с Microsoft Office – рабочий лист Excel можно открыть, активируя соответствующую функцию меню, хотя эта возможность работает крайне нестабильно. В данной программе предусмотрены матричные окна, по внешнему виду схожие с рабочими листами Excel, но позволяющие использовать трехмерные матрицы. В программе также предусмотрен внутренний язык программирования Labtalk, обеспечивающий доступ ко всем операциям Origin и позволяющий пользователю оптимизировать свою работу. В данной программе возможно применение линейной и нелинейной регрессий. Модели, определяемые пользователем, могут иметь только явный вид, но, за счет применения Labtalk, могут быть достаточно сложными.

SAAM II. Весь интерфейс представлен в виде набора окон (одно главное, остальные дочерние). Модель может вводиться в виде явного алгебраического уравнения, либо задаваться кинетической схемой, которая автоматически переводится в систему ОДУ. В программе предусмотрены все возможности модификации целевой функции взвешивание, Байесовское оценивание, абсолютная/относительная ошибка. Однако надо отметить низкую точность вычислений данной программы.

SigmaPlot – программный продукт хорошо совместимый с Microsoft Office. Позволяет открыть рабочий лист Excel путем нажатия кнопки, позволяет вставлять и редактировать графики в Word, Excel и PowerPoint, строить графики и запускать статистические тесты непосредственно из рабочего листа Excel. Рабочий лист SigmaPlot по внешнему виду схож со стандартным рабочим листом Excel, однако сильно уступает ему по возможностям работы с таблицами данных. Программа нелинейной регрессии допускает модели, содержащие не более чем 25 параметров и 10 независимых переменных.

XLSTAT и **Solver** – это две надстройки для Excel. Они предоставляют пользователю возможность проводить анализ данных внутри Excel, что очень удобно для тех, кто хорошо знаком с этой программой и использует Microsoft Office в своей повседневной практике. Надо отметить низкую эффективность XLSTAT. Это связано с тем, что код программы написан на VBA [7], что значительно замедляет скорость вычислений. Данная программа позволяет пользователю вводить не только модель, но и уравнения для вычисления производных. Результаты, приведенные на Рис. 4.5, были получены именно таким образом. **Solver** – это надстройка, которая распространяется вместе с пакетом Excel. Эффективность Solver очень высока, однако полученные результаты имеют низкую точность из-за численного вычисления производных по параметрам.

Все перечисленные выше программы используют современные и надежные алгоритмы оптимизации, которые показали свою пригодность для решения нелинейных регрессионных задач. Для оценки точности программ был использован пример MGH10 (см. Рис. 4.4).

На Рис. 4.5 приведены результаты сопоставительного тестирования программных средств, перечисленных в Табл. 4.9 по описанным выше критериям. Результаты представлены бальными оценкам (от низшего балла 1 до наивысшего балла 4).

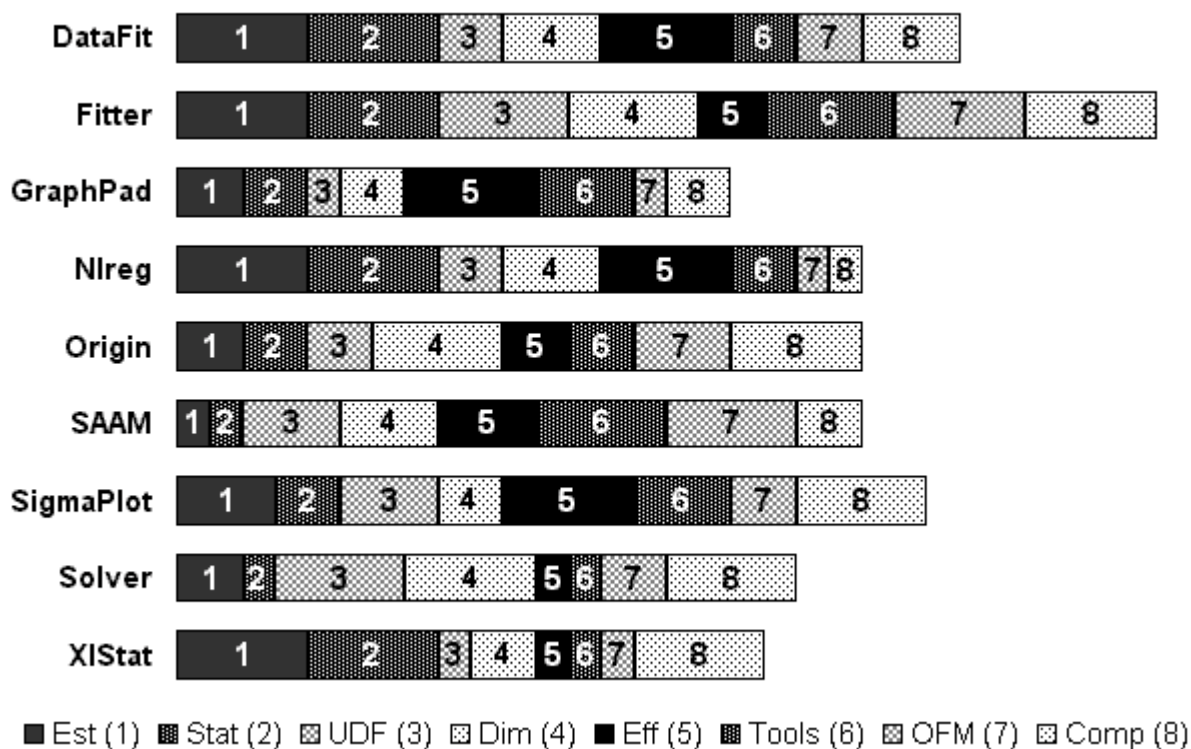


Рис. 4.5 Результаты сопоставительного тестирования программных средств

Разумеется, трудно быть беспристрастными, но получается, что Fitter наилучшим образом отвечает выдвинутым выше критериям.

4.4. Мультиколлинеарность

Мультиколлинеарность – это вырождение матрицы Гессе целевой функции (4.9). Величину этого вырождения можно ценить с помощью величины $N(A)$ (4.21) – характеризующей разброс собственных значений матрицы A .

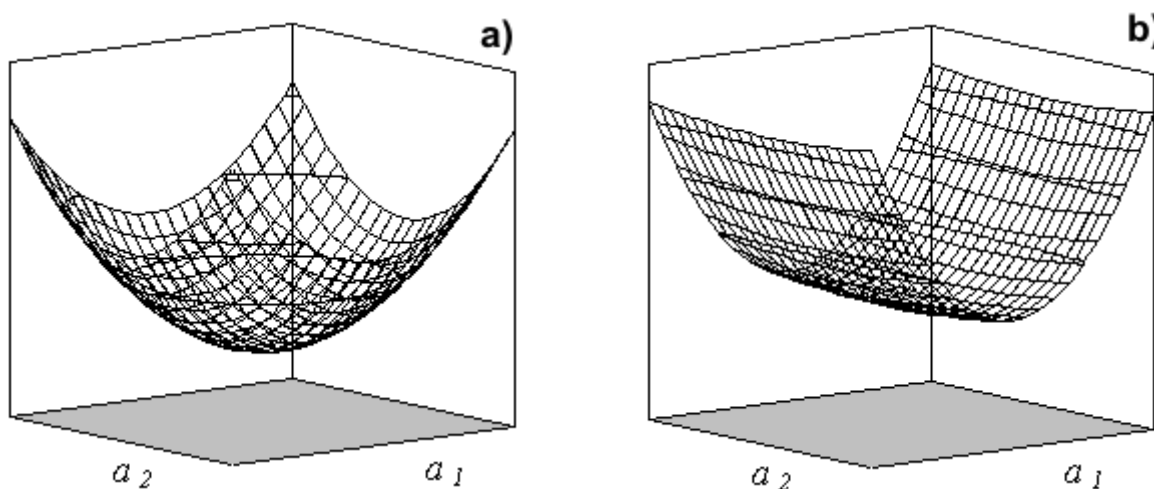


Рис. 4.6 Форма поверхности целевой функции для $N(A)=1$ (a) и для $N(A)=7$ (b)

Чем больше эта величина, тем хуже обусловлена матрица.

Эта же величина тесно связана с формой поверхности целевой функции – чем она больше, тем хуже выражен минимум, и тем больше эта поверхность напоминает овраг. На Рис. 4.6 показаны две такие поверхности для малого и большого разброса собственных значений. Видно, что во втором случае минимум не определяется, т.к. все точки на дне «оврага» равнозначны. Таким образом, главное и наиболее неприятное проявление мультиколлинеарности – это неопределимость оценок параметров модели [2, 94].

Рассмотрим то, как возникает мультиколлинеарность на простом, но поучительном примере. Пусть модель имеет следующий вид

$$y = a_1 a_2 x \quad (4.27)$$

Разумеется, это очень «глупая» модель, т.к. оценить отдельно параметры a_1 и a_2 невозможно. Модели такого типа, но, разумеется, гораздо более сложные, и поэтому, неочевидно вырожденные, часто встречаются в задачах химической кинетики. В работе [54] они названы *локально неидентифицируемыми*. Решить задачу оценивания в таких моделях, можно только предварительно выделив комбинации параметров, такие как $a_1 a_2$ в нашем примере. Этой задаче посвящено много работ [92, 93, 94] в которых дано ее исчерпывающее решение. Характерной особенностью моделей типа (4.27) является то, что вырожденность не исчезает ни при каком плане эксперимента, поэтому такую ситуацию естественно назвать *строгой мультиколлинеарностью*.

Рассмотрим теперь другую модель

$$y = a_1 \left(1 - e^{-a_2 x} \right) \quad (4.28)$$

которая является широко применяемым уравнением кинетики первого порядка.

На Рис. 4.7 представлены кинетические кривые уравнения (4.28) полученные для различных значений параметра a_2 . Видно, что при его уменьшении кривые все более подходят на прямую (0 на Рис. 4.7), соответствующую «глупому» уравнению (4.27)

$$y = a_1 \left(1 - e^{-a_2 x} \right) \approx a_1 a_2 x, \quad \text{при } a_2 x \ll 1 \quad (4.29)$$

При этом увеличивается и разброс собственных значений от $N(A)=2$ (при $a_2=2$) до $N(A)=6$ (при $a_2=0.1$). Строго говоря, матрица A остается все время невырожденной, однако обратить ее все труднее и труднее. Этот случай естественно называть *нестрогой мультиколлинеарностью*. Разумеется, в этом примере важна не величина параметра a_2 , а его соотно-

шение со значениями фактора x – ведь на начальной стадии этот процесс всегда хорошо приближается простой линейной моделью реакции нулевого порядка.

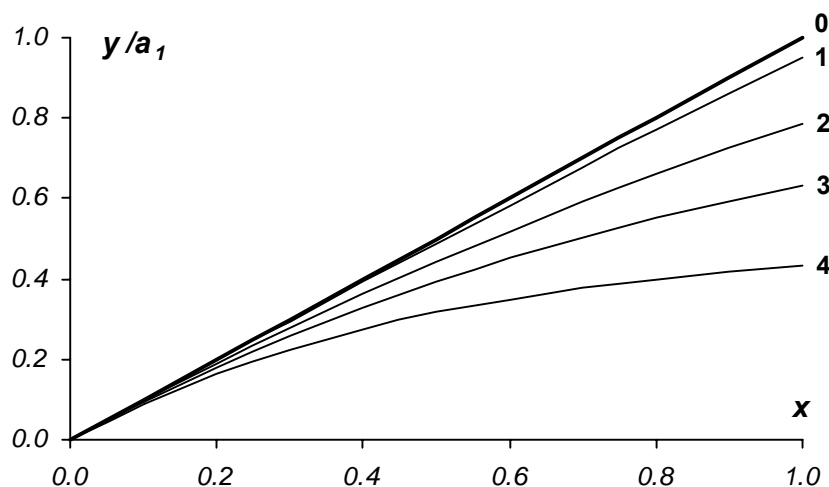


Рис. 4.7 Кинетические кривые уравнения 1-го порядка $y/a_1 = 1 - \exp(-a_2 x)$ для разных значений a_2 : 0.1 (1), 0.5 (2), 1.0 (3), 2.0 (4)

Из этого простейшего примера видно, что мультиколлинеарность теснейшим образом связана с тем, как спланированы и организованы экспериментальные данные. Во многих случаях причиной мультиколлинеарности в физико-химических данных является неудачное планирование эксперимента.

На проблему мультиколлинеарности можно смотреть с разных точек зрения. Так в линейном регрессионном анализе она проявляется как слишком большое число неизвестных параметров [108, 123, 127] в многомерной модели. Соответственно для борьбы с мультиколлинеарностью используются методы, понижающие размерность задачи – регрессия на главные компоненты (PCR), проекция на латентные структуры (PLS). В нелинейном регрессионном анализе, как правило, имеется не избыток параметров, а недостаток экспериментальных данных. Это связано с двумя обстоятельствами. С одной стороны все параметры нелинейной модели имеют некоторый важный физический смысл и не могут быть просто удалены из модели. Исключение – это случай строгой мультиколлинеарности, рассмотренный выше. С другой стороны, физико-химический эксперимент – это вещь сложная и дорогостоящая, и, подчас, имеются принципиальные ограничения, не позволяющие провести измерения в той области, которая была бы желательной с точки зрения планирования эксперимента. Поэтому для решения проблемы нестрогой мультиколлинеарности нужно применять некоторые специальные методы, специфические для нелинейной регрессии.

Важно понимать, что нестрогая мультиколлинеарность – это также и вычислительная проблема, связанная с точностью представления чисел в компьютере. Как известно, действительное число записывается в виде длинной последовательности из нулей и единиц (бит) в машинном слове. Не смотря на его значительную длину, часто она оказывается недостаточной, и некоторые значащие цифры теряются в ходе вычислений. Это происходит, например, тогда, когда к очень большому числу прибавляется очень маленькое и оказывается, что

$$10^{+20} + 10^{-20} + 10^{-20} + 10^{-20} + \dots + 10^{-20} = 10^{+20}$$

Именно такой случай имеет место, когда обращается матрица с большим разбросом собственных значений. Ясно, что эта операция является ключевой при поиске оценок параметров (4.13) и определения их статистических характеристик (1.36). Наша практика показывает, что проблемы с обращением возникают, когда разброс превышает 8 порядков.

Существуют простые приемы, позволяющие модифицировать исходную проблему так, чтобы разброс $N(A)$ существенно понизился. Они не изменяют сути решаемой задачи, но переформулируют ее так, чтобы облегчить компьютеру проведение вычислений. Цель этих преобразований – уменьшение разброса собственных значений, т.е. придание матрице A или, что то же самое, матрицы F (1.42) более регулярного вида.

Центрирование. Это простое преобразование предикторов, т.е. матрицы плана X (1.10)

$$X \rightarrow X' = X - X_0 \quad (4.30)$$

где

$$X_0 = \begin{bmatrix} x_{10} & x_{20} & \dots & x_{m0} \\ x_{10} & x_{20} & \dots & x_{m0} \\ \dots & \dots & \dots & \dots \\ x_{10} & x_{20} & \dots & x_{m0} \end{bmatrix}$$

Цель этого преобразования – перенос начала координат, поэтому, обычно,

$$x_{k0} = \frac{1}{N} \sum_{i=1}^N x_{ki}$$

Шкалирование. Это тоже преобразование матрицы плана X

$$X \rightarrow X' \quad (4.31)$$

где

$$X' = \begin{bmatrix} \frac{x_{11}}{s_1} & \frac{x_{21}}{s_2} & \dots & \frac{x_{m1}}{s_m} \\ \frac{x_{12}}{s_1} & \frac{x_{22}}{s_2} & \dots & \frac{x_{m2}}{s_m} \\ \dots & \dots & \dots & \dots \\ \frac{x_{1N}}{s_1} & \frac{x_{2N}}{s_2} & \dots & \frac{x_{mN}}{s_m} \end{bmatrix}$$

Цель этого преобразования – масштабирование всех предикторов, поэтому, обычно

$$s_k^2 = \frac{1}{N} \sum_{i=1}^N (x_{ki} - x_{k0})^2$$

Замена переменных. В отличие от первых двух, это более сложное и неопределенное преобразование, как предикторов, так и отклика, но, главное, – это замена параметров, так называемая перепараметризация.

$$a \rightarrow \varphi(a) \quad x \rightarrow \psi(x) \quad y \rightarrow \chi(y) \quad (4.32)$$

Подбор подходящих функций в формулах (4.32) – это не простая задача, которая решается индивидуально в каждом случае; однако этот прием может дать впечатляющий результат. Продемонстрируем все три приема на простом, но очень важном примере.

Рассмотрим уравнение Аррениуса [95]

$$k = k_0 \exp\left(-\frac{E}{RT}\right), \quad (4.33)$$

описывающее зависимость константы скорости химической реакции от температуры T . Это уравнение содержит одну константу R – газовую постоянную и два параметра: предэкспонент k_0 и энергию активации E , которые часто оцениваются по экспериментальным данным.

На Рис. 4.8 показаны типичный пример такой оценки. Видно, что в этом случае задача отличается большой жесткостью – разброс собственных значений составляет 17 десятичных порядков. Это же заметно и в F -матрице, детерминант которой равен 1.5×10^{-26} .

Простая замена переменной

$$k_0 \rightarrow a = \ln(k_0), \quad (4.34)$$

преобразует модель (4.33) к виду

$$k = \exp\left(a - \frac{E}{RT}\right), \quad (4.35)$$

для которого разброс собственных значений уже значительно меньше – только 9 порядков. Смысл преобразования (4.34) очень прост – если величина параметра $k_0 = 1.18 \times 10^{+11}$, то величина нового параметра $a = 25.49$. При этом детерминант F -матрицы уже равен 0.000215 .

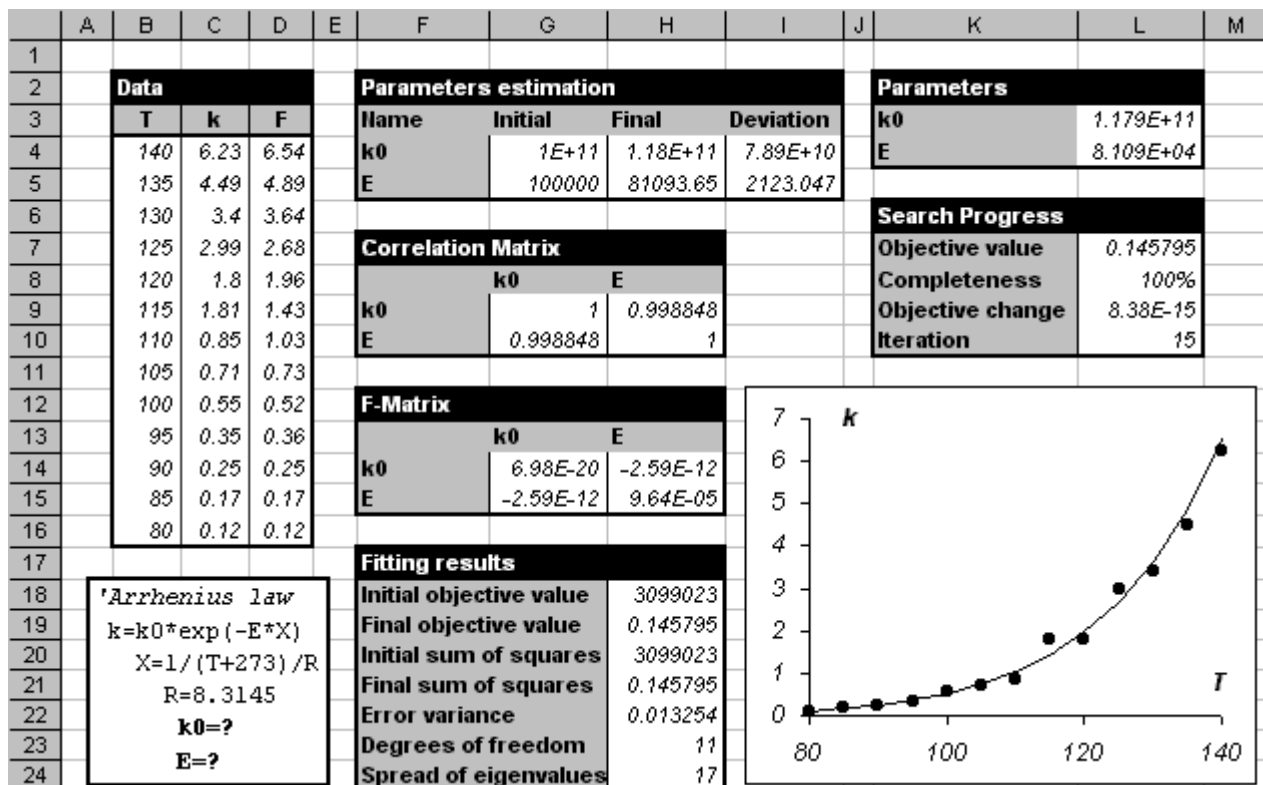


Рис. 4.8 Оценка параметров в уравнении Аррениуса

Однако, значение второго параметра – E все еще сильно отличается от величины первого параметра. Это можно исправить, применив *шкалирование* предиктора T .

$$T \rightarrow X = \frac{E_0}{R} \frac{1}{T + 273} \quad (4.36)$$

Масштабный фактор $E_0 = 62309.46$ подбирается так, чтобы вариация нового предиктора X была равна единице

$$\frac{1}{13} \sum_{i=1}^{13} (X_i - \bar{X})^2 = 1, \quad \text{где } \bar{X} = \frac{1}{13} \sum_{i=1}^{13} X_i \quad (4.37)$$

Тогда модель Аррениуса примет вид

$$k = \exp(a - bX), \quad (4.38)$$

где новый параметр b выражается через исходный параметр E по формуле

$$b = \frac{E}{E_0} = 1.301. \quad (4.39)$$

Разброс собственных значений для модели (4.38) уже очень хороший – всего 5 порядков, но мы можем еще его уменьшить.

Для этого применим *центрирование* предиктора X

$$X \rightarrow X' = X - X_0, \quad (4.40)$$

выбрав величину X_0 так, чтобы среднее значение нового предиктора X' было равно нулю

$$X_0 = \frac{1}{13} \sum_{i=1}^{13} X_i = 19.6137. \quad (4.41)$$

После этого модель Аррениуса примет вид

$$k = \exp(a' - bX'), \quad (4.42)$$

где новый параметр a' выражается через исходные параметры k_0 и E по формуле

$$a' = a - bX_0 = \ln(k_0) - \frac{E}{E_0} X_0 = -0.03362. \quad (4.43)$$

Для этой модели разброс собственных значений уже равен 1 – минимально возможной величине. Такой способ модификации уравнения Аррениуса является сейчас уже стандартной процедурой и применяется повсеместно, например, в задаче, разобранный в главе 8.

Разобранный пример показывает, что с помощью очень простых преобразований можно переписать исходную «физическую» модель (4.33) в новом «математическом» виде (4.43) и при этом решить все проблемы, связанные с мультиколлинеарностью. Здесь важно отметить два обстоятельства.

Во-первых, все преобразования исходных данных и модели ни коем образом не влияют на точность оценивания. Так среднеквадратичное отклонение оценки параметра a' равно 0.03211, что составляет 96% от его оценки -0.03362 . При этом среднеквадратичное отклонение оценки исходного параметра k_0 равно 7.89×10^{-10} , что составляет 67% от его оценки -1.18×10^{-11} . Иными словами можно сказать, что предварительная подготовка данных и модели влияет только на точность расчетов, но не на точность получаемых оценок. Это очевидно, т.к. никакие манипуляции с исходной информацией не могут ее увеличить, а могут лишь представить в новом виде, который более удобен для расчетов.

```
'Modified Arrhenius law
k=exp(a-b*X)
X=(E0/(T+273))/R-X0
X0=19.6137
E0=62309.46
R=8.3145
a=?
b=?
```

Рис. 4.9 Модифицированный закон Аррениуса

Второе обстоятельство носит чисто методический характер. Для выполнения операций с исходными данными, например, таких как (4.36) или (4.40), совсем не обязательно изменять таблицу данных, представленную на рабочем листе (см. Рис. 4.8). В большинстве случаев гораздо удобнее, проводить все изменения непосредственно в модели, например, так, как это показано на Рис. 4.9.

Третий способ борьбы с мультиколлинеарностью – последовательное байесовское оценивание. Во многих случаях исходную большую задачу удастся разбить на серию небольших и, тем самым, значительно улучшить матрицу Гессе. Пример такого подхода разобран в главе 7.

4.5. Результаты главы 4

В этой главе рассматривались основные алгоритмы и вычислительные процедуры, применяемые в нелинейном регрессионном анализе.

При обработке сложных физико-химических данных важно иметь возможность задавать модель исследуемого процесса интерактивно – вводом уравнений с клавиатуры. Разработан оригинальный алгоритм, позволяющий интерпретировать и компилировать модели, задаваемые пользователем в естественной математической форме записи – как систему явных, неявных и дифференциальных уравнений. Это дает возможность легко уточнять модель, подбирая регрессионную зависимость, наиболее точно описывающую экспериментальные данные.

Оценивание параметров тесно связано с проблемой точного вычисления производных модели по параметрам. Для ее решения был разработан другой оригинальный алгоритм, который автоматически находит эти производные в аналитической форме. Это позволяет использовать очень сложные многостадийные регрессионные зависимости и находить при этом все производные с высокой, машинной точностью, необходимой при поиске оценок.

Для задач анализа физико-химических данных характерна мультиколлинеарность или, иными словами – вырождение, которая проявляется в том, что не все искомые параметры модели можно оценить. Показано, что мультиколлинеарность определяется тремя основными причинами – формой представления модели, точностью вычислений и качеством экспериментальных данных. Предложены простые приемы модификации модели и данных, позволяющие значительно снизить вырождение при оценке. Эти приемы продемон-

стрированы на классическом примере – оценивании предэкспоненциального множителя и энергии активации в законе Аррениуса.

Качество используемых алгоритмов была проверена с помощью стандартного, общепризнанного набора тестовых данных, разработанного Национальным институтом стандартов США (NIST). Результаты этого тестирования показали высокое качество использованных процедур. Стандартной точности – 10-11 десятичных разрядов в оценках – удалось достигнуть во всех 27 примерах, представленных в базе данных NIST. Такая эффективность позволяет обрабатывать мультиколлинеарные данные, недоступные другим известным программам, предназначенным для регрессионного анализа.

5. Описание программы Fitter

Fitter – это программа, которая работает под управлением системы Excel, входящей в стандартный пакет Microsoft Office. Используя этот инструмент, можно решать нелинейные регрессионные задачи. Обработка данных предусматривает также создание дополнительной статистической информации о данных и качестве подгонки модели. Основное преимущество программы Fitter – это возможность оценивать значения параметров сложных моделей, задаваемых пользователем в естественной алгебраической записи. В ней применяется алгоритм оптимизации, который обеспечивает высокую точность даже для сильно нелинейных моделей. Вся входная и выходная информации представляется как таблицы и графики в Excel.

С помощью Fitter можно получить следующую информацию:

- *оценки параметров и их дисперсии;*
- *ковариационную, корреляционную матрицы, матрицу Фишера;*
- *начальные и конечные значения суммы квадратов и целевой функции;*
- *оценку дисперсии ошибки, вычисленную по регрессии и по выборке.*

Кроме того возможна проверка гипотез:

- *тест Стьюдента о выбросах;*
- *тест серий о корреляции остатков;*
- *тест Бартлета на гомоскедастичность;*
- *тест адекватности Фишера.*

а также

- построение доверительных интервалов методами стохастической аппроксимации и с помощью связанного моделирования.

Используя функции Fitter и VBA программирование можно создавать специальные подпрограммы, которые будут автоматически выполнять обработку данных.

5.1. Основные свойства, возможности, требования и ограничения

Fitter получает всю исходную информацию из открытой рабочей книги Excel. Эта информация должна быть помещена в таблицах непосредственно на рабочий лист (данные и параметры) или записана в текстовом поле (модель). Пользователь может организовать рабочее пространство так, как ему удобно – использовать только один лист для размещения всей информации, или несколько листов, или даже листы в разных рабочих книгах. Полученные результаты также выводятся как таблицы на листах рабочей книги (см. например,

Рис. 4.4). Все таблицы имеют схожую структуру и внешний вид. Пользователь может изменять внешний вид этих таблиц по своему вкусу.

Входная информация, которую необходимо задать для начала работы, состоит из трех частей. Первая часть – это *данные (Data)*, т.е. численные (экспериментальные) значения, которые будут обрабатываться. Таблица данных должна включать обязательные элементы – предикторы (имена и значения) и отклики (имя и значения). Кроме того, в этой таблице могут быть и необязательные элементы. Вторая часть входной информации – это модель. *Модель (Model)* состоит из системы уравнений, связывающих имена данных, параметров и промежуточных величины. Эти две части (данные и модель) необходимы для начала работы. Третья часть – *априорная информация (Bayes)* является необязательной. Она включает априорные значения параметров, информационную матрицу и другие элементы, описанные в главе 2.

Некоторые элементы входной информации – *имена*, являются общими для разных частей. Так, например, одно и то же имя отклика должно использоваться в таблице данных и в модели, одинаковые имена параметров должны участвовать в модели и в априорной информации. Пользователь должен внимательно следить за этим, иначе заданная информация будет неправильно интерпретирована.

Для того чтобы объяснить программе, какая информация будет использована при обработке, ее нужно *зарегистрировать* с помощью *Регистраторов (Wizards)*. Для задания входной информации применяются три регистратора: *Данных, Модели и Априорной Информации*. Кроме того, имеется регистратор для задания режимов работы программы – *Регистратор Настроек (Options)* и *Диалог дополнительных возможностей (Plus)*. Все они вызываются из *Главного диалога Fitter*. Используя различные регистраторы, пользователь только указывает программе, где находится входная информация или куда нужно поместить выходную информацию, выбирает способ ее интерпретации и представления. При этом он не может изменить ее содержание – все исправления возможны только вне процедуры регистрации, непосредственно на рабочих листах. Как только какая-то часть информации (данные, модель, и т.п.) зарегистрирована, она сохраняется в памяти компьютера до тех пор, пока не будет заменена в новой регистрации. Если изменить только некоторые значения внутри уже зарегистрированной таблицы или текстового поля, то их не нужно регистрировать заново – программа сделает это автоматически при начале подгонки.

Каждый регистратор состоит из нескольких шагов, выполняемых последовательно. На каждом из этих шагов можно закончить процесс, нажав кнопку **Finish** или же пройти через всю регистрацию для того, чтобы быть уверенным, что программа правильно интерпретирует входную информацию. Поначалу лучше выбрать второй вариант, потому что программа подчиняется своей внутренней логике, которая может отличаться от намерений пользователя. Выбирая правильные области для входной/выходной информации, а также используя элементы управления регистратором, можно корректировать этот процесс. Подробно регистрация описана в разделах 5.7-5.10.

Все действия, выполняемые пользователем с помощью регистраторов, могут быть проделаны в автоматическом режиме с помощью *функций Fitter*. Для этого нужно написать программу (макрос) на встроенном в Excel языке VBA [7]. Кроме того, такие программы можно создавать и автоматически. Детали того, как это делается, приведены в разделе 5.12.

Программа не имеет ограничений на размер входной информации – количество предикторов, параметров, данных, уравнений и т.п. ограничено только объемом памяти, который поддерживает операционная система, установленная на компьютере. Некоторые ограничения все же имеются. Сначала приведем несущественные, формальные ограничения.

Максимальная длина имени (идентификатора) 15 символов.

Максимальная длина численной константы, используемой в модели 15 знаков.

Максимальная длина численных значений в таблицах 39 знаков.

К существенным, принципиальным относятся следующие ограничения.

Можно использовать только модели с одним откликом.

Нельзя использовать модели с перекрестными связями переменных, т.е. запрещена, например, следующая модель

$$\begin{array}{l} y=a*x+1 \\ a=b*y \\ b=? \end{array}$$

В состав программы входят четыре файла, перечисленных в Табл. 5.1. В этой таблице использованы следующие обозначения. Символ *Root* означает диск, на котором находится соответствующая директория, например, *Root=C:*. Символ *OS* означает имя директории, в которую установлена операционная система, например *OS=Windows*. Символ *MO* означает имя директории, где находится система Microsoft Office, например *MO= Program*

Files\Microsoft Office. Символ *Username* означает имя директории, содержащей индивидуальные настройки, например *Username= All Users*.

Табл. 5.1 Размещение файлов программы Fitter

| Файл | Excel 7.0/Win 95 | Excel 8.0 (9.0) / Win 98 (NT, ME) |
|------------|------------------------|---|
| Fitter.dll | Root\OS | Root\OS |
| Fitter.xla | Root\MO\Office\Library | Root\OS\Profiles\Username\Application Data\Microsoft\Addins |
| Fitter.hlp | Root\OS | Root\OS |
| Fitter.cnt | Root\OS | Root\OS |

Для инсталляции программы разместите файлы так, как указано в таблице. После загрузки Excel выполните подключение надстройки Fitter.xla с помощью команд **Сервис/Надстройки (Tools/Add-Ins)**.

5.2. Данные

Для начала работы с программой Fitter нужно поместить на листе рабочей книги данные и модель. Данные вводят в клетки рабочего листа, а модель записывается в текстовое поле (**Drawing/Textbox**).

Fitter считывает данные из Таблицы данных.

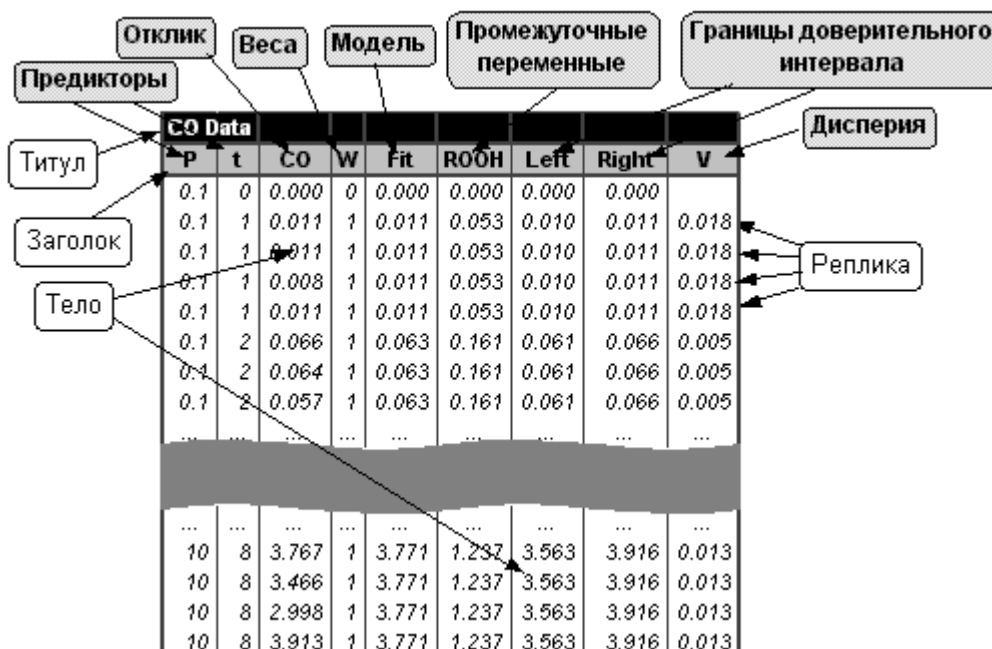


Рис. 5.1 Устройство таблицы данных

Таблица данных должна занимать одну непрерывную область на листе. Она разбивается по горизонтали на три части: *титул* – одна (необязательная) произвольная строка, *заголо-*

вок – одна обязательная строка, содержащая *имена* переменных, участвующих в модели, и *тело*, содержащее любое количество строк с численными значениями.

На Рис. 5.1 приведен пример такой таблицы. Каждая строка тела соответствует одному измерению, а каждый столбец – одной переменной. Обязательно должен присутствовать хотя бы один столбец *предикторов* (**P** и **t** на Рис. 5.1) и ровно один столбец *отклика* (**CO** на Рис. 5.1). Эти столбцы обязательно должны быть озаглавлены именами. Повторные измерения – *реплики*, записываются одно под другим в разных строках. Количество предикторов не ограничено. Каждому измерению может соответствовать индивидуальное значение *веса* – неотрицательное число (столбец **W** на рисунке). Заголовок у этого столбца не обязателен, но если он есть, то рекомендуется начинать его с буквы **W**. Если колонки с весами нет, то считается, что все веса равны единице. Если отсутствует какое-то значение отклика, то ему автоматически приписывается вес равный нулю.

В таблице могут, также, находится и другие, необязательные столбцы, в которые заносятся значения, вычисленные в ходе обработки данных.

Модельные значения – только один столбец, значения в котором меняются в ходе подгонки (**Fit** на Рис. 5.1). Заголовок не обязателен, но если он есть, то рекомендуется начинать его с буквы **F**.

Промежуточные переменные – несколько столбцов, значения в которых меняются в ходе подгонки (**ROOH** на Рис. 5.1). Заголовки обязательно должны присутствовать и соответствовать именам, использованным в модели.

Доверительные пределы – один или два столбца, значения в которых меняются в ходе подгонки (**Left** и **Right** на Рис. 5.1). Заголовок не обязателен, но если он есть, то рекомендуется начинать его с букв **L** и **R** соответственно.

Дисперсия – только один столбец, заполняемый автоматически после регистрации данных. (**V** на Рис. 5.1). Заголовок не обязателен, но если он есть, то рекомендуется начинать его с буквы **V**.

Заполнять необязательные столбцы до регистрации не нужно, однако в ходе регистрации необходимо зарезервировать для них место. Кроме того, в таблице могут присутствовать и другие столбцы с дополнительной информацией, неиспользуемой программой. В ходе регистрации (См. раздел 5.8) их нужно отметить как игнорируемые. Порядок следования столбцов произволен, однако для автоматического распознавания рекомендуется следующий порядок: предикторы, отклик, вес, модель, промежуточные, левый предел, правый предел, дисперсия.

5.3. Модель

Модель – это система уравнений, задаваемая пользователем. Она может включать *явные* и *неявные* формулы, а также *обыкновенные дифференциальные уравнения*. Используемые синтаксические правила довольно просты и близки к обычным математическим обозначениям. Каждая модель занимает отдельное текстовое поле (**TextBox**, а не прямоугольник!). На Рис. 5.2 приведен пример модели в виде, в котором она записывается на рабочем листе. Исключая, разумеется, пояснения в овальных окнах и стрелки.

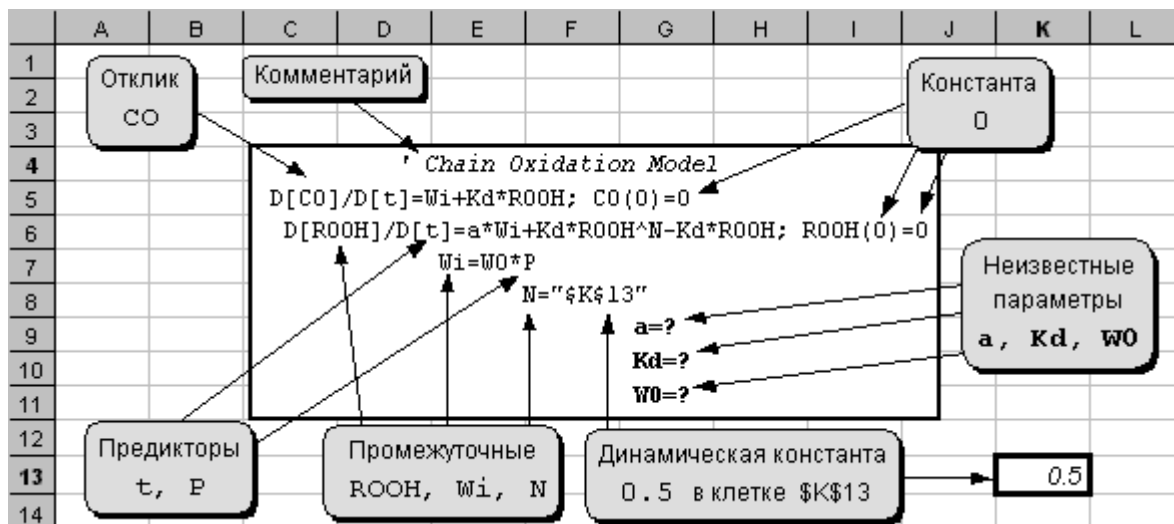


Рис. 5.2 Устройство модели

Модель состоит из *уравнений* и *неизвестных параметров*. Кроме того, в модели могут присутствовать *комментарии*. Каждое уравнение, параметр или комментарий занимает отдельную строку. Последовательность строк – произвольная. Пробелы () и пустые строки могут вводиться в модель для облегчения ее восприятия – при регистрации они игнорируются

Для записи модели используются *идентификаторы*, которые служат именами предикторов, отклика, параметров, промежуточных переменных и символических констант. Идентификатор состоит из последовательности букв и цифр длиной не более 15 символов. Первый символ обязательно должен быть буквой. Различаются заглавные и прописные буквы, так, что t1 и T1 – это разные идентификаторы. Например

1T – неверный идентификатор, т.к. первый символ – цифра.

VariableIsVeryLong – неверный идентификатор, т.к. длина более 15 символов.

wrong@Var – неверный идентификатор, т.к. символ @ является недопустимым.

В модели, представленной на Рис. 5.2 идентификаторами являются следующие величины: t, P, CO, ROOH, wi, N, a, Kd, w0.

Численные константы могут быть записаны в натуральном виде с «фиксированной точкой» (–5, 2.5 и т.п.) или в экспоненциальном виде с «плавающей точкой» (–4.9E3, 5.289E–8 и т.п.). Для разделения целой и десятичной части используется символ . (точка). Все числа считаются действительными величинами, независимо от того указана или нет десятичная точка. Длина численной константы не более 15 символов. Для стандартной константы $\pi = 3.1415926535898$ используется идентификатор PI, который следует писать только заглавными буквами. Кроме того, модель может использовать динамические константы, определяемые через адрес клетки (в любой кодировке на любом рабочем листе), в которой находится эта величина. Адрес выделяется символами кавычек (").

Примеры численных констант

```
eps=1e-8
```

```
a=-7.380502
```

```
Temp="Sheet1!$A$2"
```

В модели, представленной на Рис. 5.2 имеются две константы. Первая, равная 0, непосредственно участвует в уравнениях, а вторая, равная 0.5, вводится через промежуточную переменную N динамической ссылкой на клетку \$K\$13.

В модели должны использоваться те же имена для предикторов, откликов и промежуточных переменных, что и в заголовке зарегистрированных данных. В ходе регистрации модели программа контролирует соответствие этих идентификаторов. Сравните Рис. 5.1 и Рис. 5.2., в которых общими являются имена предикторов t и P, отклика CO и промежуточной переменной ROOH.

Неизвестные параметры вводятся в модель с помощью специальной конструкции вида

```
par=?
```

где par – это имя (идентификатор) параметра. Например, в модели, представленной на Рис. 5.2 величины a, kd и w0 являются неизвестными параметрами.

Начало *комментария* обозначается значком ` (апостроф). Весь текст, начиная от этого символа и до конца строки, считается комментарием и игнорируется при регистрации модели. Комментарии должны находиться в отдельных строках. В модели, представленной на Рис. 5.2 первая строка

```
`Chain Oxidation Model
```

является комментарием.

Идентификаторы в модели могут связываться следующими бинарными *операторами*:

Табл. 5.2 Бинарные операции программы Fitter

| № | Символ | Операция |
|----|--------|--|
| 01 | ^ | возведение в степень |
| 02 | * | умножение. |
| 03 | / | деление |
| 04 | - | вычитание (используется также как унарный минус) |
| 05 | + | сложение |

Для записи модели используются следующие *стандартные функции* одного аргумента x .

Табл. 5.3 Стандартные функции программы Fitter

| № | Имя | Функция | № | Имя | Функция |
|----|-----------|----------------------------|----|-----------|-----------------------------|
| 06 | exp(x) | экспонента | 07 | log(x) | натуральный логарифм |
| 08 | sin(x) | синус | 09 | cos(x) | косинус |
| 1a | tan(x) | тангенс | 1b | ctg(x) | котангенс |
| 1c | arcsin(x) | арксинус | 1d | arccos(x) | арккосинус |
| 1e | arctg(x) | арктангенс | 1f | arcctg(x) | арккотангенс |
| 10 | sh(x) | синус гиперболический | 11 | ch(x) | косинус гиперболический |
| 12 | th(x) | тангенс гиперболический | 13 | cth(x) | котангенс гиперболический |
| 14 | arcsh(x) | арксинус гиперболический | 15 | arcch(x) | арккосинус гиперболический |
| 16 | arth(x) | арктангенс гиперболический | 17 | arccth(x) | арккотангенс гиперболически |
| 31 | hev(x) | функция Хевисайда | 32 | ixp(x) | интегральная экспонента |
| 33 | imp(x) | импульсная функция | 34 | erf(x) | интеграл ошибок |

В Табл. 5.3, помимо элементарных, приведены и четыре специальные функции:

функция Хевисайда, –

$$\text{hev}(x) = \begin{cases} 0, & x \geq 0 \\ 1, & x < 0 \end{cases} \quad (5.1)$$

интегральная экспонента, –

$$\text{ixp}(x) = \frac{1}{x} \int_1^{\infty} t^{-2} e^{-xt} dt \quad (5.2)$$

импульсная функция, –

$$\text{imp}(x) = \begin{cases} 0, & x \neq 0 \\ 1, & x = 0 \end{cases} \quad (5.3)$$

и интеграл ошибок –

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (5.4)$$

Примеры использования этих функций приведены, например, в главе 9 (Рис. 9.3)

В Табл. 5.2 и Табл. 5.3 имеются столбцы (озаглавленные **No**), которые содержат шестнадцатеричные номера, соответствующие кодировке операций (См. раздел 4.2)

При вводе уравнений можно использовать неограниченное число *промежуточных переменных*. Соответствующие примеры приведены во всей работе, в частности на Рис. 9.15, где показана одна из сложнейших моделей. Единственное требование к модели состоит в том, чтобы последовательность всех уравнений, участвующих в ней, определяла величину отклика однозначно. При этом можно использовать различные комбинации из явных, неявных функций, а также обыкновенных дифференциальных уравнений первого порядка.

Явной функцией называется уравнение вида

$$y = f(x, a, I) \quad (5.5)$$

где y – это зависимая, т.е. вычисляемая переменная (отклик или промежуточная переменная), x – это вектор предикторов, a – это вектор неизвестных параметров и I – это набор промежуточных переменных. Нельзя использовать зависимую переменную (y) в правой части явной функции.

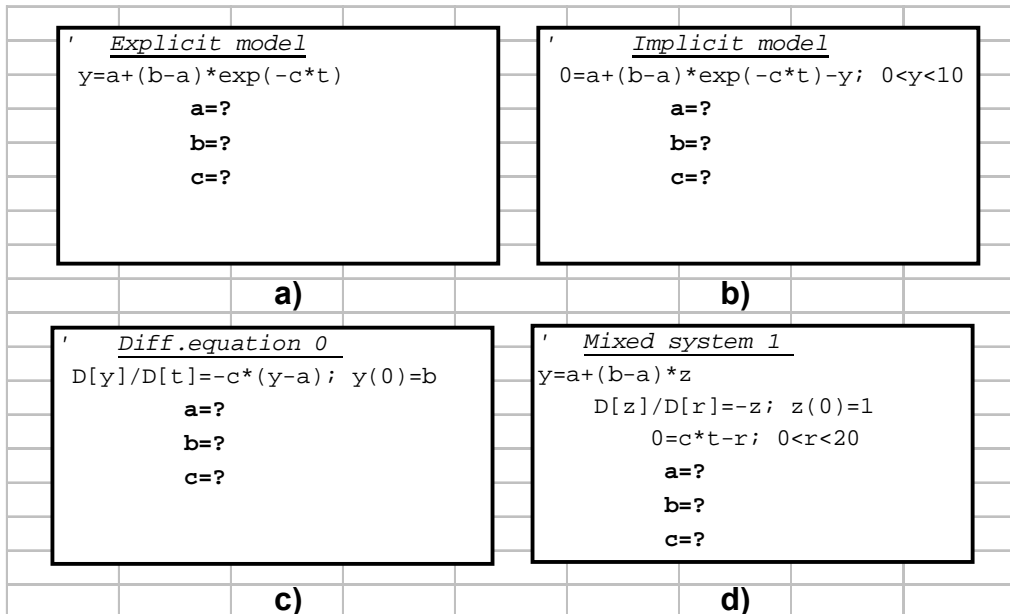


Рис. 5.3 Примеры различных форм записи одной и той же модели

Явная функция является предпочтительной формой записи модели, однако если модель невозможно выразить в этом виде, то следует использовать другие функциональные фор-

мы. На Рис. 5.3 а) показана модель, в форме явной функции. Здесь y – это отклик, t – предиктор; a , b и c – это неизвестные параметры.

Неявная функция определяется уравнением вида

$$0 = f(\mathbf{x}, \mathbf{a}, \mathbf{I}, y); Y_l < y < Y_r \quad (5.6)$$

где y – это зависимая, т.е. вычисляемая переменная (отклик или промежуточная переменная), \mathbf{x} – это вектор предикторов, \mathbf{a} – это вектор неизвестных параметров и \mathbf{I} – это набор промежуточных переменных. Зависимая переменная (y) обязательно должна присутствовать в правой части неявной функции. Неявная функция всегда вводится в модель с нулем (0) в левой части уравнения. Эта конструкция $0=$ является обязательным элементом синтаксиса, идентифицирующим неявную функцию. В правой части должен стоять символ «точка с запятой» (;), который используется как разделитель между неявным уравнением и его граничными условиями. Последние должны находиться в той же строке, что и само уравнение. Символ неравенства (<) используется для отделения зависимой переменной (y) от левого (Y_l) и правого (Y_r) граничных условий.

В ходе вычисления неявной функции ищется значение той переменной, идентификатор которой находится между двумя символами неравенства (y), а численная величина находится между двумя граничными условиями. Это сделано для того, чтобы придать определенность неявной функции, которая, в принципе, может иметь несколько корней. Если границы заданы неверно, т.е. так, что на интервале $[Y_l, Y_r]$ находятся несколько значений y , удовлетворяющих уравнению (5.6), то результат может быть непредсказуемым.

Граничные условия могут быть положительными константами, предикторами или промежуточными переменными, но не могут быть выражениями. Следует учесть, что отрицательная константа, такая как, например $Y_l = -1$ запрещена, поскольку она является выражением. В этом случае следует использовать какую-нибудь дополнительную промежуточную переменную, например, так

$$\begin{aligned} Y_l &= cVal \\ cVal &= -1. \end{aligned}$$

Неявная функция вычисляется дольше, чем явная, но быстрее, чем дифференциальное уравнение, однако, если модель невозможно выразить в этом виде, то следует использовать другие функциональные формы. На Рис. 5.3 б) показана та же модель, что и на рисунке а), но в форме неявной функции. Здесь константы 0 и 10 являются граничными условиями.

Обыкновенное дифференциальное уравнение, имеет следующий вид

$$\frac{dy}{dt} = f(y, t, \mathbf{x}, \mathbf{a}, \mathbf{I}); y(t_0) = y_0, \quad (5.7)$$

где y – это зависимая, т.е. вычисляемая переменная (отклик или промежуточная переменная), t – это другая переменная, \mathbf{x} – это вектор предикторов, \mathbf{a} – это вектор неизвестных параметров и \mathbf{I} – это набор промежуточных переменных. Переменная t может быть предиктором, параметром или промежуточной переменной, но не может быть откликом или константой. Переменные t_0 и y_0 , определяющие начальное условие, могут быть константами, предикторами и промежуточными переменными, но они не могут быть выражениями.

При вводе модели уравнение (5.7) записывается в следующем стандартном виде

$$D[y]/D[t]=f(t, y, x_1, \dots, a_1, \dots, I_1, \dots) ; y(t_0)=y_0$$

Символ «точка с запятой» (;) используется как разделитель между уравнением и начальным условием. Конструкция $D[y]/D[t]$ (допускается вариант $d[y]/d[t]$), является обязательным элементом синтаксиса, идентифицирующим дифференциальное уравнение. При этом в первых квадратных (sic!) скобках должен находиться идентификатор зависимой переменной (y), относительно которой решается уравнение, а во вторых квадратных скобках – идентификатор переменной, по которой проводится дифференцирование (t). Первый идентификатор, стоящий после точки с запятой в правой части уравнения (y) должен совпадать с идентификатором зависимой переменной.

Дифференциальное уравнение вычисляется дольше, чем неявная функция, поэтому, если возможно, то следует использовать другие функциональные формы. На Рис. 5.3 с) показана та же модель, что и на рисунках а) и б), но в форме дифференциального уравнения.

В модели можно комбинировать различные формы уравнений, например, так, как это показано на Рис. 5.3 d). Здесь приведена та же модель, что и на рисунках а), б) и с), но в смешанной форме, где первая (после комментария) строка является явной функцией, вторая строка – дифференциальным уравнением, а третья строка – неявной функцией. Эту модель следует читать снизу вверх. Тогда видно, что из последнего уравнения определяется промежуточная переменная r в зависимости от предиктора t и параметра c ; из предпоследнего находится промежуточная величина z в зависимости от промежуточной величины r , а из первого уравнения – отклик y как функция параметров a , b и промежуточной величины z .

Следующие символы и конструкции являются зарезервированными, то есть они не должны употребляться иначе, чем предписано в Табл. 5.4

Табл. 5.4 Зарезервированные символы и конструкции программы Fitter

| Символы | Описание |
|------------------------|--|
| () [] { } | используются для управления порядком вычислений |
| ; | отделяет уравнение от граничных условий в неявной модели или от начального условия в дифференциальном уравнении. |
| . | десятичная точка, используемая в константах |
| < | используется в граничных условиях неявной модели и отделяет вычисляемую переменную от левой и правой границы |
| ? | обозначает, что эта переменная является неизвестным параметром модели |
| = | разделяет правую и левую части в уравнениях |
| + - * / ^ | арифметические операторы |
| 0= | стоит в левой части неявной функции |
| D[·]/D[·] d[·]/d[·] | стоит в левой части дифференциального уравнения |
| ` | стоит в начале строки и обозначает, что эта строка является комментарием |
| ” | выделяет адрес динамической константы |
| PI | идентификатор числа $\pi=3.1415926535898$ |

5.4. Параметры

Регрессионные параметры – это неизвестные независимые переменные, величины которых оцениваются программой. Число параметров не ограничено. *Имена* параметров определяются в модели, а их значения могут выводиться на рабочий лист в область таблицы параметров, задаваемую при регистрации модели. Начальные значения параметров могут задаваться в этой области еще до регистрации модели.

Таблица параметров состоит из трех частей (см. Рис. 5.4): *титул* – одна (необязательная) произвольная строка, *заголовок* – один обязательный столбец, содержащая имена параметров, и *тело* – один обязательный столбец, содержащее численные значения параметров.

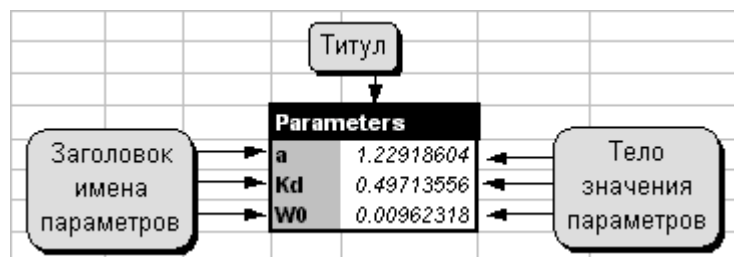


Рис. 5.4 Устройство таблицы параметров

При регистрации модели имена параметров записываются в заголовок автоматически, поэтому их можно заранее не указывать. Если в теле таблицы приведены численные значения параметров, то при регистрации модели они считываются оттуда, если же тело таблицы не заполнено, то все значения устанавливаются равными единице. Задавая начальные значения параметров близкими к оценкам, можно ускорить процесс подгонки модели. В ходе подгонки значения параметров меняются. Если нужно изменить значения параметров после регистрации модели, то можно просто ввести в тело таблицы нужные величины. При этом модель не надо регистрировать заново – программа автоматически учтет все исправления. На Рис. 5.4 приведена таблица параметров, соответствующая данным, изображенным на Рис. 5.1 и модели на Рис. 5.2.

5.5. Априорная информация

Данные можно обрабатывать с учетом *априорной информации*, так как это описано в главе 2. Перед регистрацией эта информация должна быть размещена на рабочем листе.

Таблица априорной информации состоит из трех частей (см. Рис. 5.5): *титул* – одна (необязательная) произвольная строка, обязательный *заголовок*, состоящий из одного столбца, содержащего имена параметров и одной строки, в которую автоматически помещается вспомогательная информация. Третья часть таблицы это *тело* – несколько столбцов, содержащих численные значения величин, участвующих в априорной информации.

Структура таблицы априорной информации должна быть следующей:

- первый столбец – имена параметров (**Name**);
- второй столбец – их априорные значения (**Value**);
- последующие столбцы – априорная матрица (**Matrix**);
- последний столбец – для обозначения исключенных параметров (**Exclude**).

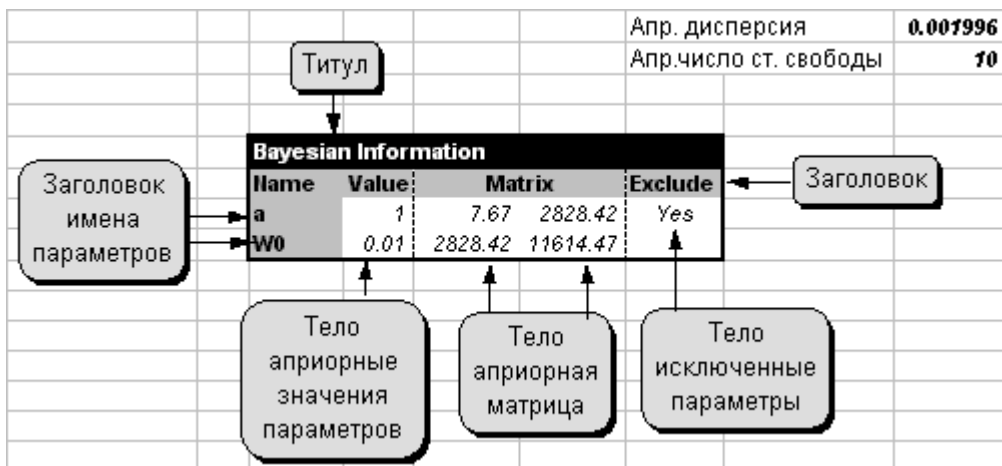


Рис. 5.5 Устройство таблицы априорной информации

Если необходимо исключить некоторый параметр из обработки и сохранить тем самым его априорное значение, то в соответствующую клетку последнего столбца нужно поместить любой символ. При регистрации он будет заменен словом *Yes*.

Число обязательных строк в таблице априорной информации равно числу параметров, о которых имеется априорная информация, плюс один ряд заголовков. Число столбцов равно числу строк плюс три. Те параметры, о которых нет априорной информации, могут также присутствовать в таблице, но с пустыми строками в теле. При этом должны быть добавлены и соответствующие пустые столбцы.

Априорные значения *дисперсии* (2.27) и *числа степеней свободы* (2.28) не участвуют в таблице и могут находиться в любом месте рабочего листа, например, так, как это показано на Рис. 5.5. Здесь приведена таблица априорной информации, соответствующая данным, изображенным на Рис. 5.1, модели на Рис. 5.2.с параметрами на Рис. 5.4. Параметр k_d не входит в таблицу, т.к. о нем нет априорной информации; а параметр a является исключенным.

5.6. Главный Диалог Fitter

Для того чтобы активировать систему Fitter первый раз, нужно использовать главное меню системы Excel, выбрав **Сервис(Tools)/Fitter**. После этого появится Главный Диалог программы Fitter (См. Рис. 5.6).

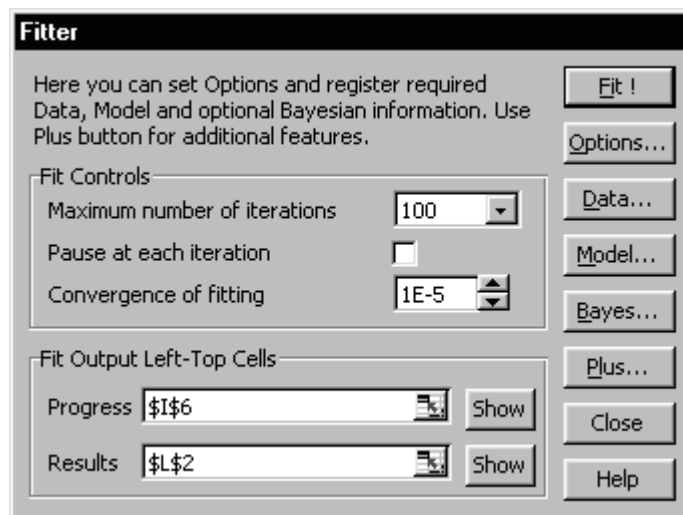



Рис. 5.6 Главный диалог программы Fitter

Его можно вызвать и с помощью *Панели инструментов* (первая кнопка  на Рис. 5.10). Если какая-то кнопка не активируется, это означает, что еще не зарегистрирована вся необходимая для нее информация.

Из этого диалога осуществляется выход на Регистраторы Настроек (**Options...**), Данных (**Data...**), Модели (**Model...**), Априорной Информации (**Bayes...**), а также в Диалог дополнительных действий (**Plus...**). Кроме того, здесь устанавливаются некоторые общие параметры настройки программы.

Maximum number of iterations. Т.к. в поиске используется итеративная процедура (см. раздел 4.1), то нужно задать *максимальное число итераций* (от 0 до 50 000). Нулевое число итераций используется для того, чтобы вывести результаты без подгонки. Процедура обработки данных заканчивается тогда, когда будет выполнено максимальное число итераций или когда программа достигнет минимума целевой функции. Рекомендованное значение – 100 устанавливается по умолчанию.

Pause at each iteration. Отметьте эту настройку, если нужно контролировать процесс подгонки. Тогда в ходе минимизации целевой функции после *каждой итерации будет делаться пауза*, и



Рис. 5.7 Диалог в поиске

выводиться специальный диалог (см. Рис. 5.7), с помощью которого можно сделать еще одну итерацию (**Next**), остановить поиск (**Cancel**) и продолжить дальше без пауз (**Continue**). Рекомендованное значение «НЕТ», устанавливается по умолчанию.

Convergence of fitting. Эта величина – *точность подгонки* – контролирует выход из процедуры минимизации. Если за несколько итераций относительное изменение целевой функции меньше, чем эта величина, то подгонка завершается. Точность подгонки должна всегда быть больше (ниже), чем точность вычислений (см. раздел 5.7). Выбор слишком высокой (малой) точности может затруднить процедуру обработки. Рекомендованное значение – $1E-5$ устанавливается по умолчанию.

Fit Output Left-Top Cells. Можно оставить два поля ввода **Progress** и **Result**, объединенных в раздел *левая верхняя клетка для вывода результатов*, пустыми, но тогда Fitter не будет выводить на рабочий лист соответствующую информацию.

Progress. Если нужно отслеживать процесс подгонки, то в это поле нужно ввести ссылку на левую верхнюю клетку для таблицы *хода поиска*. Кнопка **Show** позволяет увидеть, какую область займет эта таблица. Сама таблица будет сформиро-

| Search Progress | |
|------------------|-------------|
| Objective value | 0.486428136 |
| Completeness | 76% |
| Objective change | 5.20E-04 |
| Iteration | 6 |

Рис. 5.8 Таблица хода поиска


вана автоматически после начала подгонки (см. Рис. 5.8). На каждом шагу поиска в нее выводятся следующие значения: *целевой функции (Objective value)*, *завершенность поис-*

ка (4.20) (**Completeness**), *относительное изменение целевой функции* (**Objective change**) и *номер итерации* (**Iteration**).

Results. Для того чтобы получить *результаты подгонки* не достаточно отметить соответствующие настройки в регистраторе настроек (см. раздел 5.7.). Необходимо также ввести ссылку на левую верхнюю клетку для таблицы результатов подгонки. Кнопка **Show** позволяет увидеть, какую область на рабочем листе займет эта таблица.

Главному диалогу соответствует VBA функция `FitterFit` (см. раздел 5.12).

5.7. Регистратор настроек

Регистратор Настроек можно вызвать из Главного Диалога (кнопка **Options...** на Рис. 5.6) или с помощью *Панели инструментов* (вторая кнопка  на Рис. 5.10).

Регистратор Настроек состоит из трех шагов, на каждом из которых можно установить большое число настроек программы Fitter. На Рис. 5.9 изображено *окно первого шага*.

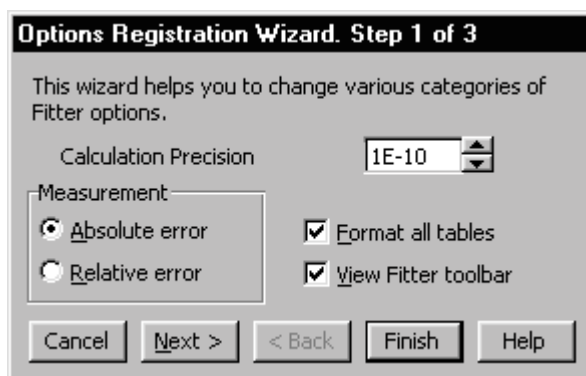


Рис. 5.9 Регистратор настроек. Первый шаг.

Calculation Precision. *Точность вычислений* может изменяться в пределах от $1.E-4$ до $1.E-15$. Эта величина определяет, насколько точно вычисляются значения неявных функций или решаются дифференциальные уравнения (см. раздел 5.3). Точность вычислений должна быть выше (меньше), чем точность подгонки. Рекомендованное значение $1.E-10$ устанавливается по умолчанию.

Measurement. Absolute or Relative error. Эта настройка позволяет выбрать *тип ошибки* – *абсолютная* (1.11) или *относительная* (1.12). Рекомендованное значение – *абсолютная* устанавливается по умолчанию.

Format all tables. Если нужно автоматически *форматировать все таблицы*, т.е. придавать им более наглядный вид, то устанавливается эта настройка. Стили, используемые для

этого форматирования, определяются на третьем шаге Регистратора Настроек (см. ниже). Рекомендованное значение – «ДА», значение по умолчанию – «НЕТ».

View Fitter toolbar. Эта настройка показывает или прячет *Панель инструментов Fitter*. Применение этой панели ускоряет работу с программой и позволяет запускать различные процедуры прямо с рабочего листа, минуя Главный Диалог. Нажатие каждой кнопка запускает соответствующую процедуру так, как это указано на Рис. 5.10. Рекомендованное значение – «ДА», значение по умолчанию – «НЕТ».

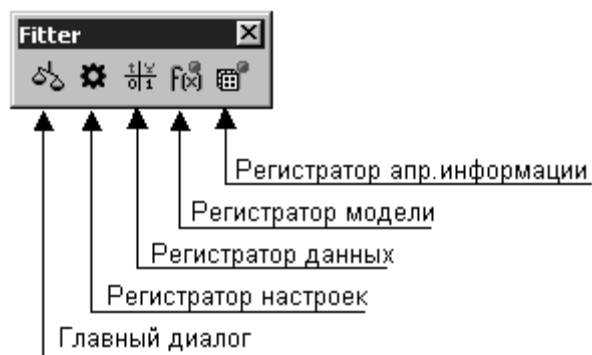


Рис. 5.10 Панель инструментов Fitter

Во всех Регистраторах Fitter применяется схожая система навигации по шагам с использованием следующих кнопок: **Cancel** – прерывание регистрации, **Next>** – переход на следующий шаг регистрации, **<Back** – возвращение на предыдущий шаг, **Finish** – завершение регистрации с установкой всех оставшихся настроек по умолчанию, **Help** – контекстная помощь.

Второе окно Регистратора Настроек позволяет установить, какие результаты будут рассчитаны и выведены в конце подгонки данных. Его вид показан на Рис. 5.11.

Важно иметь в виду, что вся заказанная информация будет выведена, только если в Главном Диалоге (Рис. 5.6) в настройке **Results** указана ссылка на верхнюю левую клетку таблиц результатов подгонки.

Первому и второму окнам Регистратора Настроек соответствует VBA функция `FitterOptions` (см. раздел 5.12).

Окно разбито на три вертикальных секции, озаглавленных: **Statistics** (Статистики), **Tests** (Тесты) и **Prediction** (Прогноз).

В первой секции (**Statistics**) собраны настройки, регулирующие вывод величин, характеризующих свойства и качество найденных оценок.

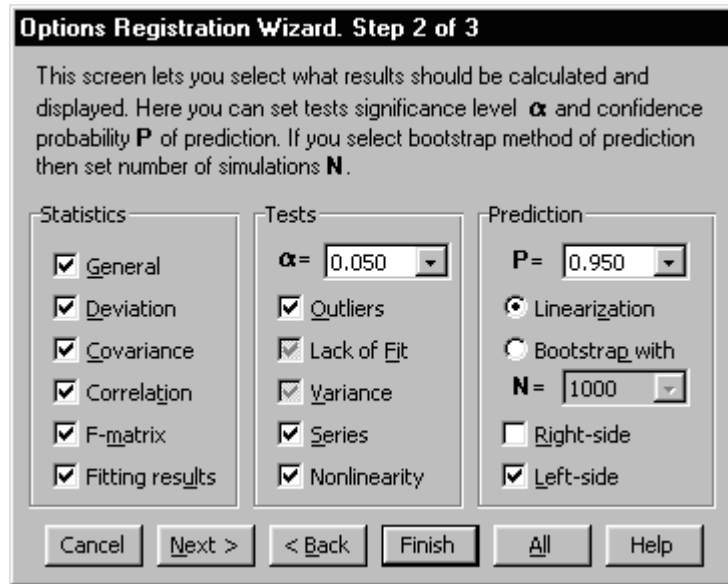


Рис. 5.11 Регистратор настроек. Второй шаг.

General. Выводится таблица (см. Рис. 5.12), содержащая общую информацию о задаче: дата (**Date**), адрес Таблицы данных (**Data**), имя Модели (**Model**), адрес Таблицы параметров (**Parameters**), адрес Таблицы априорной информации (**Bayes**), точность вычислений (**Precision**), точность подгонки (**Convergence**), тип ошибки (**Error type**), значимость (**Significance**), доверительная вероятность (**Confidence**), метод, используемый для предсказания (**Prediction**). По умолчанию устанавливается значение «НЕТ».

| General | |
|--------------|----------------|
| Date | 2.21.99 13:27 |
| Data | \$B\$2:\$J\$66 |
| Model | CO Model |
| Parameters | \$L\$2 |
| Bayes | \$7\$2:\$X\$5 |
| Precision | 0.00000001 |
| Convergence | 0.00001 |
| Error type | Relative |
| Significance | 0.05 |
| Confidence | 0.95 |
| Prediction | Linearization |

Рис. 5.12 Таблица общей информации

Deviation. Эта настройка включает в вывод таблицу оценок параметров (см. Рис. 5.13). Эта таблица состоит из четырех столбцов. Имена параметров (**Name**), их начальные значения (**Initial**), окончательные значения (1.34) оценок (**Final**) и среднеквадратичные отклонений (1.40) оценок (**Deviation**). По умолчанию устанавливается значение «НЕТ».

| Parameters estimation | | | |
|-----------------------|---------|----------|-----------|
| Name | Initial | Final | Deviation |
| a | 1 | 1.229186 | 2.983525 |
| Kd | 0.5 | 0.497136 | 0.076829 |
| W0 | 0.01 | 0.009623 | 0.006156 |

Рис. 5.13 Таблица оценок параметров

Covariance. Эта настройка включает в вывод матрицу ковариаций (1.36) оценок параметров (см. Рис. 5.14). По умолчанию для этой настройки устанавливается значение «НЕТ».

| Covariation Matrix | | | |
|--------------------|----------|----------|----------|
| | a | Kd | W0 |
| a | 8.901423 | -0.11505 | -0.01779 |
| Kd | -0.11505 | 5.90E-03 | 1.73E-04 |
| W0 | -0.01779 | 1.73E-04 | 3.79E-05 |

Рис. 5.14 Матрица ковариаций

Correlation. Эта настройка включает в вывод матрицу корреляций (1.41) оценок параметров (см. Рис. 5.15) По умолчанию для этой настройки устанавливается значение «НЕТ».

| Correlation Matrix | | | |
|--------------------|----------|----------|---------|
| | a | Kd | W0 |
| a | 1 | -0.50192 | -0.9687 |
| Kd | -0.50192 | 1 | 0.36526 |
| W0 | -0.9687 | 0.36526 | 1 |

Рис. 5.15 Матрица корреляций

F-matrix. Эта настройка включает в вывод *F-матрицу* (1.42) (см. Рис. 5.16) По умолчанию для этой настройки устанавливается значение «НЕТ».

| F-Matrix | | | |
|----------|----------|----------|----------|
| | a | Kd | W0 |
| a | 7.192471 | 47.73243 | 3159.367 |
| Kd | 47.73243 | 771.2842 | 18894.9 |
| W0 | 3159.367 | 18894.9 | 1458589 |

Рис. 5.16 F-Матрица

Fitting results. Эта настройка включает в вывод несколько полезных статистик (см. Рис. 5.17). *Начальное значение целевой функции (Initial objective value), окончательное значение целевой функции (Final objective value), начальное значение суммы квадратов (1.22) отклонений (Initial sum of squares), окончательное значение суммы квадратов (Final sum of squares), оценка взвешенной дисперсии (1.33) ошибки (Error variance), число степеней свободы (1.44) (Degrees of freedom), разброс (4.21) собственных значений (Spread of eigenvalues).* По умолчанию для этой настройки устанавливается значение «НЕТ».

| Fitting results | |
|-------------------------|---------|
| Initial objective value | 356.235 |
| Final objective value | 132.6 |
| Initial sum of squares | 356.235 |
| Final sum of squares | 132.52 |
| Error variance | 2.32492 |
| Degrees of freedom | 57 |
| Spread of eigenvalues | 6 |

Рис. 5.17 Результаты подгонки

Во *второй секции (Tests)* собраны настройки, определяющие, какие гипотезы (см. раздел 1.4) будут проверяться после подгонки. Для каждой гипотезы возможны три варианта ответа:

- мало данных (short of data)* – это означает, что решение принять нельзя;
- принято (accepted)* – гипотеза принята (для данной значимости);
- отклонено (rejected)* – гипотеза отклонена (для данной значимости).

Уровень значимости (α) – это мера строгости при проверке гипотеза. Он может изменяться в пределах от 0.5 до 0.001. По умолчанию в этой настройке устанавливается значение 0.05.

Outliers. Эта настройка включает в вывод таблицу результатов проверки *критерия выбросов* (см. Рис. 5.18). Она включает следующие значения. *Результат проверки (Result), имена предикторов и отклика, а также их значения в*

| Outliers Test | |
|---------------|----------|
| Result | Rejected |
| P | 10 |
| t | 1 |
| CO | 1.824891 |

Рис. 5.18 Тест выбросов

точке, которая является выбросом. По умолчанию для этой настройки устанавливается значение «НЕТ».

Lack of fit. Эта настройка включает в вывод таблицу результатов проверки *теста адекватности* (см. Рис. 5.19). Она включает следующие значения. *Результат* проверки (**Result**), оценка *дисперсии по регрессии* (1.33) (**Variance by Fit**), оценка *дисперсии по выборке* (1.59) (**Variance by Population**), *тестовая статистика* (1.61) (**Test Value**), *критическое значение* (1.62) (**Critical Value**). Значение по умолчанию «НЕТ».

| Lack-of-Fit | |
|------------------------|----------|
| Result | Rejected |
| Variance by Fit | 2.324917 |
| Variance by Population | 0.156576 |
| Test value | 66.78044 |
| Critical value | 1.974499 |

Рис. 5.19 Тест адекватности

Variance. Эта настройка включает в вывод таблицу результатов проверки *теста дисперсий* (см. Рис. 5.20). Она включает следующие значения. *Результат* проверки (**Result**), *имена предикторов*, а также их *значения* в точке, в которой достигнута максимальная *частная дисперсия* по репликам (1.57) (**Variance**) Значение по умолчанию «НЕТ».

| Variance Test | |
|---------------|----------|
| Result | Rejected |
| P | 10 |
| t | 1 |
| Variance | 2.223208 |

Рис. 5.20 Тест дисперсий

Series. Эта настройка включает в вывод таблицу результатов проверки *критерия серий* (см. Рис. 5.21). Она включает следующие значения. *Результат* проверки (**Result**), *число неотрицательных остатков* (**Positive residuals**), *число отрицательных остатков* (**Negative residuals**), *число серий* (**Series of sign**) (1.53). Значение по умолчанию «НЕТ».

| Series Test | |
|--------------------|----------|
| Result | Accepted |
| Positive residuals | 31 |
| Negative residuals | 32 |
| Series of sign | 39 |

Рис. 5.21 Тест серий

Nonlinearity. Эта настройка включает в вывод таблицу со значениями *коэффициента нелинейности* (см. Рис. 5.22). Она включает две величины – *точное* (**Exact Value**) и *грубое* (**Rough Value**) (3.44) значения этого коэффициента. *Число повторений* (3.42) для вычислений устанавливается в настройке N= (см. ниже). По умолчанию для этой настройки устанавливается значение «НЕТ».

| Coefficient of Nonlinearity | |
|-----------------------------|-------|
| Exact Value | 56.99 |
| Rough Value | >5 |

Рис. 5.22 Коэффициент нелинейности

В *третьей секции (Prediction)* собраны настройки, определяющие, как строятся доверительные интервалы при прогнозировании (см. Главу 3):

Достоверность прогноза (**P=**) определяет *доверительную вероятность*, для которой строятся доверительные интервалы. Она может быть установлена в пределах от 0.5 до 0.999. По умолчанию для этой настройки устанавливается значение 0.95.

Linearization. Этой настройкой для прогнозирования выбирается метод *стохастической аппроксимации* (см. S-метод в разделе 3.1). Эта настройка устанавливается по умолчанию.

Bootstrap with. Этой настройкой для прогнозирования выбирается метод *связанного моделирования* (см. A-метод в разделе 3.2). Если установлена эта настройка, то становится доступной настройка $N=$, определяющая *число повторений* в этом методе. По умолчанию для этой величины устанавливается значение *5000*.

Right-side, Left-side. Эти настройки определяют, какой доверительный интервал будет построен. Настройка **Right-side** дает *правосторонний*, а **Left-side** – *левосторонний* интервал. Если установлены обе настройки – то интервал будет *двусторонним*, если ни одна – то интервал не будет строиться. Результаты выводятся в Таблицу данных (Рис. 5.1), в которой должно быть зарезервировано соответствующее количество столбцов. По умолчанию для обеих настроек устанавливается значение «НЕТ».

Третье окно *Регистратора Настроек* позволяет изменить стили, используемые для форматирования входных и выходных таблиц с тем, чтобы сделать их более наглядными и удобными. Вид этого окна показан на Рис. 5.23.



Рис. 5.23 Регистратор настроек. Третий шаг.

В программе Fitter используется несколько типов таблиц.

- Таблица данных (**Data**)
- Таблица параметров (**Param**)
- Таблица априорной информации (**Bayes**)
- Таблица результатов (**Output**)

Все они имеют сходную структуру (см.Рис. 5.24), состоящую из трех элементов: *титул* (header), *заголовок* (title) и *тело* (body). При форматировании для каждого из этих элементов используется свой стиль. По умолчанию, применяется набор встроенных стилей, каж-

дый из которых именуется *Table_Element*, например, для форматирования заголовка таблицы данных используется стиль *Data_Header*.

| Bayesian Information | | | | |
|----------------------|-------|---------|----------|---------|
| Name | Value | Matrix | | Exclude |
| a | 1 | 7.67 | 2828.42 | Yes |
| W0 | 0.01 | 2828.42 | 11614.47 | |

| Search Progress | |
|------------------|----------|
| Objective value | 0.486428 |
| Completeness | 100% |
| Objective change | 1.33E-12 |
| Iteration | 6 |

| Parameters | |
|------------|----------|
| a | 1.229186 |
| Kd | 0.097136 |
| W0 | 0.009623 |


Рис. 5.24 Структура таблиц

Кроме того, форматируется и текстовое поле модели (Model), в котором имеются следующие элементы: *комментарии* (comment), *параметры* (parameter) и *тело* (body). Стили, используемые здесь по умолчанию, называются аналогично, например, *Model_Comment*.

С помощью третьего окна Регистратора Настроек можно изменить соответствие стилей, применяемых для форматирования. Новые стили нужно создать предварительно, используя возможности системы Excel. (**Формат/Стиль**). Для каждого элемента (список **Text Formatting**) можно выбрать соответствующий стиль из падающего листа **Style**, где показываются все стили, определенные для этой рабочей книги. Пример форматирования будет показан в окошке **Sample**. Изменения в стилях форматирования будут видны на рабочем листе, только если в первом окне Регистратора Настроек выбрана настройка **Format all tables** (Рис. 5.9).

Третьему окну Регистратора Настроек соответствует VBA функция `FitterStyles` (см. раздел 5.12)

5.8. Регистратор данных

Регистратор Данных можно вызвать из Главного Диалога (кнопка **Data...** на Рис. 5.6) или с помощью Панели инструментов (третья кнопка  на Рис. 5.10). Этот регистратор состоит из трех шагов, первый из которых показан на Рис. 5.25

В поле **Range** нужно ввести ссылку на всю область, занимаемую Таблицей данных.

Второй шаг Регистратора Данных показан на Рис. 5.26

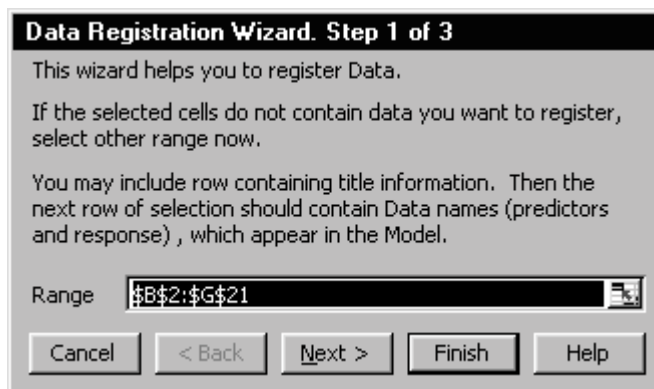


Рис. 5.25 Регистратор данных. Первый шаг.

В этом окне можно увидеть, как программа интерпретирует данные и, если необходимо, изменить настройки. Первая из них **Treat first row as Data title text** (*Считать первый ряд Таблицы Данных титулом*).

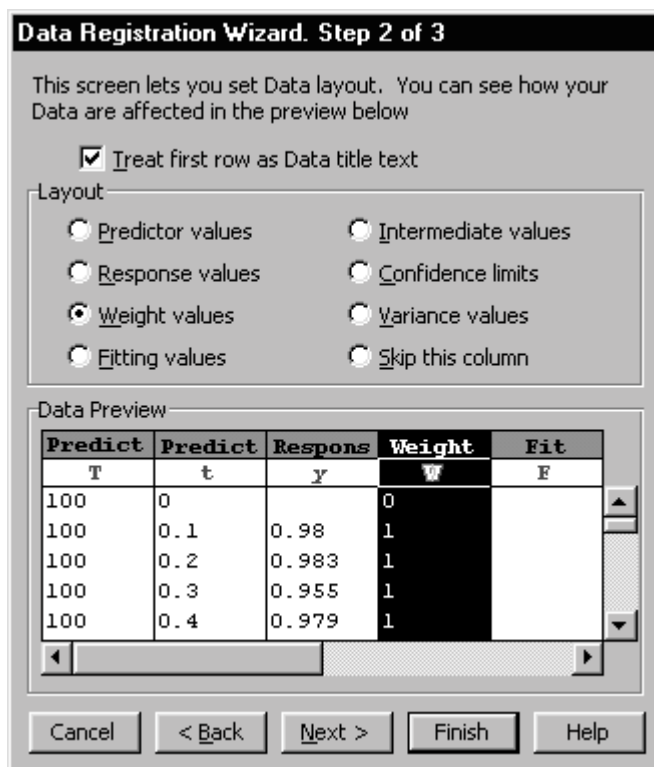


Рис. 5.26 Регистратор данных. Второй шаг.

В окне просмотра (**Data Preview**) заголовки в первом ряду (черные на сером фоне) показывают, как Fitter понимает столбцы в Таблице данных. Заголовки второго ряда (красные на белом фоне) – это имена из Таблицы Данных, ниже расположены численные значения (см. Рис. 5.1).

Если какой-то столбец интерпретирован неверно, то нужно выделить этот столбец в окне просмотра и затем установить правильную настройку, выбрав ее из списка, озаглавленного **Layout**. В нем приведены следующие варианты настроек:

Predictor values – предиктор

Intermediate values – промежуточное значение

Response values – отклик

Confidence limit – доверительная граница

Weight values – вес

Variance values – дисперсия

Fitting values – модель

Skip this column – игнорировать этот столбец

Значения, введенные ранее в Таблицу Данных, здесь изменять нельзя.

Третий шаг Регистратора Данных показан на Рис. 5.27.

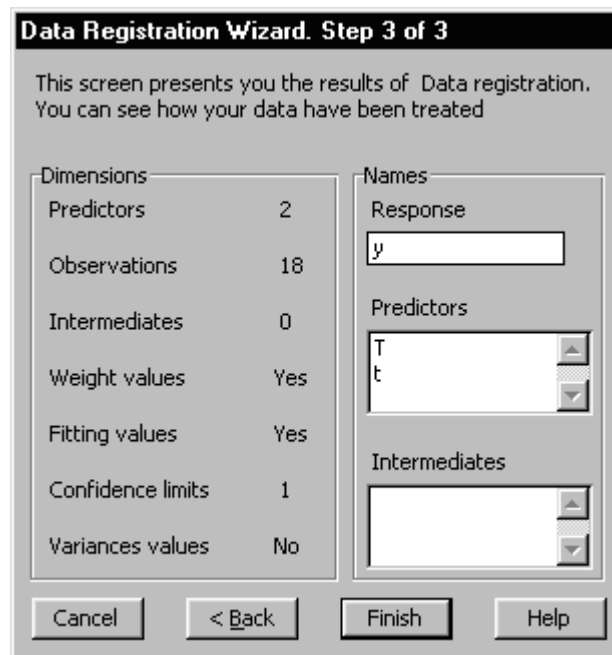



Рис. 5.27 Регистратор данных. Третий шаг.

Это окно представляет результаты регистрации данных – имена и количество элементов структуры данных. Никаких настроек здесь нет. Если эти результаты не верны, то можно вернуться на шаг назад и изменить настройки.

Регистратору Данных соответствует VBA функция `FitterData` (см. раздел 5.12)

5.9. Регистратор модели

Регистратор Модели можно вызвать из Главного Диалога (кнопка **Model...** на Рис. 5.6) или с помощью Панели инструментов (четвертая кнопка  на Рис. 5.10).

Регистрировать модель можно только после того, как зарегистрированы данные. До этой поры соответствующие кнопки будут неактивны. Этот регистратор состоит из трех шагов, первый из которых показан на Рис. 5.28.

В этом окне имеются две настройки. С помощью первой (**Text Box**) выбирается нужная модель из выпадающего списка.

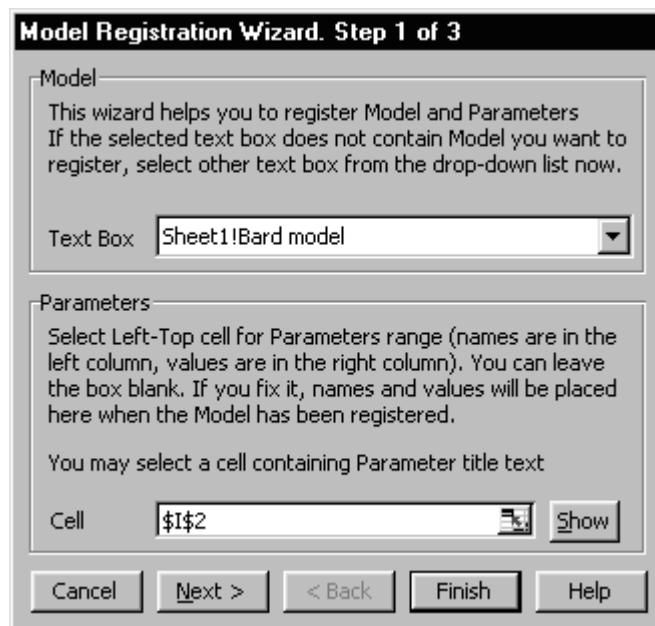


Рис. 5.28 Регистратор модели. Первый шаг.

В поле, озаглавленное **Cell** можно ввести ссылку на адрес левой верхней клетки рабочего листа, где будет размещена Таблица Параметров (Рис. 5.4). Если оставить это поле пустым, то таблица выводиться не будет.

Второй шаг Регистратора Модели показан на Рис. 5.29.

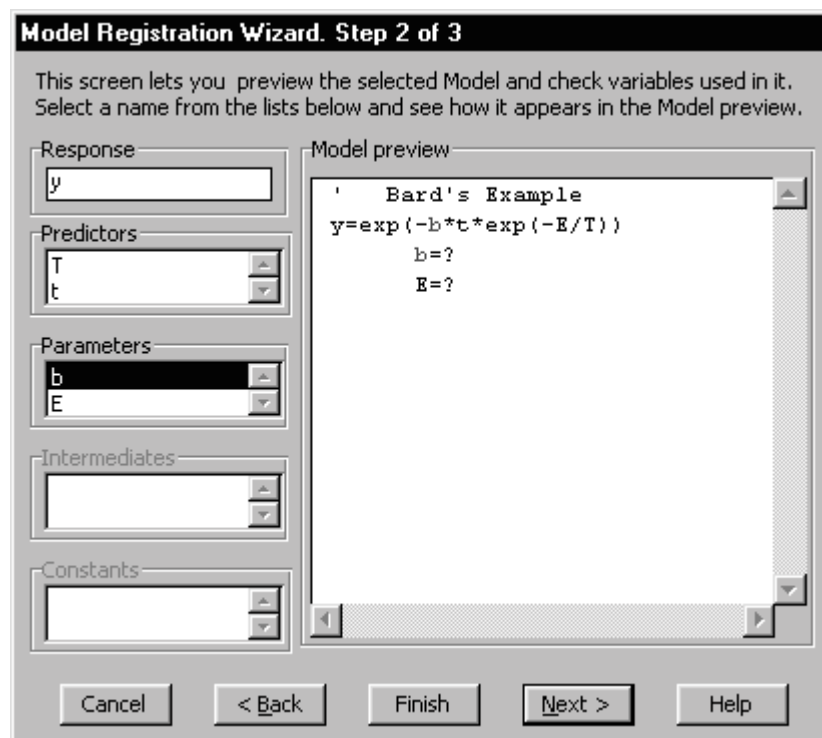


Рис. 5.29 Регистратор модели. Второй шаг.

Здесь в окошке просмотра модели (**Model preview**) можно увидеть регистрируемую модель и проверить, как и какие идентификаторы используются в ней. Для этого нужно выбрать имя из списков в левой части окна, тогда соответствующие места в окошке просмотра модели будут выделены красным шрифтом. Исправлять модель здесь нельзя.

Третий шаг Регистратора Модели показан на Рис. 5.30.

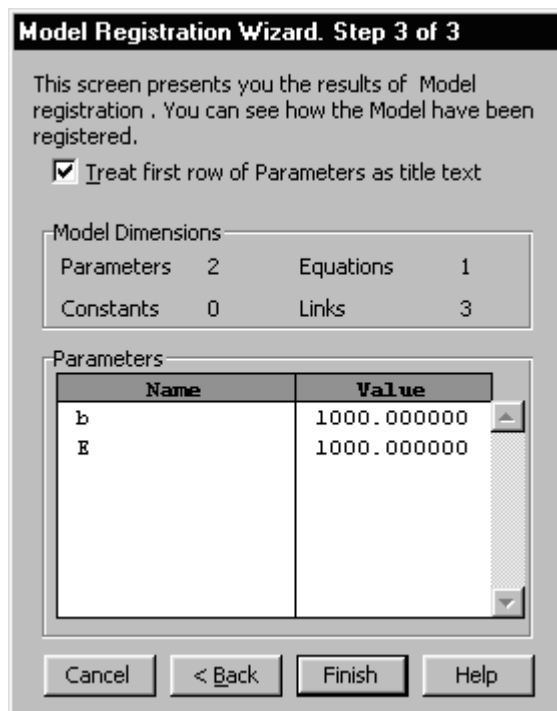



Рис. 5.30 Регистратор модели. Третий шаг.

Это окно представляет результаты регистрации модели. В нем имеется всего одна настройка.– **Treat first row of Parameters as title text** (*Считать первый ряд Таблицы Параметров титулом*). Изменить начальные значения параметров, показанные в окошке **Parameters**, здесь нельзя.

Регистратору Модели соответствуют две VBA функции `FitterModel` и `FitterParameters` (см. раздел 5.12).

5.10. Регистратор априорной информации

Регистратор Априорной информации можно вызвать из Главного Диалога (кнопка **Bayes...** на Рис. 5.6) или с помощью Панели инструментов (пятая кнопка  на Рис. 5.10). Регистрировать априорную информацию можно только после того, как будет зарегистрированы данные и модель. До этой поры соответствующие кнопки будут неактивны.

Этот регистратор состоит из трех шагов, первый из которых показан на Рис. 5.31.



Рис. 5.31 Регистратор апр. информации. Первый шаг.

В поле **Range** нужно ввести ссылку на всю область, занимаемую Таблицей априорной информации (см. Рис. 5.5). Если имеется полная априорная информация первого рода (см. Главу 2), то нужно также указать априорное число степеней свободы (**Freedom degrees**) (2.28) и априорную дисперсию (**Variance**) (2.27) В противном случае эти поля нужно оставить пустыми.

После того как априорная информация зарегистрирована, она хранится в памяти. Для того чтобы удалить ее оттуда, нужно вызвать Регистратор Априорной информации, очистить поле **Range** на первом шаге и нажать кнопку **Finish**.

Второй шаг Регистратора Априорной информации показан на Рис. 5.32.

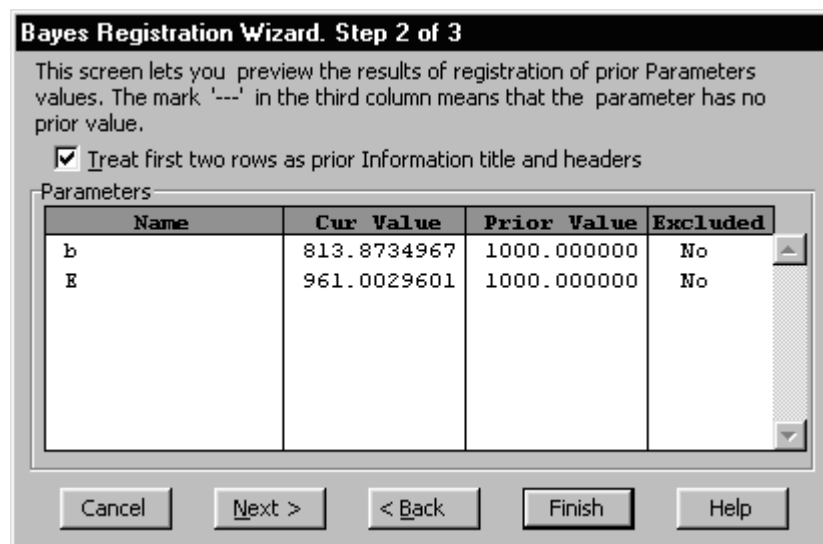


Рис. 5.32 Регистратор апр. информации. Второй шаг.

В нем имеется всего одна настройка.— **Treat first two rows as prior Information title and header** (считать первые два ряда Таблицы Априорной Информации титулом и заголовком). В нижнем окошке показаны результаты регистрации априорных параметров. Знак — во втором столбце означает, что параметр исключен из подгонки. Тот же знак в третьем

столбце означает, что у этого параметра нет априорного значения. Изменить значения параметров, показанные в окошке **Parameters**, здесь нельзя.

Третий шаг Регистратора Априорной информации показан на Рис. 5.33.

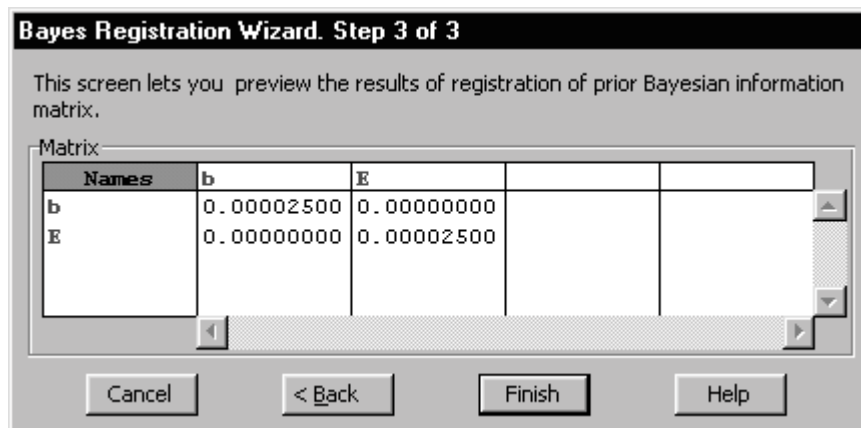


Рис. 5.33 Регистратор апр. информации. Третий шаг.

Это окно представляет результаты регистрации априорной информационной матрицы (2.26).

Регистратору Априорной информации соответствует VBA функция `FitterBayes` (см. раздел 5.12).

5.11. Диалог дополнительных действий

Диалог дополнительных действий можно вызвать только из Главного Диалога (кнопка **Plus...** на Рис. 5.6). На Панели инструментов соответствующей кнопки нет. Обратиться к этому диалогу можно только после того, как будет зарегистрированы данные и модель. До этой поры соответствующая кнопка будет неактивна.

Диалог изображен на Рис. 5.34. С его помощью можно выполнить несколько операций.

Simulate Response. Эта операция создания новых, модельных значений отклика, которые вычисляются на базе зарегистрированных данных и модели. Значения отклика включают псевдослучайные нормальные ошибки со *стандартным отклонением* (корень из дисперсии), которое задается в поле **Standard deviation**. Тип ошибки соответствует установленному в Регистраторе Настроек (см. Рис. 5.9). После нажатия кнопки **Simulate** результаты помещаются в столбец откликов Таблицы данных (Рис. 5.1). При этом предыдущие значения уничтожаются!

Этому действию соответствует VBA функция `FitterSimulate` (см. раздел 5.12).

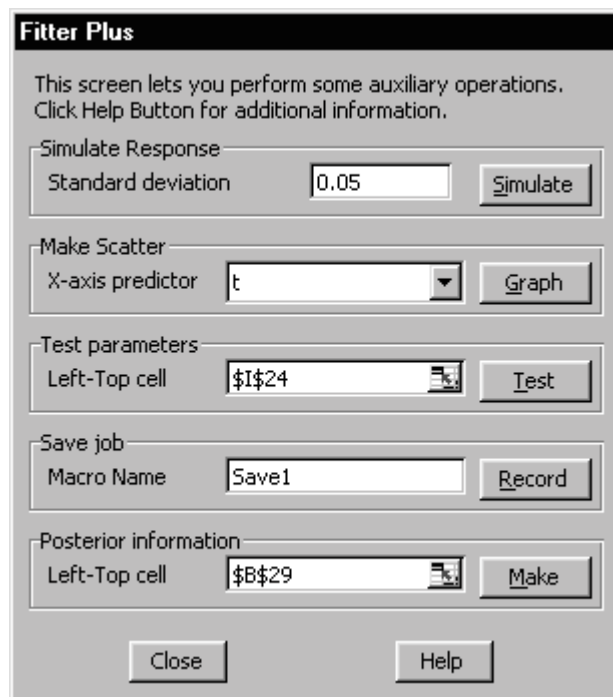


Рис. 5.34 Диалог дополнительных действий

Make Scatter. С помощью этой операции можно быстро *построить график* данных. Если в данных имеется несколько предикторов, то нужно выбрать тот предиктор, который будет использован для *оси абсцисс (X-axis predictor)* и нажать кнопку **Graph**. График всегда выводится на активный лист книги и носит имя *FitterChart*. Если нужно построить еще один график на том же рабочем листе, то сначала нужно переименовать старый, иначе он будет замещен новым.

Этому действию соответствует VBA функция `FitterGraph` (см. раздел 5.12).

Test parameters. Эта операция осуществляет проверку *альтернативных значений параметров* (см. раздел 1.4). В поле **Left-Top cell** нужно ввести ссылку на *левую верхнюю клетку* рабочего листа, где находятся альтернативные значения, и нажать кнопку **Test**. Результат выводится в новое окно.

Этому действию соответствует VBA функция `FitterAlternativeParameters` (см. раздел 5.12).

Save Job. Эта операция *запоминает текущую конфигурацию* и создает *макропроцедуру* (подпрограмму) системы Excel на языке VBA. Эта подпрограмма помещается в модуль с именем *SaveModule* в активной рабочей книге. Если модуля с таким именем не существует – он будет создан. Нужно в поле, озаглавленном **Macro Name**, установить *имя подпрограммы* и нажать кнопку **Record**. Будет создана подпрограмма, состоящая из VBA функ-

ций Fitter, которая будет восстанавливать все запомненные регистрации и настройки. Однако необходимо учесть, что в этой подпрограмме хранятся не сами значения величин, а только абсолютные адреса областей, в которых находились эти значения. Поэтому, если данные были потом перемещены или изменены, то результат ее действия может стать неверным.

Этому действию соответствует VBA функция `FitterSave` (см. раздел 5.12).

Posterior Information. Эта процедура создает *апостериорную информацию* (2.25)-(2.26) на основе зарегистрированных Данных и Модели. В поле **Left-Top cell** нужно ввести ссылку на левую верхнюю клетку рабочего листа, где будет создана информация, и нажать кнопку **M**ake. Результат выводится в форме Таблицы априорной информации (см. Рис. 5.5)

Этому действию соответствует VBA функция `FitterPosteriorInformation` (см. раздел 5.12).

5.12. Функции Fitter

Все действия, которые выполняются с помощью Диалогов и Регистраторов, могут быть осуществлены с помощью VBA [7] процедур. Для этой цели применяются функции Fitter. Используя их можно написать подпрограммы для обработки данных. Например, регистрация данных может быть выполнена следующей командой

```
FitterData(sDataRange:="$B$2:$F$21", sCodeString:="PRFI")
```

Здесь аргумент `sDataRange` – это ссылка на таблицу данных, второй аргумент `sCodeString` – это кодовая строка, объясняющая смысл столбцов в этой таблице.

Все результаты работы Fitter выводятся на рабочий лист. Однако иногда бывает удобнее только хранить их в памяти и получать по мере необходимости. Для этого также используются функции Fitter, осуществляющие обратную связь. Например, можно получить число предикторов в зарегистрированных данных, используя следующую команду

```
FitterDimensionsBack(Predictors:=nPred)
```

Здесь `nPred` – это искомое число.

Важно помнить о том, что прежде чем использовать эти функции в программах, необходимо *установить ссылку* на надстройку Fitter.xla в Microsoft Visual Basic. Для этого в активном модуле VBA надо выбрать **R**eferences в меню **T**ools, а затем отметить файл *Fitter.xla* в **A**vailable References.

В функциях Fitter применяются следующие символические константы

Табл. 5.5 Символические константы Fitter

| Имя | Численное значение | Логическое значение |
|------------|--------------------|---------------------|
| xtYes | -1 | True |
| xtNo | 0 | False |
| xtNone | 1 | - |
| xtAbsolute | 1 | - |
| xtRelative | 2 | - |

Для всех функций Fitter действует следующее правило – если возвращаемое значение равно нулю, то произошла ошибка. Для вывода этой ошибки используется функция `FitterError`.

FitterError

Выводит информацию о произошедшей ошибке.

Синтаксис

FitterError

Аргументы

Нет

FitterInit

Очищает все регистрации и настройки и перезапускает программу.

Синтаксис

FitterInit

Аргументы

Нет

FitterFit

Устанавливает основные настройки программы и осуществляет подгонку. Эта функция и ее аргументы соответствуют Главному Диалогу (Рис. 5.6).

Синтаксис

**FitterFit([sProgressLeftTop,] [sResultsLeftTop,]
[dFitConvergence,] [nPauseEachStep,] [nMaxIterations])**

Аргументы

sProgressLeftTop

Необязателен. Ссылка на одну клетку рабочего листа. Соответствует полю **Progress** в Главном Диалоге. Значение по умолчанию "".

sResultsLeftTop

Необязателен. Ссылка на одну клетку рабочего листа. Соответствует полю **Results** в Главном Диалоге. Значение по умолчанию "".

dFitConvergence

Необязателен. Число между 0.1 и $1.E-10$, которое задает точность подгонки. Соответствует полю **Convergence of fitting** в Главном Диалоге. Значение по умолчанию $1.E-5$

nPauseEachStep

Необязателен. Если **xtYes**, делается пауза на каждой итерации. Соответствует настройке **Pause at each iteration** в Главном Диалоге. Значение по умолчанию **xtNo**

nMaxIterations

Необязателен. Число между 0 и 32767 , которое определяет максимальное число итераций. Соответствует полю **Maximum number of iterations** в Главном Диалоге. Значение по умолчанию 100

FitterFitBack

Возвращает настройки, установленные в Главном Диалоге (Рис. 5.6) или функцией `FitterFit`.

Синтаксис

```
FitterFitBack( [sProgressLeftTop,] [sResultsLeftTop,]  
              [dFitConvergence,] [nPauseEachStep,]  
              [nMaxIterations] )
```

Аргументы

Аналогичны аргументам функции `FitterFit` с той разницей, что значения не устанавливаются, а возвращаются, например

sProgressLeftTop

Необязателен. Этот аргумент получает текущее значение ссылки на левую верхнюю клетку Таблицы хода поиска (Рис. 5.8). Соответствует полю **Progress** в Главном Диалоге

FitterOptions

Устанавливает настройки, которые соответствуют первым двум окнам Регистратора Настроек (Рис. 5.9 и Рис. 5.11).

Синтаксис

```
FitterOptions([kToolbar,] [dCalcPrecision,] [kErrorType,]
              [kFormat,] [kGeneral,] [kDeviations,] [kCovariances,]
              [kCorrelations,] [kFmatrix,] [kFitResults,]
              [dSignificance,] [kNonlinearity,]
              [kTestSeries,] [kTestOutliers,]
              [kTestBartlett,] [kTestFisher,] [dConfidence,]
              [kMethod,] [kLeftSide,] [kRightSide,] [kBootNumber])
```

Аргументы

kToolbar

Необязателен. Если **xtYes**, открывается *Панель инструментов Fitter* (Рис. 5.10). Соответствует настройке **View Fitter Toolbar** в первом окне Регистратора Настроек (Рис. 5.9). Значение по умолчанию **xtNo**.

dCalcPrecision

Необязателен. Число между *1.E-3* и *1.E-15*, которое определяет *точность вычислений*. Соответствует настройке **Calculation Precision** в первом окне Регистратора Настроек (Рис. 5.9). Значение по умолчанию *1.E-10*.

kErrorType

Необязателен. Может быть **xtAbsolute** или **xtRelative**. Устанавливает *тип ошибки*. Соответствует настройке **Error** в первом окне Регистратора Настроек (Рис. 5.9). Значение по умолчанию **xtAbsolute**.

kFormat

Необязателен. Если **xtYes**, то все *таблицы форматируются*. Соответствует настройке **Format all tables** в первом окне Регистратора Настроек (Рис. 5.9). Значение по умолчанию **xtNo**.

kGeneral

Необязателен. Если **xtYes**, то выводится *Таблица общей информации* (Рис. 5.12). Соответствует настройке **General** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию **xtNo**.

kDeviations

Необязателен. Если **xtYes**, то выводится *Таблица оценок параметров* (Рис. 5.13). Соответствует настройке **Deviation** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию **xtNo**.

kCovariances

Необязателен. Если **xtYes**, то выводится *матрица ковариаций* параметров (Рис. 5.14). Соответствует настройке **Covariance** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию **xtNo**.

kCorrelations

Необязателен. Если **xtYes**, то выводится *матрица корреляций* параметров (Рис. 5.15). Соответствует настройке **Correlation** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию **xtNo**.

kFmatrix

Необязателен. Если **xtYes**, то выводится F-матрица (Рис. 5.16). Соответствует настройке **F-matrix** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию **xtNo**.

kFitResults

Необязателен. Если **xtYes**, то выводится *Таблица результатов подгонки* (Рис. 5.17). Соответствует настройке **Fitting results** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию **xtNo**.

dSignificance

Необязателен. Число между *0.001* и *0.5*, которое определяет *уровень значимости*. Соответствует настройке **α** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию *0.05*.

kNonlinearity

Необязателен. Если **xtYes**, то вычисляется *коэффициент нелинейности* (Рис. 5.22). Соответствует настройке **Nonlinearity** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию **xtNo**.

kTestSeries

Необязателен. Если **xtYes**, то проводится *тест серий*. Соответствует настройке **Series** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию **xtNo**.

kTestOutliers

Необязателен. Если **xtYes**, то проводится *тест выбросов*. Соответствует настройке **Outliers** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию **xtNo**.

kTestBartlett

Необязателен. Если **xtYes**, то проводится *тест дисперсий*. Соответствует настройке **Variance** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию **xtNo**.

kTestFisher

Необязателен. Если **xtYes**, то проводится *тест адекватности*. Соответствует настройке **Lack of Fit** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию **xtNo**.

dConfidence

Необязателен. Число между *0.75* и *0.999*, которое определяет *уровень достоверности*. Соответствует настройке **P=** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию *0.95*.

kMethod

Необязателен. Если **xtYes**, то для построения доверительных интервалов используется *метод связанного моделирования*, а если (значение по умолчанию) **xtNo**, то *метод стохастической аппроксимации*. Соответствует настройке **Linearization** или **Bootstrap with** во втором окне Регистратора Настроек (Рис. 5.11).

kLeftSide

Необязателен. Если **xtYes**, то строится *левосторонний доверительный интервал*. Если (значение по умолчанию) **xtNo**, то не строится. Соответствует настройке **Left-side** во втором окне Регистратора Настроек (Рис. 5.11).

kRightSide

Необязателен. Если **xtYes**, то строится *правосторонний доверительный интервал*. Если (значение по умолчанию) **xtNo**, то не строится. Соответствует настройке **Right-side** во втором окне Регистратора Настроек (Рис. 5.11).

kBootNumber

Необязателен. Число между *1000* и *5000*, которое определяет *количество повторений* при моделировании. Соответствует настройке **N=** во втором окне Регистратора Настроек (Рис. 5.11). Значение по умолчанию *3000*.

FitterOptionsBack

Возвращает значения настроек, которые были установлены Регистратором Настроек в первом и втором окнах или функцией `FitterOptions`.

Синтаксис

```
FitterOptionsBack([kToolbar,] [dCalcPrecision,] [kErrorType,]
```

```
[kFormat,] [kGeneral,] [kDeviations,] [kCovariances,]  
[kCorrelations,] [kFmatrix,] [kFitResults,]  
[dSignificance,] [kNonlinearity,]  
[kTestSeries,] [kTestOutliers,]  
[kTestBartlett,] [kTestFisher,] [dConfidence,]  
[kMethod,] [kLeftSide,] [kRightSide,] [kBootNumber])
```

Аргументы

Аналогичны аргументам функции FitterOptions с той разницей, что значения не устанавливаются, а возвращаются, например

kToolbar

Необязателен. Этот аргумент получает текущее значение настройки *Панели инструментов* Fitter: **xtYes** или **xtNo**.

FitterStyles

Устанавливает стили, используемые при форматировании. Эта функция и ее аргументы соответствуют третьему окну Регистратора Настроек (Рис. 5.23).

Синтаксис

```
FitterStyles ( [sDataTitleStyle,] [sDataHeaderStyle,]  
              [sDataBodyStyle,] [sParamTitleStyle,]  
              [sParamHeaderStyle,] [sParamBodyStyle,]  
              [sBayesTitleStyle,] [sBayesHeaderStyle,]  
              [sBayesBodyStyle,] [sOutputTitleStyle,]  
              [sOutputHeaderStyle,] [sOutputBodyStyle,]  
              [sModelBodyStyle,] [sModelCommentStyle,]  
              [sModelParameterStyle] )
```

Аргументы

sDataTitleStyle

Необязателен. Стиль *титла Таблицы данных*. Значение по умолчанию **Data_Title**.

sDataHeaderStyle

Необязателен. Стиль *заголовка Таблицы данных*. Значение по умолчанию **Data_Header**.

sDataBodyStyle

Необязателен. Стиль *тела Таблицы данных*. Значение по умолчанию **Data_Body**.

sParamTitleStyle

Необязателен. Стиль *титула Таблицы параметров*. Значение по умолчанию **Param_Title**.

sParamHeaderStyle

Необязателен. Стиль *заголовка Таблицы параметров*. Значение по умолчанию **Param_Header**.

sParamBodyStyle

Необязателен. Стиль *тела Таблицы параметров*. Значение по умолчанию **Param_Body**.

sBayesTitleStyle

Необязателен. Стиль *титула Таблицы апр. информации*. Значение по умолчанию **Bayes_Title**.

sBayesHeaderStyle

Необязателен. Стиль *заголовка Таблицы апр. информации*. Значение по умолчанию **Bayes_Header**.

sBayesBodyStyle

Необязателен. Стиль *тела Таблицы апр. информации*. Значение по умолчанию **Bayes_Body**.

sOutputTitleStyle

Необязателен. Стиль *титула Таблиц результатов*. Значение по умолчанию **Output_Title**.

sOutputHeaderStyle

Необязателен. Стиль *заголовков Таблиц результатов*. Значение по умолчанию **Output_Header**.

sOutputBodyStyle

Необязателен. Стиль *тела Таблиц результатов*. Значение по умолчанию **Output_Body**.

sModelBodyStyle

Необязателен. Стиль *тела Текстового поля Модели*. Значение по умолчанию **Model_Body**.

sModelCommentStyle

Необязателен. Стиль *комментариев Модели*. Значение по умолчанию **Model_Comment**.

sModelParameterStyle

Необязателен. Стиль *параметров Модели*. Значение по умолчанию **Model_Parameter**.

FitterStylesBack

Возвращает имена стилей форматирования, которые были установлены Регистратором Настроек на третьем шагу (Рис. 5.23) или функцией `FitterStyles`.

Синтаксис

```
FitterStylesBack ( [sDataTitleStyle,] [sDataHeaderStyle,]  
[sDataBodyStyle,] [sParamTitleStyle,]  
[sParamHeaderStyle,] [sParamBodyStyle,]  
[sBayesTitleStyle,] [sBayesHeaderStyle,]  
[sBayesBodyStyle,] [sOutputTitleStyle,]  
[sOutputHeaderStyle,] [sOutputBodyStyle,]  
[sModelBodyStyle,] [sModelCommentStyle,]  
[sModelParameterStyle] )
```

Аргументы

Аналогичны аргументам функции `FitterStyles` с той разницей, что значения не устанавливаются, а возвращаются, например

sDataTitleStyle

Необязателен. Этот аргумент получает *имя* текущего стиля, используемого для форматирования *титла Таблицы данных*.

FitterData

Регистрирует данные. Эта функция и ее аргументы соответствуют Регистратору Данных (см. раздел 5.8).

Синтаксис

```
FitterData ( sDataRange [, sCodeString,] [ kTitle] )
```

Аргументы

sDataRange

Обязателен. Ссылка на *область Таблицы Данных* (Рис. 5.1) на рабочем листе. Соответствует полю **Range** в первом окне Регистратора Данных (Рис. 5.25).

sCodeString

Необязателен. *Кодированная строка*, содержащая информацию об устройстве Таблицы Данных (Рис. 5.1). Соответствует настройкам **Layout** во втором окне Регистратора Данных (Рис. 5.26). Длина этой строки должна быть равна числу столбцов в

Таблице Данных, а каждая буква определяет тип столбца так, как это указано в Табл. 5.6.

Табл. 5.6 Кодировка строки `sCodeString`

| Буква | Интерпретация столбца |
|-------|---|
| P | Предиктор (Predictor value) |
| R | Отклик (Response value) |
| W | Вес (Weight value) |
| F | Модель (Fitting value) |
| I | Промежуточная (Intermediate value) |
| C | Доверительный предел (Confidence limit) |
| V | Дисперсия (Variance value) |
| S | Игнорировать этот столбец (Skip this column) |

kTitle

Необязателен. Если **xtYes**, то первая строка Таблицы Данных (Рис. 5.1) интерпретируется как *титул*. Если **xtNo**, то – как заголовок. Если **xtNone**, то программа принимает решение самостоятельно. Соответствует настройке **Treat first row as Data title text** во втором окне Регистратора Данных (Рис. 5.26). Значение по умолчанию **xtNone**.

FitterDataBack

Возвращает настройки, которые были установлены Регистратором Данных или функцией `FitterData`.

Синтаксис

```
FitterDataBack ( [sDataRange] [, sCodeString,] [kTitle] )
```

Аргументы

Аналогичны аргументам функции `FitterData` с той разницей, что значения не устанавливаются, а возвращаются, например

sDataRange

Необязателен. Возвращает *адрес текущей Таблицы Данных* (Рис. 5.1). Соответствует полю **Range** в первом окне Регистратора Данных (Рис. 5.25).

FitterModel

Регистрирует модель. Эта функция и ее аргументы соответствуют Регистратору Модели (см. раздел 5.9).

Синтаксис

FitterModel (sModelTextBox)

Аргументы

sModelTextBox

Обязателен. Ссылка на *текстовое поле* модели (Рис. 5.2) на рабочем листе. Соответствует настройке **Text Box** в первом окне Регистратора Модели (Рис. 5.28).

FitterModelBack

Возвращает настройки, которые были установлены Регистратором Модели или функцией `FitterModel`.

Синтаксис

FitterModelBack (sModelTextBox)

Аргументы

sModelTextBox

Обязателен. Этот аргумент получает *имя текстового поля* (Рис. 5.2), в котором находится текущая модель. Соответствует настройке **Text Box** в первом окне Регистратора Модели (Рис. 5.28).

FitterParameters

Регистрирует параметры модели. Эта функция и ее аргументы соответствуют Регистратору Модели (см. раздел 5.9).

Синтаксис

FitterParameters ([sParametersLeftTop,] [kTitle])

Аргументы

sParametersLeftTop

Необязателен. Ссылка на *левую верхнюю клетку* Таблицы Параметров (Рис. 5.4). Соответствует настройке **Cell** в первом окне Регистратора Модели (Рис. 5.28). Значение по умолчанию ""

kTitle

Необязателен. Если **xtYes**, то первая строка Таблицы Параметров интерпретируется как *титул*. Если **xtNo**, то, как заголовок – имена параметров. Если **xtNone**, то программа принимает решение самостоятельно. Соответствует настройке **Treat first row of Parameters as title text** в третьем окне Регистратора Модели (Рис. 5.30). Значение по умолчанию **xtNone**.

FitterParametersBack

Возвращает настройки, которые были установлены Регистратором Модели или функцией `FitterParameters`.

Синтаксис

```
FitterParametersBack ( [sParametersLeftTop,] [kTitle] )
```

Аргументы

Аналогичны аргументам функции `FitterParameters` с той разницей, что значения не устанавливаются, а возвращаются, например –

sParametersLeftTop

Возвращает адрес *левой верхней клетки* текущей Таблицы Параметров (Рис. 5.4).

Соответствует настройке **Cell** в первом окне Регистратора Модели (Рис. 5.28).

FitterBayes

Регистрирует априорную информацию. Эта функция и ее аргументы соответствуют Регистратору Априорной Информации (см. раздел 5.10).

Синтаксис

```
FitterBayes ( sBayesRange [, lFreedom] [, dDispersion] [, kTitle] )
```

Аргументы

sBayesRange

Обязателен. Ссылка на *область Таблицы априорной информации* на рабочем листе.

Соответствует полю **Range** в первом окне Регистратора Априорной Информации (Рис. 5.31).

lFreedom

Необязателен. Число между 0 и 2147483647, которое определяет *априорное число степеней свободы*. Соответствует настройке **Freedom degrees** в первом окне Регистратора Априорной Информации (Рис. 5.31). Значение по умолчанию 0.

dDispersion

Необязателен. Число между 0 и 1.79769313486232E+308, которое определяет *априорное значение дисперсии* ошибки. Соответствует настройке **Variance** в первом окне Регистратора Априорной Информации (Рис. 5.31). Значение по умолчанию 0.

kTitle

Необязателен. Если **xtYes**, то первые две строки Таблицы априорной информации интерпретируется как *титул*. Если **xtNo**, то, как значения. Если **xtNone**, то программа принимает решение самостоятельно. Соответствует настройке **Treat first two**

rows as Prior Information title and headers во втором окне Регистратора Априорной Информации (Рис. 5.32). Значение по умолчанию **xtNone**.

FitterBayesBack

Возвращает настройки, которые были установлены Априорной Информации или функцией FitterBayes.

Синтаксис

```
FitterBayesBack ( [sBayesRange] [, lFreedom] [, dDispersion]  
                [, kTitle] )
```

Аргументы

Аналогичны аргументам функции FitterBayes с той разницей, что значения не устанавливаются, а возвращаются, например

sBayesRange

Необязателен. Возвращает *адрес области* текущей Таблицы априорной информации. Соответствует полю **Range** в первом окне Регистратора Априорной Информации (Рис. 5.31).

FitterSimulate

Создает и выводит модельные значения откликов. Эта функция и ее аргументы соответствуют операции **Simulate Response** в Диалоге дополнительных действий (Рис. 5.34)

Синтаксис

```
FitterSimulate( [sResultLeftTop,] [dDeviation] )
```

Аргументы

sResultLeftTop

Необязателен. Ссылка на *клетку* рабочего листа. Модельные значения откликов будут выведены в столбец, начинающийся с этой клетки. Соответствия с Диалогом дополнительных действий нет. Значение по умолчанию "" – результаты не выводятся.

dDeviation

Необязателен. Число между 0 и $1.79769313486232E+308$, которое определяет *стандартное отклонение* ошибки. Соответствует настройке **Standard deviation** в Диалоге дополнительных действий (Рис. 5.34). Значение по умолчанию 0.

FitterGraph

Строит график на активном листе. Эта функция и ее аргументы соответствуют операции **Make Scatter** в Диалоге дополнительных действий (Рис. 5.34)

Синтаксис

```
FitterGraph ( [nAxis] )
```

Аргументы

nAxis

Необязателен. Число между 0 и числом предикторов минус один. Обозначает *предиктор*, значения которого будут отложены на *оси абсцисс*. Соответствует настройке **X-axis predictor** в Диалоге дополнительных действий (Рис. 5.34). Значение по умолчанию 0 – первый предиктор.

FitterAlternativeParameters

Проверяет гипотезу об альтернативных значениях параметров. Эта функция и ее аргументы соответствуют операции **Test parameters** в Диалоге дополнительных действий (Рис. 5.34)

Синтаксис

```
FitterAlternativeParameters ( sAlterLeftTop, nResult )
```

Аргументы

sAlterLeftTop

Обязателен. Ссылка на *левую верхнюю клетку* рабочего листа, где находится таблица альтернативных значений параметров. Соответствует настройке **Left-Top cell** в Диалоге дополнительных действий.

nResult

Обязателен. Сюда выводится *результат проверки*: если **xtYes** – принята, если **xtNo** – отвергнута, и если **xtNone** – мало данных.

FitterPosteriorInformation

Строит апостериорную информацию. Эта функция и ее аргументы соответствуют операции **Posterior Information** в Диалоге дополнительных действий (Рис. 5.34)

Синтаксис

```
FitterPosteriorInformation( sResultLeftTop )
```

Аргументы

sResultLeftTop

Обязателен. Ссылка на *клетку* рабочего листа. Построенная информация будет выведены в таблицу, начинающуюся с этой клетки. Соответствует настройке **Left-Top cell** в Диалоге дополнительных действий.

FitterSave

Сохраняет текущее состояние в подпрограмме. Эта функция и ее аргументы соответствуют операции **Save Job** в Диалоге дополнительных действий (Рис. 5.34)

Синтаксис

```
FitterSave ( sMacroName )
```

Аргументы

sMacroName

Обязателен. *Имя подпрограммы*. Соответствует настройке **Macro Name** в Диалоге дополнительных действий.

FitterDimensionsBack

Возвращает текущие размерности зарегистрированных данных и модели.

Синтаксис

```
FitterDimensionsBack ( [Observations,] [Predictors,]  
[Intermediates,] [ConfLimits,]  
[Replicas,] [Parameters] )
```

Аргументы

Observations

Необязателен. Этот аргумент получает *число наблюдений* в зарегистрированной Таблице данных (Рис. 5.1), что соответствует числу ее строк (без строк титула и заголовка).

Predictors

Необязателен. Этот аргумент получает *число предикторов* в зарегистрированной Таблице данных (Рис. 5.1).

Intermediates

Необязателен. Этот аргумент получает *число промежуточных величин* в зарегистрированной Таблице данных (Рис. 5.1).

ConfLimits

Необязателен. Этот аргумент получает *число границ доверительных интервалов* (0, 1 или 2) в зарегистрированной Таблице данных (Рис. 5.1).

Replicas

Необязателен. Этот аргумент получает *число реплик* (см. раздел 1.4) в зарегистрированной Таблице данных (Рис. 5.1). Если это число больше нуля, то данные содержат реплики.

Parameters

Необязателен. Этот аргумент получает *число параметров* в зарегистрированной Модели, что соответствует числу строк (без строки титула) в Таблице параметров (Рис. 5.4).

FitterNamesBack

Возвращает имена, используемые в зарегистрированных данных и модели

Синтаксис

```
FitterNamesBack ( [Response,] [Predictors,] [Intermediates,]  
                 [Parameters] )
```

Аргументы

Некоторые аргументы являются массивами. При этом предполагается установка Option Base 1 в модуле VBA. Необходимо заранее зарезервировать достаточное число элементов этих массивов.

Response

Необязателен. Этот аргумент получает *имя отклика* в зарегистрированных Данных.

Predictors

Необязателен. Этот аргумент получает массив, содержащий все *имена предикторов* в зарегистрированных Данных..

Intermediates

Необязателен. Этот аргумент получает массив, содержащий *имена промежуточных переменных* в зарегистрированных Данных.

Parameters

Необязателен. Этот аргумент получает массив, содержащий *имена параметров* в зарегистрированной Модели, которые соответствует заголовкам в Таблице параметров (Рис. 5.4).

FitterStatisticsBack

Возвращает различные величины, вычисляемые программой в конце обработки данных. Аргументы этой функции соответствуют второму окну Регистратора Настроек (Рис. 5.11)

Синтаксис


```
FitterStatisticsBack ( [SumSquares,] [Objective,] [VarianceFit,]  
                      [,NDF_Fit,] [VariancePop,] [NDF_Pop,]  
                      [Spread,] [Estimates,] [Deviations,]  
                      [Covariances,] [Correlations,] [FMatrix] )
```

Аргументы

Некоторые аргументы являются массивами. При этом предполагается установка Option Base 1 в модуле VBA. Необходимо заранее зарезервировать достаточное число элементов этих массивах.

SumSquares

Необязателен. Этот аргумент получает величину *суммы квадратов отклонений*, соответствующую четвертой строке в Таблице результатов подгонки (Рис. 5.17)

Objective

Необязателен. Этот аргумент получает величину *целевой функции*, соответствующую второй строке в Таблице результатов подгонки (Рис. 5.17)

VarianceFit

Необязателен. Этот аргумент получает величину *дисперсии ошибки по регрессии*, соответствующую пятой строке в Таблице результатов подгонки (Рис. 5.17)

NDF_Fit

Необязателен. Этот аргумент получает величину *числа степеней свободы по регрессии*, соответствующую шестой строке в Таблице результатов подгонки (Рис. 5.17)

VariancePop

Необязателен. Этот аргумент получает величину *дисперсии ошибки по выборке*, соответствующую третьей строке в Таблице теста адекватности (Рис. 5.19)

NDF_Pop

Необязателен. Этот аргумент получает величину *числа степеней свободы по выборке*, которое на лист не выводится.

Spread

Необязателен. Этот аргумент получает величину *разброса собственных значений*, соответствующую последней строке в Таблице результатов подгонки (Рис. 5.17)

Estimates

Необязателен. Этот аргумент получает массив, содержащий значения *оценок параметров*, соответствующий третьему столбцу в Таблице оценок параметров (Рис. 5.13).

Deviations

Необязателен. Этот аргумент получает массив, содержащий значения *среднеквадратичных погрешностей оценок параметров*, соответствующий последнему столбцу в Таблице оценок параметров (Рис. 5.13).

Covariances

Необязателен. Этот аргумент получает двумерный массив, содержащий *матрицу ковариаций*, соответствующий Таблице матрицы ковариаций (Рис. 5.14).

Correlations

Необязателен. Этот аргумент получает двумерный массив, содержащий *матрицу корреляций*, соответствующий Таблице матрицы корреляций (Рис. 5.15).

FMatrix

Необязателен. Этот аргумент получает двумерный массив, содержащий F-матрицу ковариаций, соответствующий Таблице F-матрицы (Рис. 5.16).

FitterValuesBack

Возвращает различные величины, помещаемые программой в конце обработки данных в Таблицу Данных (Рис. 5.1).

Синтаксис

```
FitterValuesBack ( [Responses,] [Fittings,] [Weights,]  
                  [Predictors,] [Intermediates,] [Lefts,]  
                  [Rights,] [Variances,] [NDFs ] )
```

Аргументы

Некоторые аргументы являются массивами. При этом предполагается установка Option Base 1 в модуле VBA. Необходимо заранее зарезервировать достаточное число элементов этих массивах.

Response

Необязателен. Этот аргумент получает массив, содержащий значения всех *откликов* в зарегистрированной Таблице данных. (Рис. 5.1).

Fittings

Необязателен. Этот аргумент получает массив, содержащий вычисленные *значения модели*, даже если они не участвуют в зарегистрированной Таблице данных. (Рис. 5.1).

Weights

Необязателен. Этот аргумент получает массив, содержащий значения всех *весов*, даже если они не участвуют в зарегистрированной Таблице данных. (Рис. 5.1).

Predictors

Необязателен. Этот аргумент получает двумерный массив, содержащий значения всех *предикторов* в зарегистрированной Таблице данных. (Рис. 5.1).

Intermediates

Необязателен. Этот аргумент получает двумерный массив, содержащий значения всех *промежуточных переменных* в зарегистрированной Таблице данных. (Рис. 5.1).

Lefts

Необязателен. Этот аргумент получает массив, содержащий значения всех *левых границ доверительных интервалов* в зарегистрированной Таблице данных. (Рис. 5.1).

Rights

Необязателен. Этот аргумент получает массив, содержащий значения всех *правых границ доверительных интервалов* в зарегистрированной Таблице данных. (Рис. 5.1).

Variances

Необязателен. Этот аргумент получает массив, содержащий значения всех *частных дисперсий* в зарегистрированной Таблице данных. (Рис. 5.1).

NDFs

Необязателен. Этот аргумент получает массив, содержащий значения всех *частных степеней свободы дисперсий* (1.57) для данных из зарегистрированной Таблицы данных. (Рис. 5.1). Заметим, однако, что в самой Таблице такого столбца нет.

FitterTestsBack

Возвращает результаты проверок гипотез. Аргументы этой функции соответствуют второму окну Регистратора Настроек (Рис. 5.11)

Синтаксис

```
FitterTestsBack ( [Outliers,] [LackOfFit,] [Series,]  
                 [Variances,] [Nonlinearity] )
```

Аргументы

Outliers

Необязателен. Этот аргумент получает результат проверки *теста выбросов* (Рис. 5.18). Он может быть

- xtYes** – гипотеза принята,
- xtNo** – гипотеза отвергнута,
- xtNone** – тест не проводился или мало данных.

LackOfFit

Необязателен. Этот аргумент получает результат проверки *теста адекватности* (Рис. 5.19). Результат может быть тот же, что и в тесте выбросов.

Series

Необязателен. Этот аргумент получает результат проверки *теста серий* (Рис. 5.21).
Результат может быть тот же, что и в тесте выбросов.

Variances

Необязателен. Этот аргумент получает результат проверки *теста дисперсий* (Рис. 5.20). Результат может быть тот же, что и в тесте выбросов.

Nonlinerarity

Необязателен. Этот аргумент получает значение *коэффициента нелинейности*, соответствующий второй строке Таблицы приведенной на Рис. 5.22. Если результат равен единице, это означает, что коэффициент не рассчитывался.

Пример программы с функциями Fitter

Это простая программа автоматически выполняет все необходимые настройки и регистрации, а затем проводит обработку данных.

```
Option Explicit
Sub BardExample()
Dim nRes As Integer
Worksheets("Sheet1").Activate

'Инициализация
If (FitterInit() = 0) Then GoTo ErrorEnd

'Установка стилей
If (FitterStyles(sDataTitleStyle:="Data_Title", _
sDataHeaderStyle:="Data_Header", _
sModelParameterStyle:="Model_Parameter") = 0) _
Then GoTo ErrorEnd

'Установка настроек
If (FitterOptions(kToolBar:=xtYes, dCalcPrecision:=0.0000000001, _
kErrorType:=xtAbsolute, kFormat:=xtYes, _
kGeneral:=xtYes, kDeviations:=xtYes, _
kCovariances:=xtYes, kCorrelations:=xtYes, _
kFmatrix:=xtYes, kFitResults:=xtYes, _
kTestSeries:=xtYes, kTestOutliers:=xtYes, _
kLeftSide:=xtYes, kRightSide:=xtNo, kNonlinearity:=xtYes, _
kBootNumber:=1000) = 0) Then GoTo ErrorEnd

'Регистрация данных
If (FitterData(sDataRange:="$B$2:$G$21", _
sCodeString:="PPRWFC", _
kTitle:=xtYes) = 0) Then GoTo ErrorEnd

'Регистрация модели
If (FitterModel(sModelTextBox:="Sheet1!Bard model") = 0) _
Then GoTo ErrorEnd

'Регистрация параметров
If (FitterParameters(sParametersLeftTop:="$I$2", _
kTitle:=xtYes) = 0) Then GoTo ErrorEnd
```

Описание программы Fitter

```
'Регистрация априорной информации
  If (FitterBayes(sBayesRange:="$B$24:$F$27", _
    lFreedom:=0, _
    dDispersion:=0, kTitle:=xtYes) = 0) _
    Then GoTo ErrorEnd

'Подгонка
  If (FitterFit(sProgressLeftTop:="$I$6", _
    sResultsLeftTop:="$L$2", dFitConvergence:=0.00001, _
    nPauseEachStep:=xtNo, nMaxIterations:=100) = 0) _
    Then GoTo ErrorEnd

'Построение графика
  If (FitterGraph(kAxis:=1) = 0) Then GoTo ErrorEnd

'Создание модельных откликов
  If (FitterSimulate(sResultLeftTop:="$J$31", _
    dDeviation:=0.005) = 0) _
    Then GoTo ErrorEnd

'Проверка альтернативных параметров
  If (FitterAlternativeParameters(sAlterLeftTop:="$I$24", _
    nResult:=nRes) = 0) Then GoTo ErrorEnd
  If (nRes = xtNo) Then MsgBox "Hypothesis is rejected"

'Построение постериорной информации
  If (FitterPosteriorInformation(sResultLeftTop:="$B$29") = 0) _
    Then GoTo ErrorEnd

'Запись макро
  If (FitterSave(sMacroName:="BardMacro") = 0) _
    Then GoTo ErrorEnd
  Exit Sub

'Обработка ошибок
ErrorEnd:
  FitterError

End Sub
```

5.13. Результаты главы 5

В этой главе было дано описание компьютерной системы Fitter – надстройки для программы Excel. С ее помощью можно проводить регрессионную обработку данных и выполнять статистический анализ полученных закономерностей. Устройство системы Fitter соответствует современным требованиям к подобным программам.

Вся входная информация задается на листах рабочей книги Excel. Она состоит из трех частей. Первая часть – это таблица *данных*, состоящая из численных значений и имен. Вторая часть входной информации – это *модель*, являющаяся текстовым полем, в котором записана система уравнений, связывающих имена данных и параметров. Эти две части необходимы для начала работы. Третья часть – таблица *априорной информации* является необязательной. Она включает априорные значения параметров, информационную матрицу

и прочие элементы. Были даны объяснения того, как устроены все эти компоненты, какие правила следует соблюдать при их создании.

Для того чтобы объяснить программе, какая информация будет использована при обработке, ее нужно *зарегистрировать* с помощью *Регистраторов*. Для задания регрессионной задачи применяются три регистратора: *Данных*, *Модели* и *Априорной Информации*. Кроме того, имеется регистратор настройки режимов работы программы – *Регистратор Настроек* и *Диалог дополнительных возможностей*. Все они вызываются из *Главного диалога Fitter*. Показано, как эти инструменты используются для задания входной и выходной информации, выбора способа интерпретации и представления результатов.

Все действия, которые выполняются с помощью Диалогов и Регистраторов, могут быть осуществлены и с помощью VBA процедур. Для этого применяются функции Fitter. Используя их можно написать подпрограммы для ввода и обработки данных, получить результаты моделирования. Приведено подробное описание всех функций и приведен пример.

Система Fitter – это мощный, современный статистический инструмент, созданный для регрессионного анализа сложных данных. Он систематически используется в этой работе для решения всех теоретических и прикладных задач. Не смотря на свою новизну, она уже применяется в некоторых научных учреждениях – Алтайский государственный университет (г. Барнаул), ФГУП НИИ Приборов (г. Лыткарино МО), Институт неметаллических материалов СО РАН (г. Якутск) и Институт химической физики РАН (г. Москва). Система представлялась и демонстрировалась на международных и российских конференциях.

Приложения нелинейного регрессионного анализа к практическим задачам

В этой части работы собраны несколько задач, показывающих как нелинейный регрессионный анализ, в общем, и система Fitter, в частности, применяются при решении практических задач анализа экспериментальных физико-химических данных. Эти примеры демонстрируют основные приемы и методы, используемые при обработке данных.

Большинство моделей, использованных для описания данных, не является оригинальными. Они либо хорошо известны, либо создавались совместно с коллегами, имеющими более глубокие знания в областях, где эти задачи актуальны. Что является оригинальным, так это математические подходы, которые использовались для анализа и подгонки этих моделей. Во всех случаях были предложены новые приемы и методы, позволившие достичь более высокой точности, извлечь больше скрытой информации при обработке экспериментальных данных. При этом иногда удавалось получать результаты, которые являются недостижимыми при применении традиционных методов.

Подбор примеров для этой части работы базировался, в основном, на соображениях методического характера. Эти примеры должны продемонстрировать весь спектр приемов, используемых в нелинейном моделировании. С другой стороны, эти же примеры призваны продемонстрировать, широту областей приложения, начиная с традиционных методов аналитического анализа (глава 10), заканчивая сложнейшими проблемами обработки данных, полученных с помощью современных инструментальных методов (глава 7). Порядок следования разобранных примеров подчинен, прежде всего, принципу «от простого к сложному» в применении системы Fitter. Поэтому первые примеры, собранные в главе 6 являются наиболее простыми (с точки зрения использованных приемов нелинейного регрессионного анализа), и, в то же время, наиболее «поучительными» с методической точки зрения. Они разбираются максимально подробно. Дальнейшие примеры расположены по нарастанию сложности.

В главе 6 собраны три примера, объединенные общей темой – обработка термограмм (ТГА, ТМА и ДСК) полимеров. Они решались в сотрудничестве с Е.В. Быстрицкой и О.Е. Родионовой [32, 50-53, 96, 102]. Эти примеры показывают, что часто результаты, полученные при обработке исходных данных, служат затем новыми данными для построения

следующей регрессионной модели. Во всех этих задачах активно используется метод последовательного байесовского оценивания (ПБО).

Глава 7 посвящена актуальной задаче оценки кинетических параметров по спектральным данным. Эта задача решалась в полемике с A.Smilde [113, 116] – автором многих подобных методов. Метод последовательного байесовского оценивания, примененный к этой проблеме, позволил получить результаты с высокой точностью. С методической точки зрения пример, рассмотренный в этой главе, важен, прежде всего, потому, что он показывает, как проводится обработка однородных данных в задачах с большим количеством параметров.

Другую сторону метода ПБО раскрывает задача, разобранный в главе 8, где демонстрируется, как обрабатываются разнородные экспериментальные данные, описываемые разными моделями, зависящими от общих параметров. Это сделано на примере практически важной задачи прогнозирования старения эластомерных материалов. Модели, описывающие изменение механических свойств резин, использованные в этой главе, были разработаны Е.В. Быстрицкой [44-46].

Следующая глава 9 посвящена проблемам моделирования нормальной и аномальной диффузии. Хорошо известно [129, 130], что модели, описывающие эти процессы, имеют сложную математическую форму и требуют специальных усилий по их программированию. Прямыми выкладками удалось получить точные и достаточно удобные формулы для расчета кинетики сорбции в нефиковских моделях релаксационной и конвекционной диффузии, а также кинетики цикла «увлажнение-сушка». Оригинальная форма этих моделей позволяет использовать их в системе Fitter, что радикально облегчает процесс подбора параметров. Материал, изложенный в этой главе, показывает, что разработанное программное обеспечение может легко оперировать с очень большими и сложными моделями. Практическое приложение эти методы нашли в работах О.В. Старцева [136-138].

Глава 10 стоит несколько в стороне от главной проблематики работы. В ней рассматривается задача обработки данных титрования, актуальная для аналитической химии. Модели, использованные в этой главе, были разработаны многими авторами, в частности Б.М. Марьяновым [154-158].

В последней главе 11 показано, как, используя разработанное программное обеспечение – систему Fitter, можно создавать стандартные программы-шаблоны, которые автоматически строят модели и проводят всю необходимую обработку в скрытом от оператора режи-

ме. Это продемонстрировано на примере оценки эффективности ингибиторов хемиллюминесцентным методом, который решался в сотрудничестве с О.Н. Карпухиным.

6. Анализ термограмм полимеров

В этой главе приводятся несколько примеров приложения нелинейного регрессионного анализа к исследованию свойств полимеров. На этих простых задачах мы покажем, как применять систему Fitter для моделирования экспериментальных данных и прогнозирования на условия, отличные от условий эксперимента.

Все приложения, описанные в этой главе, относятся к количественному описанию термограмм полимеров. Под термограммой мы понимаем любой сигнал, являющийся функцией от переменной, линейно растущей со временем температуры. К термограммам относятся экспериментальные данные, полученные методами термогравиметрического анализа (ТГА), термомеханического анализа (ТМА), дифференциальной сканирующей калориметрии (ДСК), и ряд других. Если в полимере при нагреве протекают необратимые процессы, его термограмма зависит от скорости нагрева. Тогда моделирование сигнала основано на системе дифференциальных кинетических уравнений с аррениусовской зависимостью параметров, которые интегрируются при условии постоянной скорости нагрева.

Рассмотренные в этой главе примеры представляют интерес с методической точки зрения, так как на них можно проследить технику практического применения нелинейного регрессионного анализа. Часто результаты, полученные при обработке исходных данных, служат затем данными для построения следующей регрессионной модели и т.д. шаг за шагом. Первый разобранный пример – наиболее простой, т.к. в нем окончательное решение находится за один шаг. Во втором примере присутствуют уже два, а в третьем – три таких шага. Кроме того, во всех этих задачах активно используется метод последовательного байесовского оценивания.

6.1. Прогнозирование старения ПВХ методом ТМА

Этот пример опубликован в [32, 96, 128]. Тепловое старение ПВХ материалов является сложным физико-химическим процессом, включающим три основных механизма [97]: десорбция пластификатора, автокаталитическое дегидрохлорирование и цепное окисление. Однако, известно [98-100], что при температурах ниже 100С основным процессом, определяющим изменение свойств ПВХ-пластиков, является десорбция пластификатора.

Процесс десорбции протекает, вообще говоря, в две стадии. Сначала происходит диффузия пластификатора из объема к поверхности, а затем удаление с поверхности (при десорбции в газовую среду – испарение). Какая из этих стадий будет лимитирующей, зави-

сит от толщины изделия и температуры. Из данных, приведенных в [101] следует, что во всем температурном диапазоне эксплуатации, диффузия не является лимитирующей стадией, а процесс десорбции пластификатора из ПВХ пластиката лимитируется испарением, как при эксплуатации, так и при ускоренных испытаниях.

Поскольку исследуемый процесс представляет собой десорбцию пластификатора, которая сопровождается соответствующим уменьшением массы, наиболее подходящим инструментом исследования является метод термогравиметрического анализа (ТГА) [103]. Этот метод, во-первых, обеспечивает высокую точность измерения массы при минимальных размерах образцов и, во-вторых, позволяет проводить измерения в динамическом режиме (при постоянной скорости нагрева), что существенно сокращает время испытаний.

На рисунке Рис. 6.1 представлены типичные экспериментальные данные (точки), полученные методом термогравиметрического анализа (ТГА) для образцов ПВХ пластиката с начальной концентрацией пластификатора $C_0=30\%$. Показано относительное изменение массы $y=m/m_0$, где m и m_0 – это текущая и исходная массы, соответственно, описанное моделью (кривая **1**, левая ось ординат), которая объясняется ниже и рост температуры в ТГА эксперименте (прямая **2**, правая ось ординат).

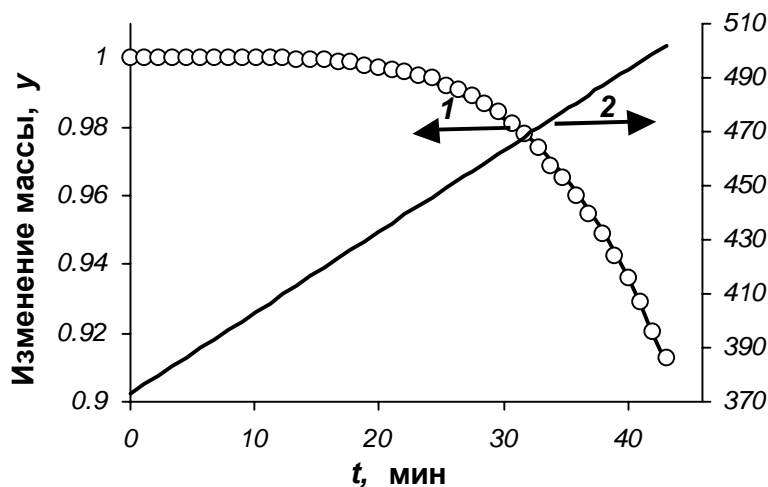


Рис. 6.1 Данные ТГА (O), модель (1) и рост температуры (2).

Обычно, когда описывается кинетика испарения неполимерных компонент из полимера, предполагается [129], что скорость испарения пропорциональна концентрации этого компонента:

$$W_e = kC \quad (6.1)$$

где k – константа скорости испарения, C – концентрация пластификатора. Тогда уравнение для описания изменения величины $y=m/m_0$, имеет вид

$$\frac{dy}{dt} = -k \cdot C, \quad y(0) = y_0 \quad (6.2)$$

Следует учесть, что потеря пластификатора сопровождается соответствующим уменьшением объема

$$C = 1 - \frac{1 - C_0}{y}, \quad (6.3)$$

где C_0 – исходная концентрация пластификатора. Константа скорости испарения k в уравнении (6.2) пропорциональна удельной поверхности образца F и зависит от температуры T по закону Аррениуса

$$k = k_0 \cdot F \cdot e^{-E/RT}, \quad (6.4)$$

где R – универсальная газовая постоянная.

Наконец, температура T в ТГА эксперименте растет со временем по линейному закону

$$T = T_0 + \nu t, \quad (6.5)$$

где ν скорость нагрева.

Модель для ТГА сигнала y представляет собой дифференциальное уравнение

$$\frac{dy}{dt} = -k \cdot \left(1 - \frac{1 - C_0}{y} \right) \quad (6.6)$$

которое получается подстановкой (6.3) в (6.2). В качестве независимой переменной используется время (а не температура, как это принято в ТГА), чтобы упростить процедуру прогноза.

Система уравнений (6.4)-(6.6) составляет математическую модель, которая описывает ТГА кривые для десорбции пластификатора из ПВХ. С регрессионной точки зрения (см. раздел 1.1) значение y является откликом, значения t , C_0 , F , ν и T_0 – независимыми переменными (предикторами), значения y_0 , k_0 и E – неизвестными параметрами. Начальная концентрация пластификатора C_0 измеряется независимым методом (хроматографически) и подставляется в уравнение (6.3) как условие эксперимента. Удельная поверхность F в уравнении (6.4) – это отношение площади поверхности образца к его объему. Эта величина может быть определена простым геометрическим расчетом. Например, для образцов, имеющих форму длинного цилиндра (изоляция электропроводов),

$$F = \frac{2R}{R^2 - r^2}$$

где R и r – соответственно, внешний и внутренний радиусы изоляции.

Эта система не имеет аналитического решения, что типично для всех динамических методов с изменяющейся температурой. Как правило, решение может быть выражено только с помощью интегральной экспоненты [143]

$$E_n(z) = \int_z^{\infty} t^{-n} e^{-zt} dt, \quad n = 0, 1, 2, \dots; \quad z > 0 \quad (6.7)$$

что очень неудобно для практических расчетов. Однако с помощью программы Fitter можно выполнить регрессионный анализ модели, заданной дифференциальным уравнением. Подробности всех расчетов можно найти в [126], где приведен файл TGA.XLS, содержащий данные и порядок расчетов.

На Рис. 6.2 показаны часть экспериментальных данных, модель и оценки параметров в том виде, в каком они представлены на рабочем листе Excel и используются в программе Fitter. Рассмотрим, как устроена эта страница.

| Data | | | | | | | | | Parameters | |
|------|-----|-----|------|-------|-------|---|-------|-------|---|----------|
| v | T0 | C0 | F | t | y | w | f | Left | Parameter | Value |
| 3 | 373 | 0.3 | 2.43 | 0 | 1 | 1 | 1.001 | 1 | y0 | 1.000905 |
| 3 | 373 | 0.3 | 2.43 | 1.15 | 1 | 1 | 1.001 | 1 | k0 | 13.96805 |
| 3 | 373 | 0.3 | 2.43 | 2.25 | 1 | 1 | 1.001 | 1 | E | 18024.43 |
| 3 | 373 | 0.3 | 2.43 | 3.4 | 1 | 1 | 1.001 | 1 | 'TGA desorbtion model D[y]/D[t]=-k*[1-(1-C0)/y];y(0)=y0 k=F*exp(k0-E/T/R) T=T0+v*t R=1.98717 y0=? k0=? E=? | |
| 3 | 373 | 0.3 | 2.43 | 4.55 | 1 | 1 | 1.001 | 1 | | |
| 3 | 373 | 0.3 | 2.43 | 5.65 | 1 | 1 | 1.001 | 1 | | |
| 3 | 373 | 0.3 | 2.43 | 6.75 | 1 | 1 | 1.001 | 1 | | |
| 3 | 373 | 0.3 | 2.43 | 7.85 | 1 | 1 | 1 | 1 | | |
| 3 | 373 | 0.3 | 2.43 | 9 | 1 | 1 | 1 | 1 | | |
| 3 | 373 | 0.3 | 2.43 | 10.1 | 1 | 1 | 1 | 1 | | |
| 3 | 373 | 0.3 | 2.43 | 11.2 | 1 | 1 | 1 | 1 | | |
| 3 | 373 | 0.3 | 2.43 | 12.3 | 1 | 1 | 1 | 0.999 | | |
| 3 | 373 | 0.3 | 2.43 | 13.4 | 0.999 | 1 | 1 | 0.999 | | |
| 3 | 373 | 0.3 | 2.43 | 14.5 | 0.999 | 1 | 0.999 | 0.999 | | |
| 3 | 373 | 0.3 | 2.43 | 15.6 | 0.999 | 1 | 0.999 | 0.999 | | |
| 3 | 373 | 0.3 | 2.43 | 16.7 | 0.999 | 1 | 0.999 | 0.998 | | |
| 3 | 373 | 0.3 | 2.43 | 17.75 | 0.999 | 1 | 0.998 | 0.998 | | |
| 3 | 373 | 0.3 | 2.43 | 18.85 | 0.998 | 1 | 0.997 | 0.997 | | |
| 3 | 373 | 0.3 | 2.43 | 19.95 | 0.997 | 1 | 0.997 | 0.996 | | |
| 3 | 373 | 0.3 | 2.43 | 21.05 | 0.997 | 1 | 0.996 | 0.996 | | |
| 3 | 373 | 0.3 | 2.43 | 22.1 | 0.996 | 1 | 0.995 | 0.995 | | |
| 3 | 373 | 0.3 | 2.43 | 23.2 | 0.995 | 1 | 0.994 | 0.994 | | |
| 3 | 373 | 0.3 | 2.43 | 24.25 | 0.994 | 1 | 0.993 | 0.992 | | |
| 3 | 373 | 0.3 | 2.43 | 25.35 | 0.992 | 1 | 0.991 | 0.991 | | |
| 3 | 373 | 0.3 | 2.43 | 26.4 | 0.991 | 1 | 0.99 | 0.989 | | |

| Parameters summary | | | | |
|--------------------|---------|----------|-----------|-------|
| Name | Initial | Final | Deviation | % |
| y0 | 1 | 1.000905 | 0.000206 | 0.02% |
| k0 | 10 | 13.96805 | 0.246881 | 1.77% |
| E | 10000 | 18024.43 | 231.7135 | 1.29% |

| Fitting results | |
|-------------------------|----------|
| Initial objective value | 2.528929 |
| Final objective value | 2.53E-05 |
| Initial sum of squares | 2.528929 |
| Final sum of squares | 2.53E-05 |
| Error variance | 6.65E-07 |
| Degrees of freedom | 38 |
| Spread of eigenvalue | 12 |

Рис. 6.2 Данные и модель ТГА эксперимента

В ее левой части расположена таблица данных (см. раздел 5.2), из которой показана только верхняя часть. В первых пяти ее столбцах указаны значения всех предикторов: скорости нагрева (**v**), начальной температуры (**T0**), начальной концентрации пластификатора (**C0**), удельной поверхности (**F**) и времени (**t**). Первые четыре предиктора в эксперименте

неизменны, а меняется только время. В следующих столбцах приведены экспериментальные значения отклика (y) и веса (w), которые в этой части таблицы равны 1. Оставшиеся два столбца предназначены для величин, которые не вводятся, а вычисляются в ходе работы программы. Это значения модели (f) и величина левой (нижней) границы доверительного интервала (**Left**). Последний столбец введен для прогноза на условия эксплуатации и его роль будет ясна далее.

Рассмотрим теперь устройство модели, которая записана в текстовом поле в правой части страницы (см. раздел 5.3). Первая строка является комментарием и поэтому она отмечена апострофом. Вторая строка в точности соответствует уравнению (6.6). Важно, что начальное условие $y(0) = y_0$ записано в той же строке и отделено от основного уравнения точкой с запятой. В третьей строке записано уравнение Аррениуса для константы скорости, которое отличается от формулы (6.4) только тем, что предэкспоненциальный множитель k_0 внесен в показатель экспоненты. Это сделано для уменьшения мультиколлинеарности при оценивании, в соответствии с общими принципами, изложенными в разделе 4.4. Четвертая строка в точности соответствует формуле (6.5), описывающей рост температуры. Затем определяется константа универсальной газовой постоянной, и перечисляются все неизвестные параметры; каждый в отдельной строке. Важно то, что при записи этой модели мы использовали те же обозначения – v , T_0 , C_0 , F , t и y , что и в таблице данных. А вот порядок в котором записаны строки в текстовом поле модели как раз не важен – он мог быть и другим.

Выше модели находится таблица параметров (см. раздел 5.4). В ней перечислены те же параметры y_0 , k_0 и E , что указаны в модели как неизвестные. Значения, приведенные рядом, соответствуют оценкам, определенным программой по экспериментальным данным. Начальное условие для уравнения (6.6) строго говоря, равно 1. Однако, в реальном эксперименте, как это видно на Рис. 6.1, могут наблюдаться небольшие отклонения, связанные со случайными колебаниями приборного нуля. Поэтому в модели для обработки данных начальное условие считается неизвестным.

Наконец, на странице в правом нижнем углу показаны таблицы оценок параметров и результатов подгонки, которые создаются программой. В верхней таблице (см. раздел 5.7) приведены начальные, конечные значения параметров и их среднеквадратичные отклонения. В последнем столбце те же отклонения указаны в процентах от среднего значения (коэффициент вариации). Точность оценивания энергии активации E составляет около 1.3%, а константы k_0 - около 1.8%. Фоновый параметр y_0 близок к единице, хотя и не ра-

вен ей. Во второй таблице приведены некоторые результаты оценивания. Полученные оценки дисперсии ошибки не превышают сотой долей процента, что согласуется с паспортными данными прибора. В целом, не смотря на большой разброс собственных значений – 12 порядков, качество обработки вполне удовлетворительное.

Полученное решение может быть использовано для прогнозирования срока службы изделий, сделанных из данного полимера. Известно [98], что отказ изделия из ПВХ пластиката связан с потерей пластификатора. Эксперименты показывают, что концентрация пластификатора $C_c=0.12$ является критической, так что время эксплуатации может быть определено как момент, когда концентрация падает до этого критического значения. Экстраполяция была сделана для следующих значений независимых переменных: время t в интервале 0–25 лет; исходная концентрация пластификатора C_0 равна 0.35; форм-фактор F определяется размерами объекта (1.5 или 2.0); скорость нагрева $v=0$; температура T_0 равна температуре эксплуатации (293K или 303K). На Рис. 6.3 показано, как это делается в системе Fitter.

Этот рисунок представляет нижнюю часть того же рабочего листа, который был показан на Рис. 6.2. В левой части видно продолжение таблицы данных. Заметим некоторые характерные особенности тех данных, которые расположены ниже пунктирной линии. Это данные для прогноза на условия эксплуатации, которые состоят из двух серий, соответствующих условиям, перечисленные выше: скорость нагрева $v=0$, температура $T_0=293$ (или 303) и т.д. Важно, что время приведено в часах, а также то, что все веса прогнозных данных равны нулю. Это сделано потому, что у нас нет измерений отклика в условиях прогноза (соответствующий столбец y – пуст). В следующем столбце (\hat{y}) помещаются модельные значения отклика, вычисленные программой для соответствующих условий при найденных оценках неизвестных параметров – прогноз. Однако необходимо учитывать неточность оценок параметров. Это делается с помощью построения 95%-го доверительного интервала для прогнозируемого значения отклика. Соответствующие значения, построенные методом стохастической линеаризации (S -метод, см. раздел 3.1), приведены в последнем столбце таблицы данных (Left).

Значения, представленные в таблице данных, соответствуют измеряемому (и прогнозируемому) отклику $y=m/m_0$, показывающему относительное изменение массы, тогда, как критическое значение известно нам для концентрации пластификатора. Однако, используя уравнение (6.3) можно найти соответствующие прогнозируемые значения концентрации. Они приведены в небольшой таблице в правом нижнем углу Рис. 6.3. Здесь время (t) вы-

ражено в годах, следующей столбец (C) соответствует ожидаемой, а последний (Cleft) – доверительной концентрации пластификатора при прогнозе. Теперь, сравнивая эти величины с критическим значением концентрации $C_c=0.12$ можно определить сроки эксплуатации.

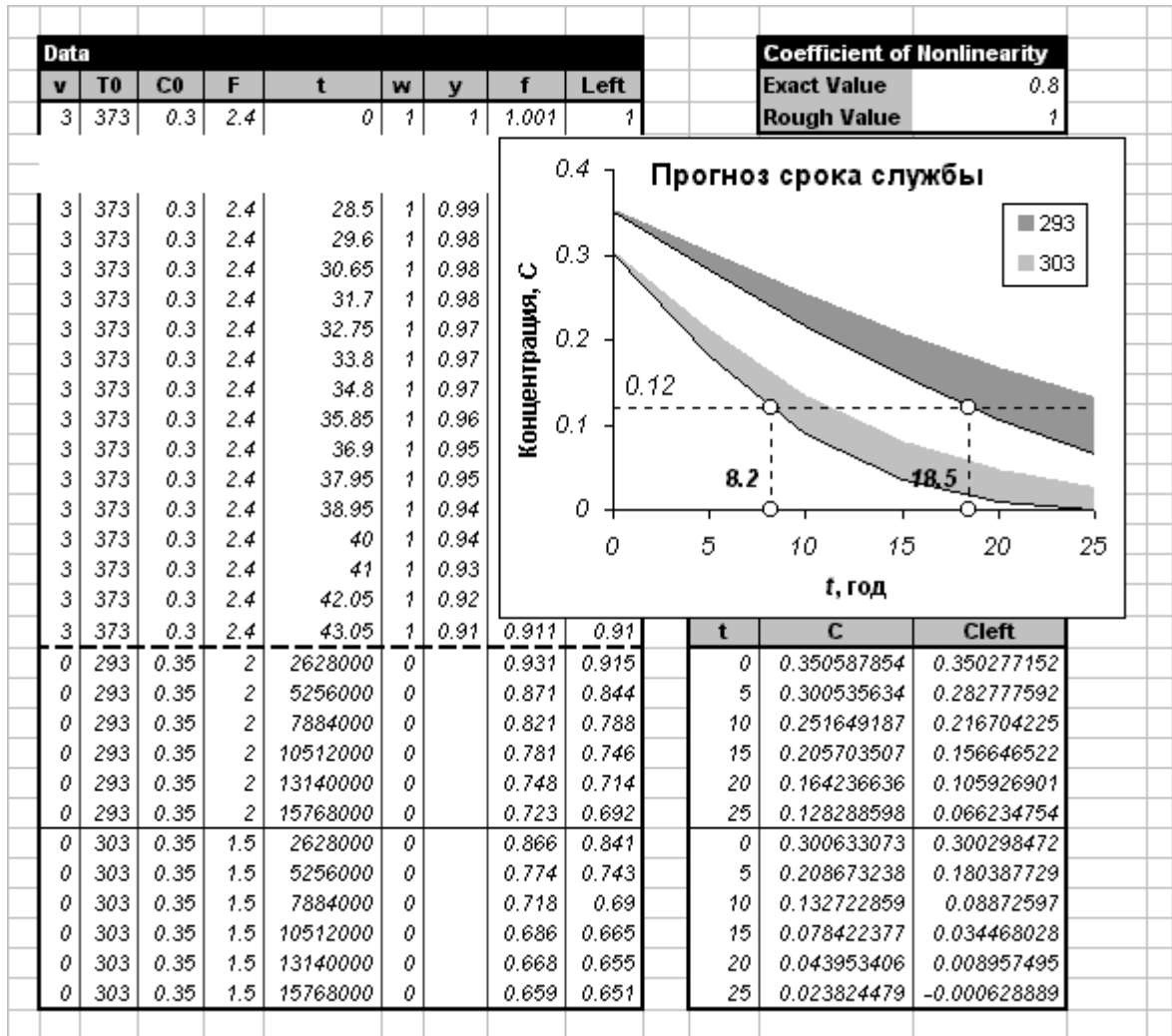


Рис. 6.3 Прогноз срока службы по данным ТГА эксперимента

Прогноз показан на графике, где представлены две кривые падения концентрации пластификатора и их нижние доверительные интервалы. Очевидно, что с достоверностью 0.95, первый образец может использоваться приблизительно 18 лет, а второй приблизительно 8 лет.

Эти результаты хорошо согласуются с известными данными [100] о сроках эксплуатации подобной продукции. Для проверки точности прогноза, доверительный интервал был рассчитан и методом связанного моделирования (А-метод в разделе 3.2). Результаты оказались так близки друг к другу, что их не возможно различить на графике. Это доказывает, что при прогнозировании на вышеприведенные условия данная модель может трактовать-

ся как линейная. Расчет коэффициента нелинейности (см. раздел 3.4), показанный в верхнем правом углу Рис. 6.3, подтверждает этот вывод.

По результатам этого примера мы можем прийти к следующему очевидному заключению. Не смотря на “ужасный вид”, приведенная здесь модель не намного сложнее, чем линейная. Более того, она даже проще. Чтобы убедиться в этом, вновь проследим путь анализа. Выбор модели прост и очевиден. Уравнения (6.1)-(6.6) совершенно ясны для любого студента-химика. Конечно, необходимо учесть, что диффузионной стадией процесса можно пренебречь, о чем, впрочем, хорошо известно. Оценивание начинается со стандартной исходной точки $y_0=1$, $k_0=1$ и $E=10000$ (это энергия активации, так что было бы неразумно задать ее меньше чем 5000). Вычисления занимают всего 5 секунд на обычном персональном компьютере. Конечно, эти данные могут быть подогнаны и некоторой линейной моделью, например $\ln(y-1)=\ln(k)+nt$, но это уравнение невозможно использовать для экстраполяции на постоянную температуру. Очень трудно создать линейную модель, зависящую от всех переменных, участвующих в задаче. Во всяком случае, мы не смогли этого сделать, даже зная правильный вид модели.

6.2. Анализ структуры сетки в радиационно модифицированном ПЭ методом ТМА

Настоящий пример иллюстрирует применение нелинейной регрессии при решении задачи оптимизации технологического процесса [32, 102]. Объект исследования – полиэтиленовое волокно, предназначенное для термоусаживающихся изделий. Как известно [104], эффект памяти, обуславливающий способность полиэтилена к термоусадке, достигается за счет формирования в нем сетки поперечных связей. Эту сетку получают, облучая материал небольшими дозами радиации. Поскольку состав и структура полиэтилена разных марок могут сильно различаться, для каждой марки нужно подбирать свою оптимальную дозу облучения.

Обычно степень сшивания полиэтилена определяют по содержанию геля. Однако этот метод не дает надежных результатов в термоусаживающемся полиэтилене, где концентрация поперечных сшивок, как правило, невысока. При низких степенях сшивания, когда содержание геля не превышает 20 %, оно измеряется с большой ошибкой. Кроме того, содержание геля в этих условиях зависит не только от густоты сетки, но и от структуры самого ПЭ, в первую очередь от молекулярно-массового распределения и степени кристалличности. Для разных промышленных марок полиэтилена эти характеристики варьируются в широких пределах, поэтому одинаковое содержание геля может отвечать разным степе-

ням сшивания в разных материалах. Таким образом, можно констатировать, что обычные методы измерения недостаточно чувствительны для низкой степени сшивания, которая применяется в усаживающихся полиэтиленах. Поэтому, мы использовали термомеханический анализ (ТМА) для контроля плотности сетки в ходе облучения.

Промышленное ориентированное волокно HDPE диаметром 0.35 мм облучали пучком электронов на воздухе, дозы облучения составляли от 5 до 20 Мрад. Образец облученного волокна выдерживался 5 мин при температуре 150С для устранения ориентации. При этом длина волокон уменьшалась в 8.6 раза, а диаметр увеличивался до 1.05 мм.

ТМА измерения выполнялись на стандартном оборудовании. Образец нагревался с постоянной скоростью нагрева (v) до 180С под постоянной нагрузкой (P). Скорость нагрева варьировалась в разных опытах от 5 до 15 град/мин, а нагрузка от 2 до 30 г. Нагрузки подбирались таким образом, чтобы максимальное относительное удлинение в эксперименте не превышало 50%. Затем, при постоянной температуре измерялась деформация $Y=L/L_0$ как функция времени. Эта деформация растет и стремится к некоторому постоянному пределу, величина которого зависит от плотности сетки.

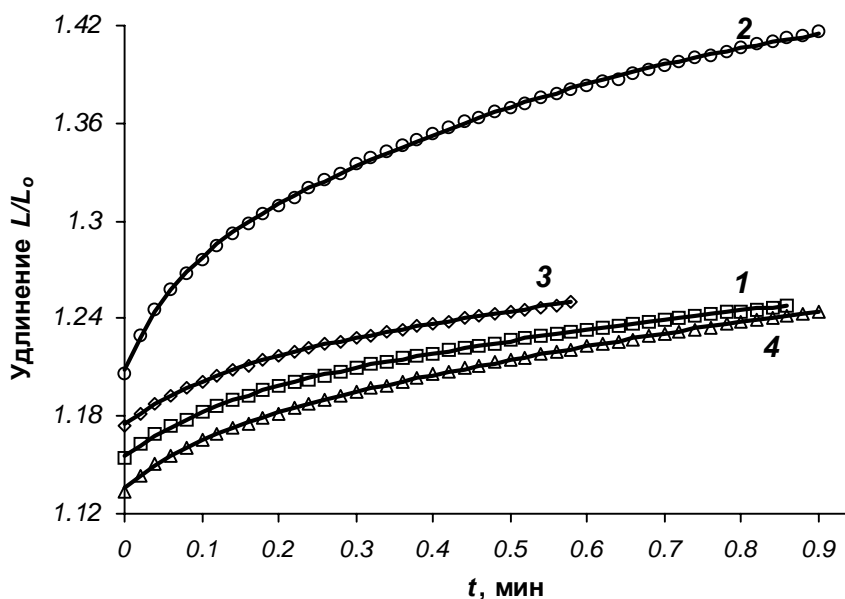


Рис. 6.4 ТМА данные и кривые для полиэтиленового волокна, облученного дозой в 10 Мрад. Экспериментальные условия представлены в Табл. 6.1

На Рис. 6.4 показаны экспериментальные данные, полученные при условиях, представленные в Табл. 6.1

Очевидно, что оценка плотности сетки требует экстраполяции экспериментальных данных далеко за область наблюдений. В такой ситуации применение формальной модели может быть причиной значительных ошибок. Поэтому, необходимо использовать содержатель-

ные модели, основанные на хорошо известном уравнении деформации идеальной сетки [105]

$$\frac{P}{S} = G \left(Y - \frac{1}{Y^2} \right) \quad (6.8)$$

Здесь P – это нагрузка; S – это площадь поперечного сечения недеформированного волокна, $Y=L/L_0$ – это отношение текущей длины образца к исходной длине и G – это модуль эластичности, который пропорционален общей концентрации узлов сетки.

Далее предполагалось [32], что сетка состоит из двух видов узлов [106]: химических узлов, то есть сшивок, которые были образованы в течение радиационной вулканизации и физических узлов, вызванных молекулярным взаимодействием. При этом модуль G представляется как сумма двух компонент: «химического» модуля – G^c и «физического» модуля – G^p .

$$G = G^p + G^c \quad (6.9)$$

Химический модуль G^c определяется концентрацией поперечных сшивок, он не меняется в ходе измерения и не должен зависеть от условий эксперимента (нагрузки и скорости предварительного нагрева). Физический модуль G^p со временем падает под действием нагрузки в ходе эксперимента.

Анализ экспериментальных кривых показывает, что кинетическая модель для разрушения физических узлов может быть представлена как сумма медленного и быстрого экспоненциальных слагаемых:

$$G^p = G_{10}^p \exp(-k_1 t) + G_{20}^p \exp(-k_2 t), \quad k_1 \ll k_2 \quad (6.10)$$

В этом уравнении параметр G_{10}^p пропорционален концентрации сильных узлов с маленькой константой деструкции k_1 , тогда как параметр G_{20}^p пропорционален концентрации слабых физических узлов, которые быстро деградируют с высокой кинетической константой k_2 .

Уравнение (6.10) – это типичная формальная модель, где число экспоненциальных слагаемых ограничено только вычислительными возможностями. Таким образом, используя содержательное уравнение (6.8) и формальное уравнение (6.10), получаем окончательную модель, которая является неявной функцией:

$$0 = \frac{S}{P} \left[G^c + G_{10}^p \exp(-k_1 t) + G_{20}^p \exp(-k_2 t) \right] \left(Y - \frac{I}{Y^2} \right) - 1 \quad (6.11)$$

В этом уравнении Y – это отклик, а t , P и S – это независимые переменные. Это нелинейная модель с 5 неизвестными параметрами. Параметр G^c – это искомая характеристика плотности сетки. Другие параметры (G_{10}^p , k_1 , G_{20}^p , k_2) изменяются в зависимости от условий эксперимента и они нас не интересуют. Модель ТМА измерений показана на Рис. 6.5, в том виде, как она представляется на рабочем листе.

```
' TMA model
0=S/P*G*[Y-Y^(-2)]-1; 1<Y<1.5
G=Gp+Gc
Gp=G1*exp(-k1*t)+G2*exp(-k2*t)
S=0.8659
Gc=?
G1=?
G2=?
k1=?
k2=?
```

Рис. 6.5 ТМА модель, представленная на рабочем листе

Рассмотрим устройство этой модели, которая представляет неявную функцию. Первая строка является комментарием и поэтому она отмечена апострофом. Вторая строка соответствует уравнению (6.8), записанному в стандартной форме, принятой в Fitter для неявных функций (см. раздел 5.3). Важно, что граничное условие $1 < Y < 1.5$ записано в той же строке и отделено от основного уравнения точкой с запятой. Величины граничных условий легко определяются из экспериментальных данных, показанных на Рис. 6.4. В третьей и четвертой строках записаны уравнения для модуля (6.9) и (6.10). Затем определяется константа S равная площади поперечного сечения волокна, и перечисляются все неизвестные параметры, каждый в отдельной строке.

При решении проблемы экстраполяции за область наблюдений, проблема адекватности модели имеет первостепенное значение. В данном случае истинность модели подтверждается независимыми оценками значения G^c полученными в разных условиях измерений, т.е. нагрузки и нагрева. Результаты, представленные в Табл. 6.1 показывают, что искомый параметр G^c действительно не зависит от условий измерения.

В первых столбцах этой таблицы указаны условия экспериментов, которые представлены на Рис. 6.4. Затем приведены соответствующие оценки химического модуля вместе со среднеквадратичными отклонениями, а в последнем столбце – оценки дисперсии ошибки измерения. В последних трех строках таблицы приведены значения оценок, полученные

при совместной обработке четырех серий экспериментальных данных различными способами. Остановимся на этом для более подробного рассмотрения.

Табл. 6.1 Оценка химических модулей для различных экспериментов

| | P , г | v , град/мин | $G^c \pm \text{СКО}(G^c)$, г/мм ² | s^2 |
|--------------|---------|----------------|---|----------------------|
| 1 | 5 | 10 | 8.708 ± 0.046 | $6.5 \cdot 10^{-08}$ |
| 2 | 7 | 10 | 8.312 ± 0.038 | $1.2 \cdot 10^{-06}$ |
| 3 | 5 | 5 | 8.707 ± 0.026 | $7.5 \cdot 10^{-09}$ |
| 4 | 5 | 15 | 8.587 ± 0.066 | $1.9 \cdot 10^{-07}$ |
| Joint | | | 8.326 ± 0.022 | $4.3 \cdot 10^{-07}$ |
| 1234 | | | 8.334 ± 0.021 | $4.3 \cdot 10^{-07}$ |
| 4321 | | | 8.335 ± 0.021 | $4.3 \cdot 10^{-07}$ |

В строке озаглавленной **Joint** приведены результаты, полученные при совместной обработке всех четырех кривых обычным методом наименьших квадратов (МНК). В строках **1234** и **4321** находятся результаты, полученные методом последовательного байесовского оценивания (ПБО). Этот метод был объяснен в главе 2, в общем. Здесь же мы покажем, как алгоритм ПБО (2.38) применяется для решения конкретной задачи. Подробности всех расчетов можно найти в [126], где приведен файл *TMA.XLS*, содержащий исходные данные и порядок расчетов.

Сначала методом МНК обрабатывается первая серия данных и выводится таблица **Fitting results** (см. Рис. 5.17). По итогам обработки создается апостериорная информация с помощью диалога *дополнительных действий* (см. раздел 5.11). Получившаяся таблица апостериорной информации показана на Рис. 6.6 в левом верхнем углу.

В этой таблице стираются все имена параметров G_1 , G_2 , k_1 и k_2 , которые являются частными и не переносятся (в отличие от G_c) в следующую серию. Далее регистрируется вся необходимая информация (данные, модель, параметры) для второй серии данных. Наконец, проводится регистрация априорной информации (см. раздел 5.10). В поле **Range** нужно ввести ссылку на всю область, занимаемую таблицей апостериорной информации. В поле **Freedom degrees** вводится ссылка на клетку, где по итогам обработки первой серии было выведено число степеней свободы (**Degrees of freedom**), а в поле **Variance** ссылку на клетку, в которой было выведена взвешенная дисперсия ошибки (**Error variance**).

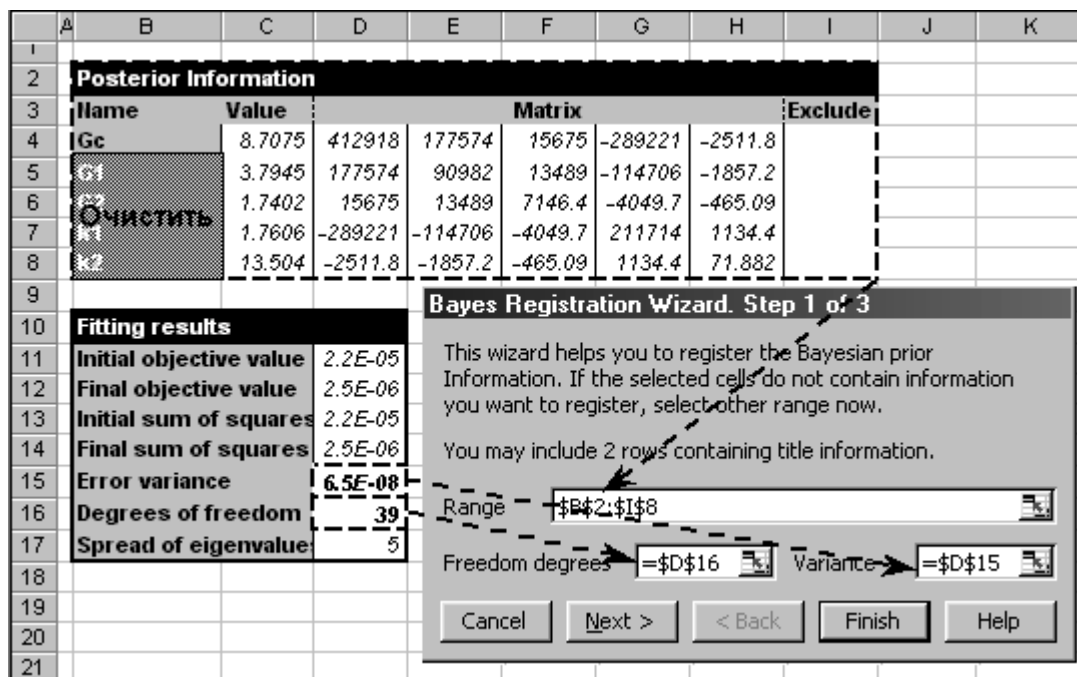


Рис. 6.6 Создание априорной информации в примере ТМА

Схема регистрации показана на Рис. 6.6. После завершения регистрации априорной информации на листе появляется новая таблица – априорной информации (см. Рис. 6.7), перестроенная в соответствии с правилами, изложенными в разделе 2.3.

| | A | B | C | D | E | F | G | H | I |
|---|---|-----------------------------|--------------|---------------|---|---|---|----------------|---|
| 1 | | | | | | | | | |
| 2 | | A Priori Information | | | | | | | |
| 3 | | Name | Value | Matrix | | | | Exclude | |
| 4 | | Gc | 8.7075 | 465.52 | 0 | 0 | 0 | 0 | |
| 5 | | | 0 | 0 | 0 | 0 | 0 | 0 | |
| 6 | | | 0 | 0 | 0 | 0 | 0 | 0 | |
| 7 | | | 0 | 0 | 0 | 0 | 0 | 0 | |
| 8 | | | 0 | 0 | 0 | 0 | 0 | 0 | |
| 9 | | | | | | | | | |

Рис. 6.7 Априорная информация в примере ТМА

Вторая серия обрабатывается уже с учетом этой априорной информации. Затем процедура повторяется – создается апостериорная информация, которая затем регистрируется как априорная для третьей серии и т.д. до последней, четвертой серии.

В рассматриваемом примере имеются несколько особенностей, которые следует выделить особо. Во-первых, есть веские основания считать, что дисперсия ошибки постоянна во всех сериях (см. Табл. 6.1), поэтому мы используем здесь априорную информацию первого типа (см. раздел 1.1). Во-вторых, порядок обработки серий может быть и другим – не **1234**, а, например, обратный **4321**. В таблице Табл. 6.1 приведены результаты именно этих процедур ПБО. Те же значения показаны и на Рис. 6.8, где каждый способ обработки

представлен точкой с разбросом, соответствующим среднеквадратичному отклонению, приведенному в четвертом столбце Табл. 6.1.

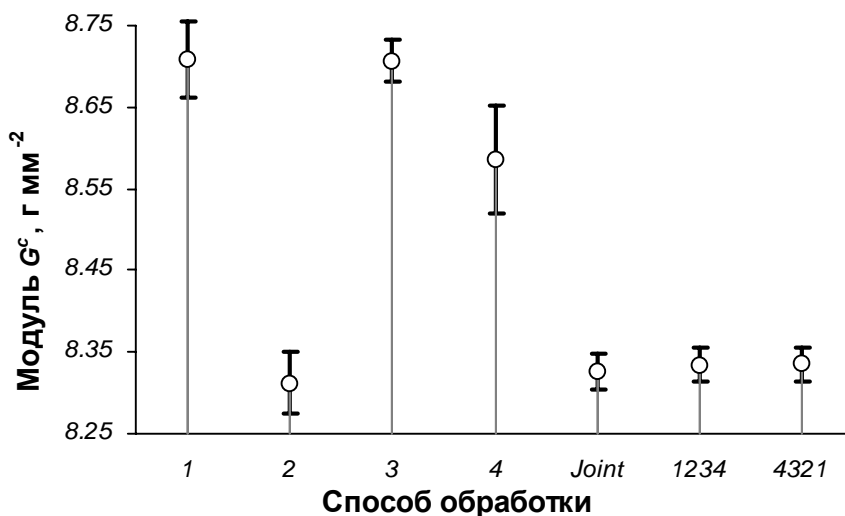


Рис. 6.8 Оценки химического модуля, полученные различными способами

Из этого рисунка очевидно, что совместные оценки химического модуля G^c полученные различными способами (*Joint*, *1234* и *4321*) практически совпадают. Этот результат подтверждает теорему о свойствах оценок метода ПБО, доказанную в разделе 2.3.

Зависимость параметра G^c от дозы облучения для трех различных сортов полиэтилена показаны на Рис. 6.9.

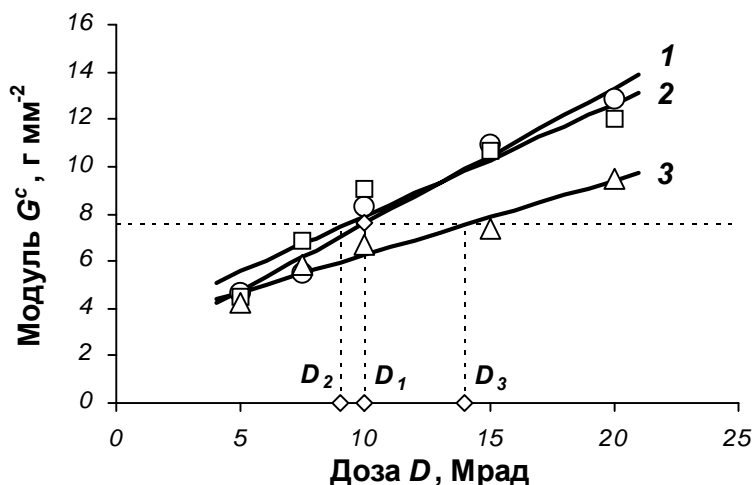


Рис. 6.9 Зависимость химического модуля G^c от дозы облучения для трех различных сортов полиэтиленовых волокон

Видно, что при одинаковой дозе волокно **3** имеет значительно менее густую сетку, чем **1** и **2**, т.е. для формирования сетки нужной структуры ему потребуются более высокие дозы. Оптимальная доза облучения была получена посредством интерполяции этих данных. Для этой цели мы использовали простую линейную модель

$$G^c = G_0^c + qD$$

Волокно 1 использовалось как стандартное калибровочное. Например, если для волокна 1 стандартная доза 10 Мрад, то для получения такой же плотности сетки волокно 2 должно быть облучено с дозой 9 Мрад, а волокно 3 – дозой 14 Мрад.

Этот пример показывает эффективность практического сочетания содержательной и формальной моделей. Часто такая комбинация помогает уменьшить число экспериментов и получить очень полезную информацию. При этом модель может иметь очень сложную математическую форму, но ясную физическую основу. Важно только использовать надежный инструмент для оценки параметров модели.

6.3. Оценка активности антиоксидантов методом ДСК

Исследование эффективности антиоксидантов (АО) в полиолефинах – это длительный и дорогой процесс. Стандартная процедура тестирования АО предполагает выдержку образцов в печах в течение 1-3 месяцев. В этом разделе излагается метод, который позволяет быстро предсказывать активность АО по данным, полученным методом дифференциальной сканирующей калориметрии (ДСК) [32, 51, 52, 128]. Рассматриваемый пример интересен тем, что здесь нужно три раза использовать методы нелинейного регрессионного анализа до получения желаемого результата. При этом решения, найденные на предыдущем шаге используются на следующем шаге как исходные данные.

Антиоксиданты – это специальные добавки, которые замедляют термоокислительное старение полимеров. Они защищают материал от окисления в ходе изготовления материала и при эксплуатации изделий. Взаимодействуя со свободными радикалами, АО обрывает цепи, но при этом и сам расходуется. Антиоксидант полностью подавляет окисление до тех пор, пока его концентрация не упадет ниже некоторого критического значения. Поэтому, основной характеристикой эффективности АО является *период индукции* – время, в течение которого концентрация АО достаточна велика. Чем больше период индукции (при той же температуре) – тем более эффективен АО. По сложившейся практике период индукции определяется с помощью длительной выдержки образцов в печах при температуре 120-140С. Время выдержки зависит от качества АО и может составлять от 1 дня для плохого АО до 100 дней для хорошего АО. В течение этого времени экспериментатор обследует образцы в поисках признаков деградации – растрескивания, пожелтения и т. п. Разумеется, такой метод является ненадежным и требует больших затрат времени и средств.

Альтернативой этому методу является подход, использующий ДСК измерения, с последующей обработкой полученных данных методами нелинейного регрессионного анализа. Дифференциальная сканирующая калориметрия – это метод, в котором сигналом является величина теплового потока от образца, нагреваемого с постоянной скоростью. Если в образце происходит химическая реакция с не нулевым тепловым эффектом, то величина ДСК сигнала пропорциональна скорости реакции. Применительно к рассматриваемой проблеме, это означает, что пока концентрация АО в образце еще достаточна для подавления цепного окисления, то ДСК сигнал является константой, однако, начиная с некоторой температуры, он начинает резко расти. Эта температура называется *температурой начала окисления* (ТНО). Проводя эксперименты при разной скорости нагрева v , можно получить несколько значений ТНО.

На Рис. 6.10 представлены экспериментальные данные (точки), полученные методом дифференциальной сканирующей калориметрии с разной скоростью нагрева на образцах полипропилена. Там же показаны и температуры начала окисления для каждой кривой. Главная проблема ДСК – это количественное описание данных. Нелинейная регрессия помогает решить эту задачу и определить температуры начала окисления.

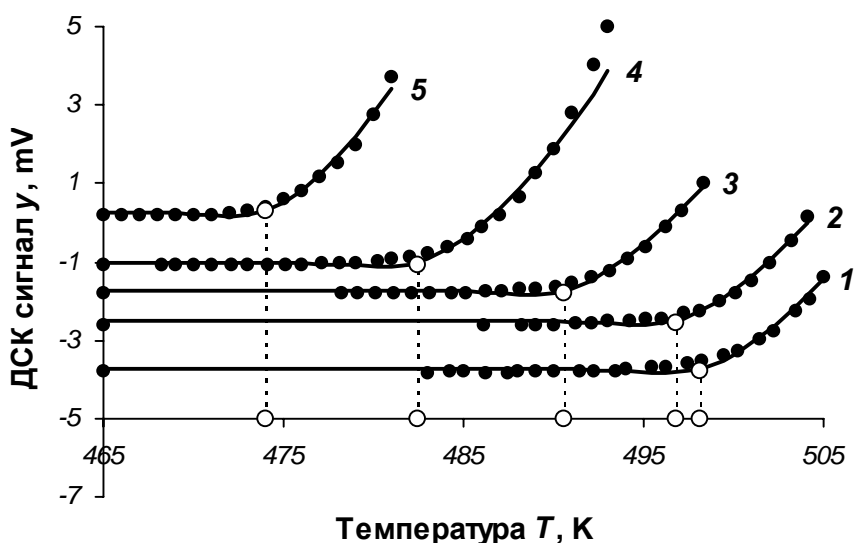


Рис. 6.10 Данные ДСК и регрессионные кривые для различных скоростей нагрева v (град/мин): 1 – 20; 2 – 15; 3 – 10; 4 – 5, 5 – 2 в полипропилене.

Как отмечалось выше, ДСК сигнал y пропорционален скорости реакции

$$y = mD \frac{dc}{dt} \quad (6.12)$$

где m – масса образца, D – константа. На начальной стадии, когда идет автоокисление, [107], реакция описывается уравнением ($c=[\text{ROOH}]$)

$$\frac{dc}{dt} = k \exp\left(-\frac{E}{R(T_0 + vt)}\right) \sqrt{c}; \quad c(T_v) = 0 \quad (6.13)$$

где E – это энергия активации, которую можно считать [32] общей для всех типов АО, k – предэкспонента, T_0 – температура начала разогрева в ДСК, R – универсальная газовая константа. Объединяя уравнения (6.12) и (6.13), получаем

$$y = f_v + \begin{cases} 0, & T < T_v \\ mc e^{-\frac{E}{RT}}, & T > T_v \end{cases} \quad (6.14)$$

$$\frac{dc}{dT} = \frac{k}{v} e^{-\frac{E}{RT}}, \quad c(T_v) = 0$$

Эта модель включает три независимые переменные (предиктора), а именно: v (скорость нагрева), m (масса образца), T (температура) и четыре неизвестных параметра, а именно: k (предэкспонента), E (энергия активации), T_v – значение ТНО и фоновый параметр: f_v . Последние два параметра зависят от скорости нагрева v .

Текстовое поле с моделью, предназначенной для описания данных, представленных на Рис. 6.10, показано на Рис. 6.11.

```
'Модель для оценки ТНО
y=f0+m*exp(-E/T/R)*c*hev(T-T0)
D[c]/D[T]=k/V*exp(-E/T/R); c(T0)=0
R=1.98717
k=?
E=?
T0=?
f0=?
```

Рис. 6.11 Модель для оценки ТНО в системе Fitter

Таким образом, можно определить ТНО для образцов, содержащих разное начальное количество АО для различных скоростей нагрева (см. Рис. 6.12). Эти данные содержат в себе информацию, характеризующую эффективность АО. Остается только применить нелинейный регрессионный анализ для того, чтобы извлечь эту информацию. Подробности всех расчетов можно найти в [126], где приведен файл *DSC.XLS*, содержащий исходные данные и порядок расчетов.

В ходе старения материала происходит расход антиоксиданта. Период индукции определяется моментом времени, когда АО израсходуется до некоторого критического значения A_c , которое зависит от температуры по закону Аррениуса

$$A_c = k_c \exp\left(-\frac{E_c}{RT}\right) \quad (6.15)$$

Расход АО в ходе окисления идет термически

$$\frac{dA}{dt} = -kA \quad (6.16)$$

$$A(0) = A_0$$

где A – текущая концентрация АО, A_0 – начальная концентрация АО, k – константа скорости, зависящая от температуры по закону Аррениуса

$$k = k_a \exp\left(-\frac{E_a}{RT}\right) \quad (6.17)$$

где R – газовая постоянная.

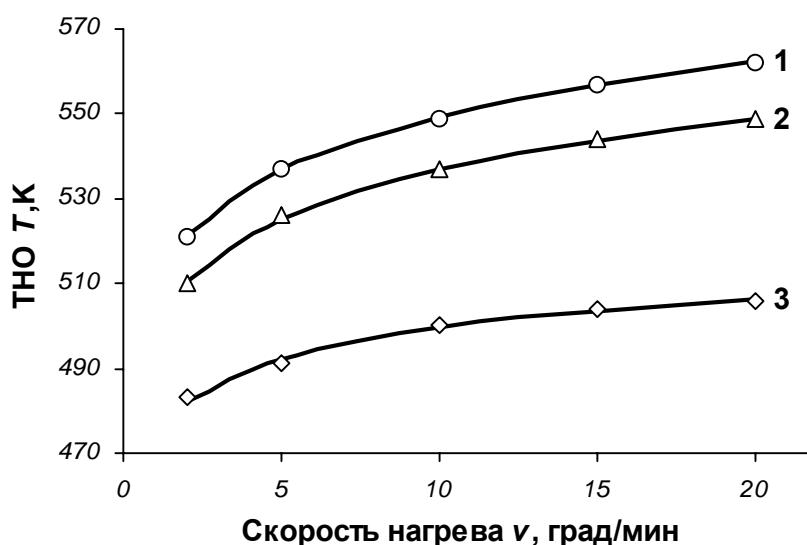


Рис. 6.12 Зависимость ТНО от скорости нагрева при разных начальных концентрациях АО: 0.1 (1, O), 0.05 (2, □), 0.025 (3, △). Полиэтилен + Naugard 76

В ДСК эксперименте образцы разогреваются со скоростью v от начальной температуры $T_0=293K$ до температуры начала окисления

$$T(t) = T_0 + vt \quad (6.18)$$

Расход АО в таком эксперименте происходит по уравнению (6.16), где константа скорости k зависит от времени

$$k = k_a \exp\left(-\frac{E_a}{R(T_0 + vt)}\right) \quad (6.19)$$

В этом случае решение уравнения (6.16) можно представить в виде

$$A(t) = A_0 \exp(-k_0 Z(t)) \quad (6.20)$$

где

$$Z(t) = \int_0^t \exp\left(-\frac{E_a}{R(T_0 + vs)}\right) ds \quad (6.21)$$

Этот интеграл может быть выражен через интегральную показательную функцию $E_n(x)$ (6.7) в виде, зависящем не от времени, а от температуры (6.18) –

$$Z(T) = \frac{I}{v} \left[TE_2 \left(\frac{E_a}{RT} \right) - T_0 E_2 \left(\frac{E_a}{RT_0} \right) \right] \quad (6.22)$$

Подставим (6.22) в (6.20) и приравняем текущую концентрацию АО критической (6.15)

$$A_0 \exp(-k_0 Z(T)) = k_c \exp\left(-\frac{E_c}{RT}\right) \quad (6.23)$$

Полученное уравнение неявно определяет зависимость температуры начала окисления T в методе ДСК от неизвестных параметров и известных условий эксперимента (предикторов). Однако практически использовать это уравнение очень неудобно. Приведем его в форму, которая более подходит для применения этой модели в системе Fitter.

Прологарифмируем уравнение (6.23)

$$k_0 \left[TE_2 \left(\frac{E_a}{RT} \right) - T_0 E_2 \left(\frac{E_a}{RT_0} \right) \right] + v \left[\ln(k_c) - \ln(A_0) - \frac{E_c}{RT} \right] = 0 \quad (6.24)$$

Будем рассматривать это уравнение как неявную функцию $T=T(v)$, выраженную в форме $F(T, v)=0$. Найдем производную $\frac{dT}{dv}$. Для этого воспользуемся следующим известным соотношением [144]

$$\frac{dT}{dv} = -\frac{\partial F}{\partial v} \left[\frac{\partial F}{\partial T} \right]^{-1} \quad (6.25)$$

и свойствами интегральной экспоненты [143]

$$E_2(x) = e^{-x} - xE_1(x); \quad \frac{dE_1(x)}{dx} = -\frac{e^{-x}}{x} \quad (6.26)$$

В результате получим

$$\frac{dT}{dv} = \frac{\frac{E_c}{RT} + c - \ln(A_0)}{\exp\left(a - \frac{E_a}{RT}\right) + v \frac{E_c}{RT^2}} \quad (6.27)$$

где $c=\ln(k_c)$ и $a=\ln(k_a)$. Это уравнение нужно дополнить начальным условием

$$T(0) = T_0 \quad (6.28)$$

В уравнении (6.27) температура T рассматривается как отклик, а скорость v и начальная концентрация A_0 – как предикторы. Величины a , E_a , c и E_c – это неизвестные параметры,

которые оцениваются по экспериментальным данным. В такой форме модель уже пригодна для применения в системе Fitter. На Рис. 6.13 приведено соответствующее текстовое поле с этой моделью.

```
'Модель для оценки параметров ТНО
D[T]/D[v]=(Ec/R/T-b)/[exp(a-Ea/R/T)+v*Ec/R/T/T];T(0)=T0
b=c-log(A0)
T0=293
R=1.98717
a=?
Ea=?
c=?
Ec=?
```

Рис. 6.13 Модель для оценки параметров ТНО в системе Fitter

Таким образом, мы видим, что оценки ТНО, найденные из кривых ДСК на первом этапе, используются на втором этапе как исходные «экспериментальные» данные, из которых можно оценить параметры a , E_a , c и E_c , характеризующие процесс старения материала. На Рис. 6.12 показано, как модель (6.27) (кривые) описывает данные – ТНО (точки), полученные из кривых ДСК.

Теперь, опираясь на полученные оценки, можно построить прогноз периода индукции. Для этого нужно проинтегрировать уравнение (6.16) при постоянных условиях

$$A(t) = A_0 \exp(-kt) \tag{6.29}$$

Чтобы определить период индукции t_i надо концентрацию $A(t)$ из (6.29) приравнять A_c из уравнения (6.15)

$$A_0 \exp(-kt) = A_c \tag{6.30}$$

или

$$A_0 \exp\left[-\exp\left(a - \frac{E_a}{RT_e}\right)t\right] = \exp\left(c - \frac{E_c}{RT_e}\right) \tag{6.31}$$

где T_e – это температура эксплуатации. Отсюда получаем

$$t_i = \left[\frac{E_c}{RT_e} + \ln(A_0) - c \right] \exp\left(\frac{E_a}{RT_e} - a\right) \tag{6.32}$$

В этом уравнении используются оценки параметров a , E_a , c и E_c , найденные на втором шаге. Для того чтобы получить не только среднее значение прогнозируемого периода индукции, но и его доверительные границы необходимо учесть неопределенность в оценках параметров, участвующих в уравнении (6.32).

Проблема состоит в том, что при прогнозировании нужно использовать модель (6.32), для которой нет экспериментальных данных и, соответственно, нет и суммы квадратов откло-

нений. Однако, используя байесовский подход, можно обойти эту трудность. Действительно, в разделе 2.1 показано, что в общем случае целевая функция $Q(\mathbf{a})$ состоит из двух частей

$$Q(\mathbf{a}) = S(\mathbf{a}) + B(\mathbf{a}) \quad (6.33)$$

где $S(\mathbf{a})$ – это сумма квадратов отклонений, а $B(\mathbf{a})$ – это байесовский член. При обработке данных, являющихся значениями ТНО, с помощью модели (6.27), априорная информация не используется, поэтому в этой задаче целевая функция состоит только из суммы квадратов отклонений

$$Q(\mathbf{a}) = S(\mathbf{a}) \quad (6.34)$$

После того, как оценки параметров найдены, можно построить апостериорную информацию (см. раздел 2.2), состоящую из \mathbf{b} , \mathbf{H} , s_0^2 и N_0 . Из теоремы, доказанной в разделе 2.3 следует, что исходную экспериментальную информацию можно заменить на байесовскую апостериорную информацию, т.е.

$$Q(\mathbf{a}) = S(\mathbf{a}) \approx B(\mathbf{a}) \quad (6.35)$$

При этом байесовский член $B(\mathbf{a})$ имеет вид

$$B(\mathbf{a}) = s_0^2 \left[N_0 + (\mathbf{a} - \mathbf{b})' \mathbf{H} (\mathbf{a} - \mathbf{b}) \right] \quad (6.36)$$

Этот подход используется при прогнозировании периода индукции.

На Рис. 6.14 показано как это делается практически. В верхнем левом углу представлена таблица данных. В первых двух столбцах этой таблицы указаны значения концентраций **АО** и температуры **T**, на которые выполняется прогноз. Важно, что в столбце экспериментальных значений отклика, озаглавленном **ti** нет ни одного значения и соответствующие веса **w** все равны нулю. Затем идут столбцы левой границы доверительного интервала (**Left**), прогноз (**Fit**) и правая граница (**Right**). На рисунке также показаны некоторые результаты, полученные при обработке данных ТНО, – таблица оценок параметров (**Parameters**), результаты поиска (**Fitting Results**) и таблица апостериорной информации (**Posterior Information**). Видно, как эти результаты используются при регистрации априорной информации (см. раздел 5.10).

После регистрации обработка проводится без оптимизации, т.е. при нулевом значении максимального числа итераций (см. раздел 5.6). Кроме того, следует использовать метод линеаризации (см. раздел 3.1) при построении доверительных интервалов (см. раздел 5.7).

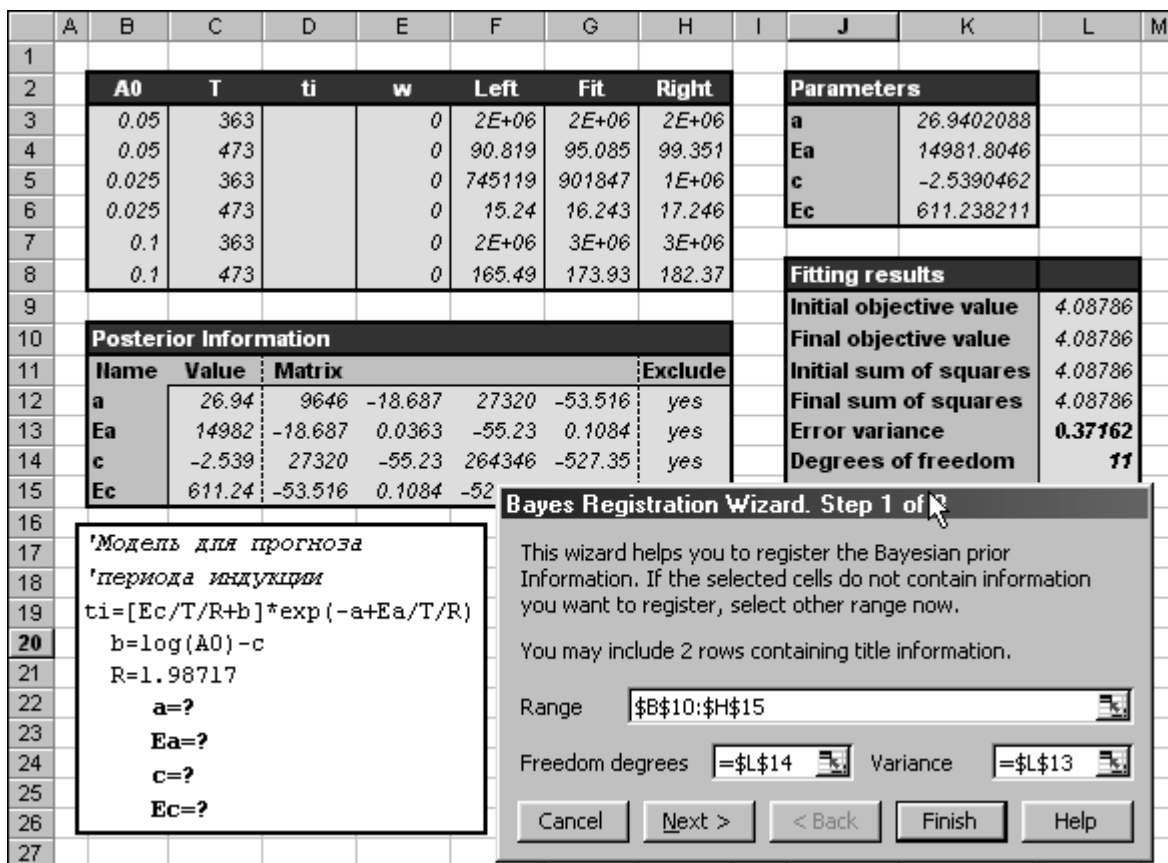


Рис. 6.14 Использование априорной информации для прогнозирования в примере ДСК

Результаты показаны в таблице данных в последних трех столбцах.

С помощью такой техники байесовского прогнозирования были получены результаты, представленные на Рис. 6.15 и Рис. 6.16.

На первом рисунке показаны результаты предсказания периода индукции при $T=200C$ для полипропиленовых образцов с различными антиоксидантами при концентрации 0.05 . Доверительные интервалы с достоверностью 0.85 представлены серыми колонками, средние значения – черными точками (●), а тестовые значения, померенные традиционным способом – белыми точками (○).

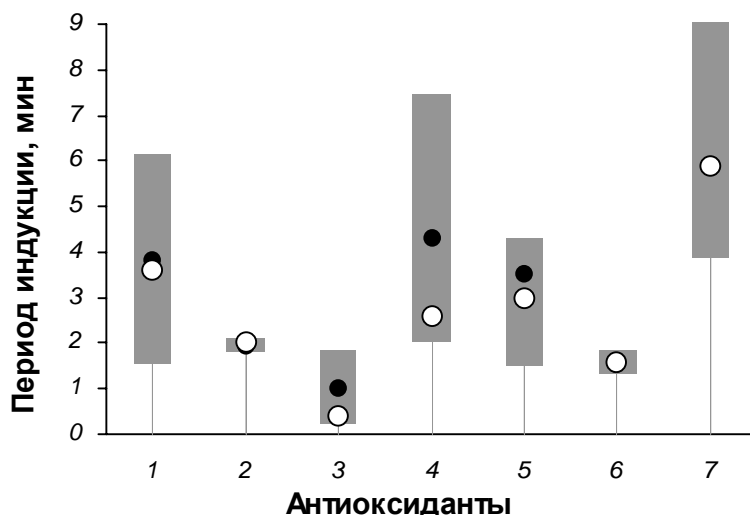


Рис. 6.15 Прогноз периода индукции полипропилена с различными антиоксидантами при $T=473\text{K}$. Доверительные интервалы с достоверностью 0.85 показаны затененными областями, средние значения черными точками (●), а контрольные значения – белыми точками (○). Концентрация АО равна 0.05 .

На Рис. 6.16 показан прогноз периода индукции полиэтилена с АО *Naugard 76*, построенный по данным, представленным на Рис. 6.12. Здесь также показаны доверительные интервалы (серая область) и одно тестовое значение – белая точка (○).

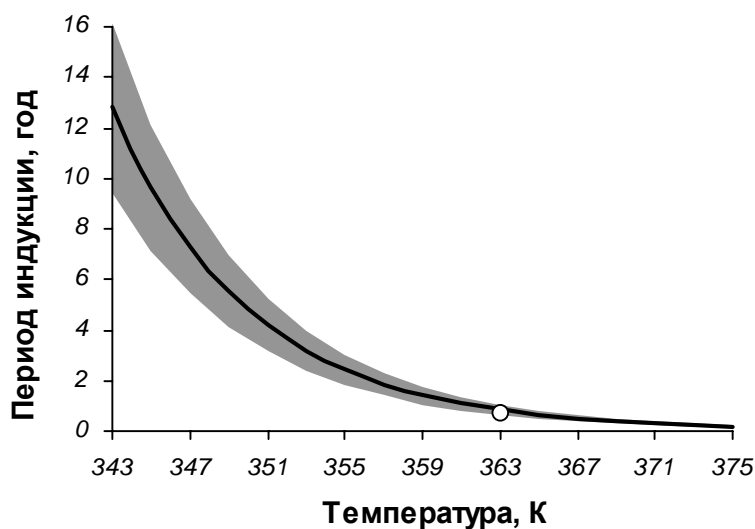


Рис. 6.16 Прогноз периода индукции полиэтилена с *Naugard 76* для различных температур. Доверительные интервалы с достоверностью 0.85 показаны затененными областями, а среднее значение черной кривой. Показано также и контрольное значение – белая точка (○). Концентрация АО равна 0.025 .

Из этих результатов видно, что предсказанные величины хорошо согласуются с референтными данными, хотя точность прогноза не всегда хороша, особенно для полипропиленовых образцов.

Разобранный пример показывает, что, применяя методы нелинейного регрессионного анализа к данным ДСК можно получить кинетические параметры, характеризующие эффективность АО (константу скорости расщепления и критическую концентрацию с соответствующими энергиями активации), а также прогнозировать период индукции при разных температурах.

6.4. Результаты главы 6

Примеры обработки термограмм полимеров, представленные в этой главе, показывают, что НЛРА позволяет извлекать и использовать важную информацию, содержащуюся в подобных экспериментальных данных. Не смотря на сложную форму используемых моделей, обработка данных в системе Fitter проходит достаточно просто.

Пример ТГА анализа, разобранный в первом разделе, позволяет прийти к парадоксальному выводу, что сложность модели зависит не от ее внешнего вида, а от комплекса внутренних свойств данных и модели, характеризуемых коэффициентом нелинейности. Дифференциальное уравнение (6.6) имеет решение, выражаемое только через специальные функции, тем не менее, эта задача не сложнее, чем простейшая линейная регрессия. Нужно только использовать правильный подход для обработки данных.

Второй пример, посвященный обработке ТМА данных, дает нам два важных результата. Во-первых, прямой расчет подтвердил правильность результата, доказанного в разделе 2.3 об эквивалентности обычных оценок параметров, построенных методом максимума правдоподобия и оценок, полученных методом последовательного байесовского оценивания. Во-вторых, здесь продемонстрирован общий принцип нелинейного моделирования – для экстраполяции необходима содержательная, физически обоснованная модель, тогда как для интерполяции достаточно формальной, структурно обусловленной модели.

Третий пример иллюстрирует другой важный принцип, применяемый при обработке данных и прогнозировании. Это подход, при котором результаты обработки первичных данных становятся исходной информацией для следующей модели. При этом метод последовательного байесовского оценивания используется для построения доверительных интервалов в прогнозе.

7. Оценивание кинетических параметров по спектральным данным

Спектральные данные сейчас часто используются для определения кинетических параметров. Компьютеризированная спектроскопия дает нам очень быстрый метод измерения в режиме “on-line”. Получаемые спектры содержат важную информацию о кинетических процессах, которую нужно уметь извлекать и использовать. Все это заставляет разрабатывать сложные методы для решения таких задач. Имеется довольно много так называемых «формальных» подходов (“soft” methods) [111, 112, 113]. Они решают задачу в два приема. Сначала строится отдельная формальная калибровочная модель, связывающая спектры и концентрации. Затем предсказанные концентрации используются для решения обратной задачи – оценивания параметров кинетической модели. Часто подобные подходы не обеспечивают требуемой точности в получаемых результатах. С другой стороны, содержательное физико-химическое моделирование (“hard” modeling) [114, 115], которое базируется на основных кинетических принципах [95], позволяет получать оценки параметров с высокой точностью. Методы решения «обратных задач» могут быть применены, когда все спектры чистых компонент реакции известны. В противном случае, исследователь сталкивается с проблемой одновременного нелинейного оценивания большого числа неизвестных параметров. Обработка только нескольких «характерных» спектральных линий упрощает задачу, но при этом теряется значительная часть данных, что неприемлемо при большой ошибке измерения.

Предлагается использовать для решения подобных задач метод последовательного Байесовского оценивания (ПБО) (см. главу 2). В нем учитывается то, что не все неизвестные параметры равнозначны – есть кинетические параметры k , общие для всей модели, – и есть спектральные параметры P , работающие только на своей длине волны λ . Байесовский подход позволяет использовать все имеющиеся экспериментальные данные, но не сразу, а последовательно, одна длина волны за другой. Для того чтобы кинетическая информация, полученная на предыдущем шаге, не пропала, она преобразуется в форму Байесовской априорной информации и учитывается на следующем шаге. Такой подход позволяет разбить одну большую задачу оценивания на цепь маленьких, так чтобы, проигрывая «в пути», мы выигрывали «в силе».

Эта методика иллюстрируется двумя примерами [108, 109, 110] для последовательной двухстадийной реакции. Первый пример модельный, где экспериментальные данные были построены для проверки работоспособности подхода. Второй пример – реальные данные

ИК спектров, снятых в ходе эпексидизации 2,5-ди-*терт*-бутил-1,4-бензохинона. Модельные данные позволяют сравнить оценки с «истинными» значениями параметров, и, тем самым проверить метод ПБО на этом примере. Реальные данные дают возможность сравнить метод ПБО с другими известными методами оценивания кинетических констант.

7.1. Моделирование «экспериментальных» данные

Модель, описывающая кинетику изменения спектральных данных в ходе химической реакции, может быть выражена как функция времени t и длины спектральной волны x зависящая от неизвестных кинетических параметров k

$$y(t, x, \mathbf{k}) = \sum_{i=1}^l c_i(t, \mathbf{k}) p_i(x) \quad (7.1)$$

Здесь y – это спектральный сигнал, c_i – это концентрации компонент, p_i – это спектры чистых компонент и l – это число реагентов. В дискретном случае, когда спектры разделены на m длин волн, а время представлено n точками, можно использовать матричные обозначения для этого уравнения

$$Y = CP + E. \quad (7.2)$$

Здесь Y – это $(n \times m)$ матрица спектральных данных; C – это $(n \times l)$ матрица концентраций, которая зависит от неизвестных кинетических параметров и P – это $(l \times m)$ неизвестная матрица спектров чистых компонент. Кроме того, $(n \times m)$ матрица ошибок E участвует в модели (7.2). При таком формальном рассмотрении нам важна не природа измеряемых спектров (ИК, УФ, ЯМР и т.д.), а лишь то обстоятельство, что измеряемая величина сигнала $Y(t, x)$ явно зависит от времени t и «длины волны» x и, неявно, через концентрации C , от констант скоростей k .

Матрица концентрации C может быть получена как решение кинетической модели. Мы будем рассматривать в качестве примера двухстадийную реакцию первого порядка



которая может быть представлена системой обыкновенных дифференциальных уравнений

$$\begin{aligned} \frac{dA}{dt} &= -k_1 A; & A(0) &= A_0 \\ \frac{dB}{dt} &= k_1 A - k_2 B; & B(0) &= B_0 \\ \frac{dC}{dt} &= k_2 B; & C(0) &= C_0 \end{aligned} \quad (7.4)$$

Эта система имеет аналитическое решение

$$A = A_0 \exp(-k_1 t)$$

$$B = \frac{k_1 A_0}{k_1 - k_2} [\exp(-k_2 t) - \exp(-k_1 t)] + B_0 \exp(-k_2 t) \quad (7.5)$$

$$C = A_0 + B_0 + C_0 + \frac{A_0}{k_1 - k_2} [k_2 \exp(-k_1 t) - k_1 \exp(-k_2 t)] - B_0 \exp(-k_2 t)$$

Здесь мы используем одно и тоже обозначение для компонент реакции A , B , C и для их концентраций $[A]=A$, $[B]=B$, $[C]=C$. Это, конечно, не аккуратно, но значительно упрощает запись формул.

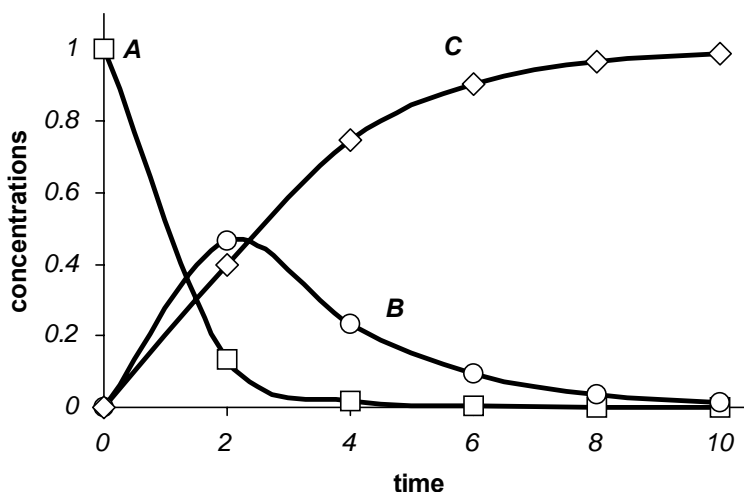


Рис. 7.1 Модельные кинетические кривые

Данные для модельного примера были вычислены с помощью уравнений (7.5) при следующих значениях исходных концентраций

$$A_0=1, \quad B_0=C_0=0.$$

«Истинные» значения констант были выбраны так –

$$k_1=1, \quad k_2=0.5$$

и, наконец, точки наблюдения, где «измерялись» спектры были взяты следующие –

$$t=0, 2, 4, 6, 8, 10.$$

Таким образом, количество наблюдений (n) равно 6. Столь малое число точек было установлено намерено, с целью создать максимальные сложности при оценивании. Шесть – это минимальное количество экспериментальных данных, при котором оценивание параметров по одной отдельно взятой кинетической кривой, соответствующей какой-то длине волны, еще возможно. Действительно, число неизвестных параметров в этой кинетике равно 5 – два кинетических параметра и три спектральных. Соответствующие кинетические кривые показаны на Рис. 7.1.

Матрица спектров чистых компонент P строилась обычным образом с использование перекрывающихся гауссовских спектральных пиков. Каждый спектр p нормировался так, что $\max(p)=1$. Спектры всех компонент реакции (7.3) показаны на Рис. 7.2. Они разбиты на 53 (m) длин волн. Содержательные значения длин волн x абсолютно не важны для этого модельного примера, поэтому мы будем использовать только некоторые условные «длины волн». Эти значения являются всего на всего числами, изменяющимися от 1 до 53.

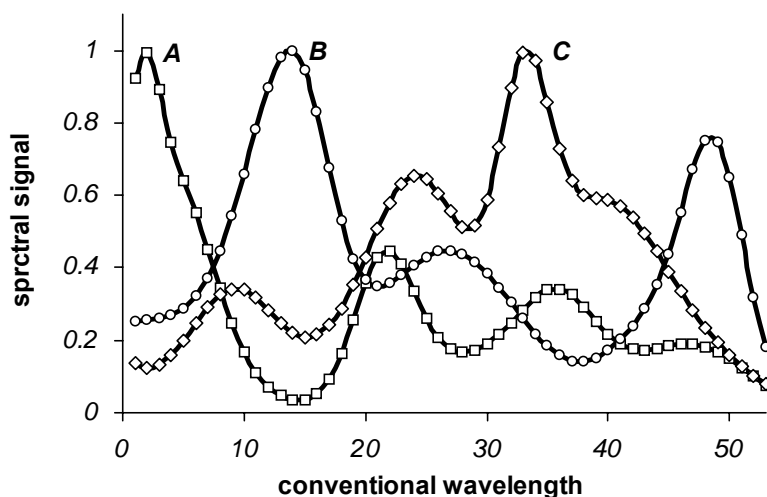


Рис. 7.2 Модельные спектры чистых компонент

Такие спектры представляются довольно сложными. Они были сконструированы для того, чтобы продемонстрировать, что предлагаемый подход может с успехом применяться даже для сложных спектральных данных.

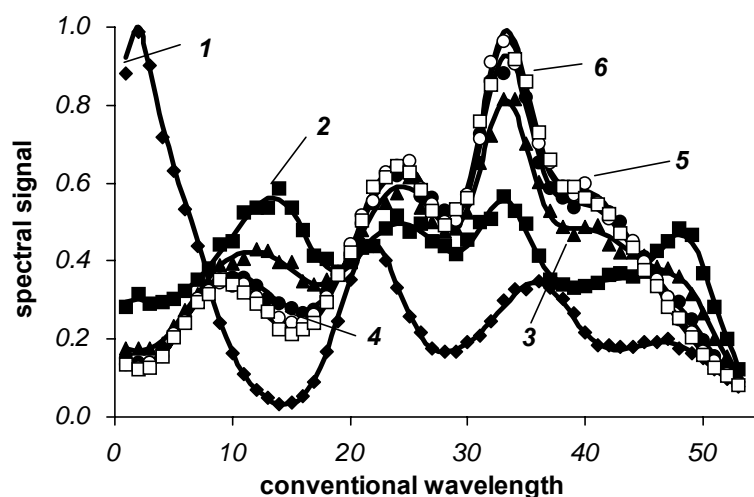


Рис. 7.3 Модельные «истинные» кинетические спектры (кривые) и соответствующие данные (точки). $t=0$ (1, \blacklozenge), $t=2$ (2, \blacksquare), $t=4$ (3, \blacktriangle), $t=6$ (4, \bullet), $t=8$ (5, \circ), $t=10$ (6, \square)

Матрица «экспериментальных» данных Y вычислялась по уравнению (7.2), где матрица концентраций C определялась по системе (7.5). Кроме того, к ней был добавлен «белый шум» с относительно ошибкой 3%.

Эти значения показаны на Рис. 7.3, где кривые представляют «точные» значения, а точками показаны «экспериментальные» данные содержащие случайные ошибки.

7.2. Метод ПБО для спектральных данных

Наша цель – это найти неизвестные кинетические параметры $\mathbf{k} = (k_1, k_2)^t$ из спектральных данных Y , которые представлены матрицей размерностью (6×53) . Если известны вектора спектров \mathbf{p} , \mathbf{q} и \mathbf{r} реагентов A , B и C , то получается довольно простая «обратная кинетическая задача» [3, 114] – найти минимум суммы квадратов отклонений –

$$\min_k \sum_{i=1}^m \sum_{j=1}^n [Y_{ij} - p_i A(t_i, \mathbf{k}) - q_i B(t_i, \mathbf{k}) - r_i C(t_i, \mathbf{k})]^2, \quad (7.6)$$

где функции A , B и C представлены уравнениями (7.5), а p_i , q_i , r_i – это известные заданные значения. Однако, если хотя бы один спектр \mathbf{p} , \mathbf{q} или \mathbf{r} неизвестен (а это – обычный случай) ситуация кардинально меняется. Практически очень трудно найти минимум суммы (7.6) по отношению к 161 неизвестному параметру (2 кинетических, плюс по 53 спектральных значений для каждого из 3 реагентов) в связи с проблемами плохо определенных матриц.

Метод последовательного байесовского оценивания может помочь в разрешении проблем оценивания. Подробное описание этого подхода было дано ранее в главе 2. Напомним, что главная идея ПБО состоит в разделении исходного массива данных на несколько частей. Затем оценивание параметров производится последовательно, порция за порцией. При этом результаты, полученные на предыдущем шагу, используются как априорные значения (в байесовской форме) для следующего шага. Первая (начальная) часть данных обрабатывается обычным методом наименьших квадратов (МНК) без априорной информации. По ходу процедуры ПБО строится последовательность оценок параметров. Последний член в этой последовательности является окончательной оценкой методом ПБО. Ранее было показано (см. раздел 2.3), что в линейном случае метод ПБО дает те же оценки, что и традиционный МНК. Отсюда следует, в частности, что окончательный результат ПБО не зависит от порядка обработки частей данных. В нелинейном случае ситуация более сложная, однако все эти свойства выполняются уже асимптотически.

Для рассматриваемой задачи оценивания кинетических параметров k_1 и k_2 , метод ПБО может быть представлен следующим алгоритмом. (Сравни алгоритм (2.38)).

Шаг 0. (Начальная стадия) Выбираются несколько длин волн (обычно 3-6) и соответствующие спектральные данные Y_1, Y_2, \dots обрабатываются совместно методом наименьших квадратов (Смотри (1.35)).

Шаг 1. Полученные результаты – оценки кинетических параметров, F -матрица, и другая апостериорная информация (см. (2.21)-(2.24)) – пересчитываются в априорную информацию в соответствии с формулам (2.34)-(2.37).

Шаг 2. Выбирается одна новая длина волны x_i и, соответствующие спектральные данные Y_i обрабатываются отдельно методом максимума правдоподобия (ММП), с учетом априорной информации, полученной на Шаге 1 (см. (2.17)).

Последний шаг. Шаги 1 и 2 повторяются до тех пор, пока не будут учтены все длины волн и получены окончательные оценки.

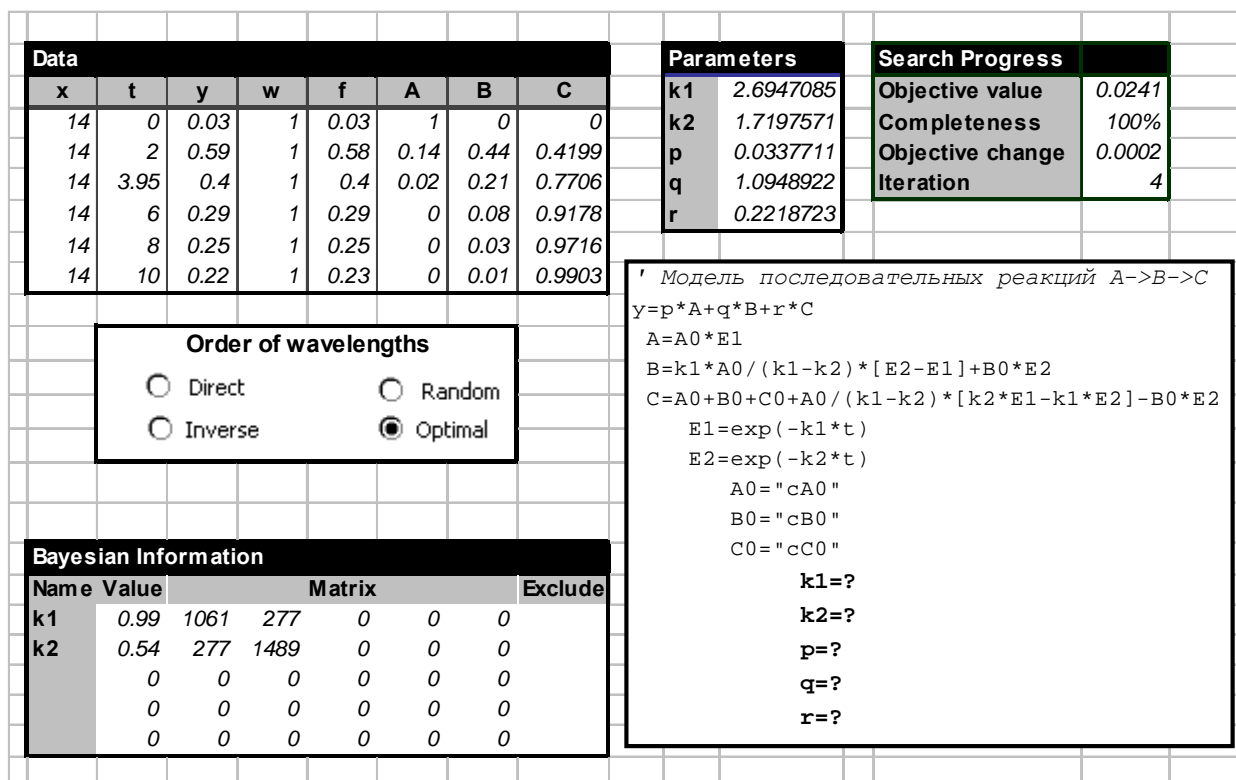


Рис. 7.4 Модель для расчета спектральных данных в системе Fitter

В соответствие с концепцией, изложенной в главе 2, здесь мы используем априорную информацию первого рода (см. раздел 1.1) т.к. нам известно, что взвешенная дисперсия ошибки одна и та же для разных длин волн. Кроме того, кинетические параметры k_1 и k_2 рассматриваются общими для всех длин волн, тогда как параметры p_i, q_i, r_i являются ча-

стными (см. раздел 2.3). Для расчетов использовалась программа Fitter, описанная выше в главе 5. При этом кинетика для каждой длины волны описывалась моделью, представленной на Рис. 7.4. Подробности можно найти в [126], где приведен файл *ABC.XLS*, содержащий исходные данные и детали обработки.

При расчетах одна и та же рутинная операция повторялась много раз для каждой длины волны. Поэтому была написана простейшая программа на языке VBA [7] с использованием функций Fitter, описанных в разделе 5.12. Фрагмент этой программы, в котором и осуществляется эта операция, представлен ниже на Рис. 7.5.

```
Function OneStep() As Integer
    Dim i As Integer
    OneStep = 1
    For i = 1 To 4
        If (dWaveValue = dSelected(i)) Then Exit Function
    Next i

    If (FitterData(sDataRange:="rData", sCodeString:="PPRWFIII", _
        kTitle:=-1) = 0) Then GoTo ErrorEnd

    If (FitterBayes(sBayesRange:="rBayes", lFreedom:=lBayesNDF, _
        dDispersion:=dBayesVar, kTitle:=-1) = 0) Then GoTo ErrorEnd

    If (FitterParameters(sParametersLeftTop:="rParam", _
        kTitle:=xtYes) = 0) Then GoTo ErrorEnd

    If (FitterFit(sProgressLeftTop:="rProgress", _
        sResultsLeftTop:="", _ dFitConvergence:=dConvergence, _
        nPauseEachStep:=0, nMaxIterations:=nBayesMax) = 0) _
        Then GoTo ErrorEnd

    If (FitterStatisticsBack(SumSquares:=dSS, _
        Objective:=dObj) = 0) Then GoTo ErrorEnd

    If (FitterParameters(sParametersLeftTop:="rParam", _
        kTitle:=xtYes) = 0) Then GoTo ErrorEnd

    If (FitterPosteriorInformation(sResultLeftTop:="rBayes") = 0) _
        Then GoTo ErrorEnd

    Range("rBayes").Cells(5, 1).Value = ""
    Range("rBayes").Cells(6, 1).Value = ""
    Range("rBayes").Cells(7, 1).Value = ""

    If (FitterStatisticsBack(VarianceFit:=dBayesVar, _
        NDF_Fit:=lBayesNDF, _ Estimates:=dEst, _
        Deviations:=dDev) = 0) Then GoTo ErrorEnd

    Exit Function
ErrorEnd:
    FitterError
    OneStep = 0
End Function
```

Рис. 7.5 Фрагмент программы для расчета спектральных данных в системе Fitter

7.3. Обработка модельных данных

Метод ПБО нуждается в начальной стадии, где кинетические данные обрабатываются без априорной информации – Шаг 0. Иногда эта стадия может вызвать проблемы. В рассматриваемом примере обработка кинетики на одной длине волны затруднительна, т.к. имеется только 6 измерений для оценки 5 неизвестных параметров. Только некоторые длины волн позволяют выполнить эту обработку, например, условная длина волны 16 (см. Рис. 7.6), но большинство – не позволяют сделать это.

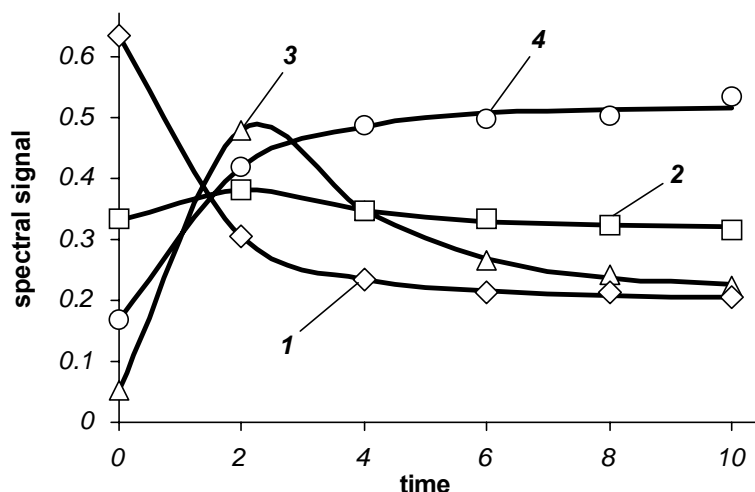


Рис. 7.6 Кинетические данные, используемые на начальной стадии ПБО со случайным порядком волн $\lambda=5$ (1, ◇), $\lambda=8$ (2, □), $\lambda=16$ (3, △), $\lambda=25$ (4, ○)

Однако можно собрать несколько длин волн и обрабатывать соответствующие кинетические данные совместно. В рассматриваемом примере оказалось, что кинетические данные для любых четырех длин волн могут решить проблему нулевого шага. Они содержат 24 измерения для оценки 14 неизвестных параметров. Причем, это могут быть первые, последние или любые четыре случайно выбранные длины волны. По результатам этой стадии может быть построена априорная информация, и процедура ПБО начинается.

На Рис. 7.6 приведен пример таких начальных данных. Эти кинетические кривые используются в случайной процедуре ПБО, описанной далее.

Известно, что в общем случае, порядок, в котором обрабатываются данные методом ПБО, влияет на результаты оценивания нелинейной модели. Однако, с практической точки зрения, этот эффект несущественен. Чтобы продемонстрировать это, мы провели последовательное оценивание для различного порядка условных длин волн в этом примере. Рассматривались следующие последовательности: прямой порядок (т.е., 1,2,3,4,5,...,53), обратный порядок (т.е., 53,52,51,50,49, ... ,1), и случайный порядок (т.е., 16,5,29,8,41,...).

Первые четыре числа в этих последовательностях представляют длины волн, используемые на начальной стадии. Результаты показаны на Рис. 7.7 (графики а-с)

Здесь толстые кривые (**1** и **2**) показывают, как меняются оценки кинетических параметров в ходе последовательного оценивания. Окрашенные области (**1a**, **1b** и **2a**, **2b**) вокруг этих кривых демонстрируют неопределенность в оценках – они образованы стандартными отклонениями прибавленными (вычтенными) к оценкам параметров. На каждом графике показаны обе оценки: верхняя – это k_1 и нижняя – это k_2 . Пунктирные линии изображают «истинные» значения кинетических констант. Все данные изображены в зависимости от чисел (ось X), которые представляют условные длины волн в порядке (слева направо) в котором они используются в процедуре ПБО. Первые четыре точки на каждом графике показывают результаты оценивания на начальной стадии.

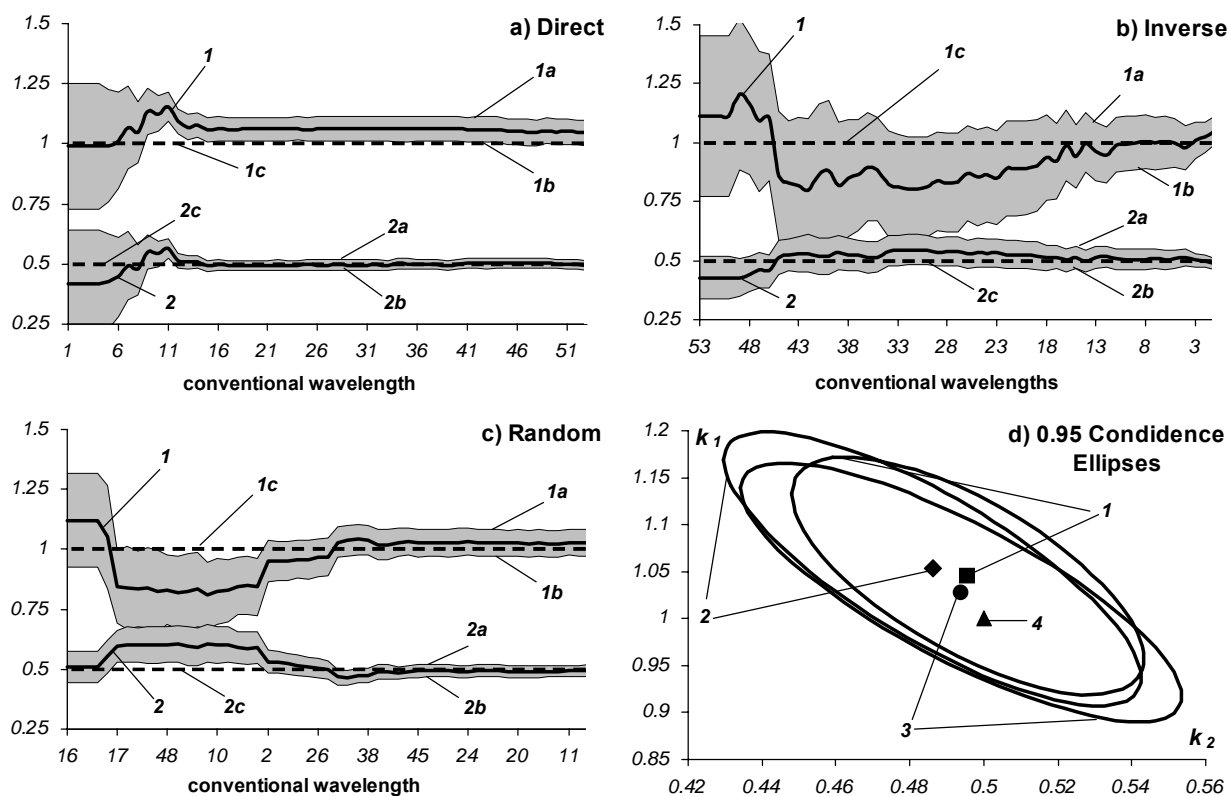


Рис. 7.7 Промежуточные результаты последовательного оценивания кинетических констант в зависимости от порядка длин волн (а-с) и 0.95-доверительные эллипсы для окончательных оценок (d). На графиках (а-с) толстые кривые представляют оценки k_1 (1) и k_2 (2); тонкие кривые (1a, 1b, 2a, 2b) показывают границы областей стандартных отклонений; пунктирные линии (1c, 2c) обозначают «истинные» значения параметров. На графике (d) эллипсы и точки представляют: прямой (1, ■), обратный (2, ◆) и случайный (3, ●) порядок волн, точка (4, ▲) показывает «истинные» значения кинетических констант.

Из графиков видно, что хотя при разном порядке обработки, промежуточные результаты оценивания разные, но окончательные оценки очень близки. Последний график (d) на Рис. 7.7 иллюстрирует этот вывод. На нем показаны 0.95 доверительные эллипсы для всех

окончательных результатов ПБО. Каждый эллипс и значок в его центре представляет результат ПБО с соответствующим порядком обработки спектров. Сравнивая траектории для разных способов обработки можно заметить, что промежуточные оценки, так же как и их неопределенности, действительно зависят от порядка длин волн. По-видимому, наилучшим является прямой порядок (график **a**), для которого и оценки и отклонения изменяются гладко и медленно без больших прыжков. Другой интересный результат можно видеть на графике (**b**), где показано последовательное оценивание с обратным порядком длин волн. Начальная стадия дает очень плохую оценку параметра k_1 . В последствии неопределенность уменьшается. Однако, эти значения все еще далеки от окончательных результатов и только две последние длины волны 2 и 1 резко изменяют оценку и ее отклонение, поднимая их до окончательного, общего уровня.

Разумеется, все эти выводы справедливы только для рассматриваемого примера, в котором спектры устроены так, как это показано на Рис. 7.2. Для других случаев наилучший порядок может быть совершенно другим. Рассмотрим этот вопрос подробнее, начав с того, что вызывает наибольшие проблемы в ПБО – с нулевого шага.

Из Рис. 7.7 видно, что удачный выбор начальных длин волн может значительно улучшить оценивание (график **a**), так и плохой выбор может ухудшить его (график **b**). Для того чтобы сделать выбор порядка расчетов автоматическим, мы предлагаем использовать следующий простой прием. Совершенно ясно, что кинетические кривые, представляющие большее изменение, являются более «информативными». Кроме того, хорошо обрабатываются и не монотонные кривые с экстремумами. Объединяя все эти идеи, мы предлагаем следующий эмпирический критерий для оценивания сравнительной «информативности» кинетических кривых.

Это – следующее выражение

$$L = \sum_{j=1}^n \sqrt{(t_j - t_{j-1})^2 + (Y_j - Y_{j-1})^2} - (t_m - t_1).$$

Здесь t_j – это значения времени (предиктор), Y_j – это значения кинетики (отклик) и n – это число измерений. Легко видеть, что главный член в этом выражении – это длина кривой. Чем больше величина критерия L , тем более «информативна» кинетическая кривая. Прямая линия, параллельная оси t имеет наименьшую «информативность» $L=0$. Например, кривые, показанные на Рис. 7.6 имеют следующие показатели: $L(16)=0.0476$, $L(5)=0.0286$, $L(29)=0.0174$, $L(8)=0.0007$. Если упорядочить все длины волн в порядке убывания «ин-

формативности» соответствующих кинетических данных, то можно ожидать «оптимального» результата ПБО. Рис. 7.8 подтверждает правильность такого подхода.

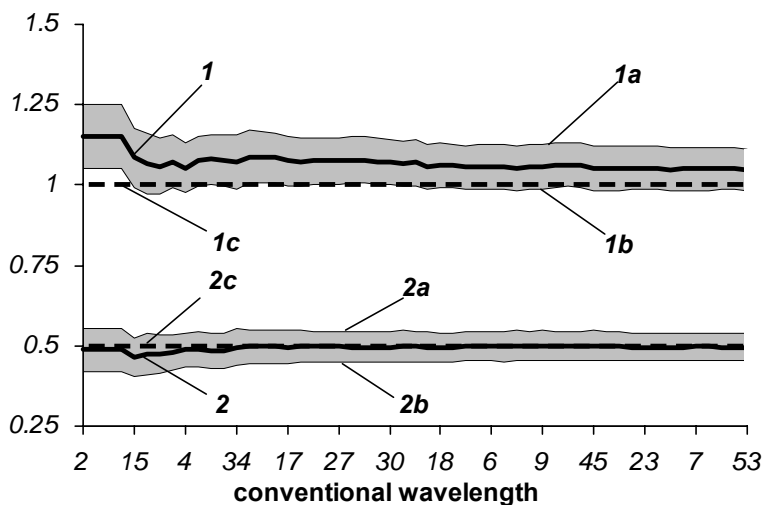


Рис. 7.8 Промежуточные результаты последовательного оценивания кинетических констант для «оптимального» порядка длин волн. Толстые кривые представляют оценки k_1 (1) и k_2 (2); тонкие кривые (1a, 1b, 2a, 2b) показывают границы областей стандартных отклонений; пунктирные линии (1c, 2c) обозначают «истинные» значения параметров.

После того, как найдены оценки общих кинетических параметров k_1 и k_2 , естественно определить и частные спектральные параметры p , q , и r . Разумеется, это можно сделать очень легко. Если константы скоростей реакций фиксированы равными их оцененным значениям, то набор спектральных параметров p_i , q_i , r_i может быть получен для каждой длины волны i линейным МНК. Однако погрешности этих оценок будут получены не верно, потому что МНК подход не может учесть неопределенности в фиксированных кинетических параметрах. Простейший пример, разъясняющий эту идею, был разобран в разделе 2.5.

Здесь можно использовать метод обратного последовательного байесовского оценивания (ОПБО), который был представлен в разделе 2.4. Для этого на последнем шаге процедуры ПБО создается «окончательная» апостериорная информация, которая затем трансформируется в «окончательную» априорную информацию с помощью формул (2.34)-(2.35). Эта информация содержит данные только о кинетических параметрах.

Эта информация используется как априорная информация второго рода (см. раздел 2.4) для каждой волны i , когда оценивается набор спектральных параметров p_i , q_i , r_i . При этом оценивание должно проводиться в соответствие с формулами (2.50)-(2.52), которые обеспечивают выполнение дополнительных ограничений на кинетические параметры, которые уже оценены и не должны переоцениваться вновь.

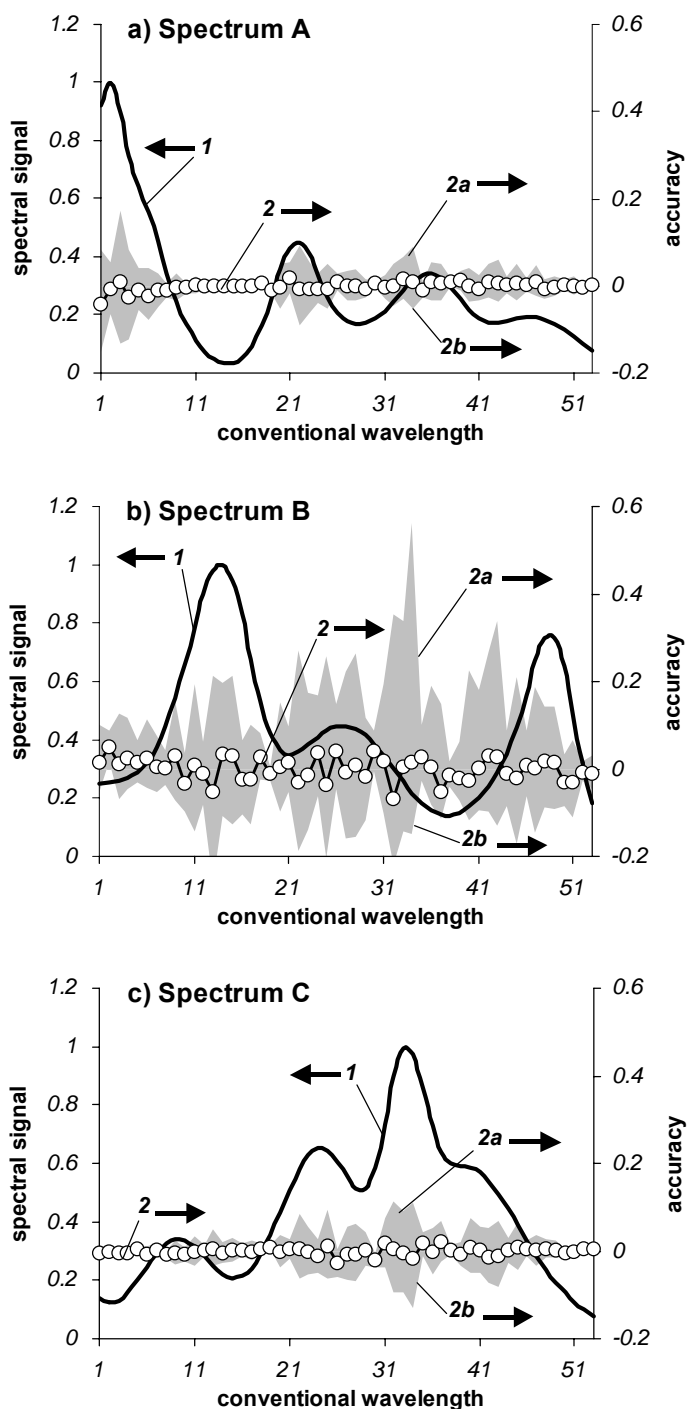


Рис. 7.9 Модельные спектры реагентов (толстые кривые, **1**, левая ось Y). Разница между оценкой и «истинным» спектром (точки **2**, правая ось Y). Области утроенных стандартных отклонений (окрашенные области **2a**, **2b**, правая ось Y)

На Рис. 7.9 на каждом графике представлены «истинные» модельные спектры (толстые кривые **1**, левая ось Y). Результаты оценивания показаны как разница между оцененным и «истинным» спектром (точки **2**). Также изображены и области утроенных стандартных отклонений (окрашенные области **2a**, **2b**). Все эти данные относятся к правой оси Y. Из графиков видно, что найденные оценки очень точны, особенно для компонент A и C.

7.4. Проверка метода ПБО на модельных данных

Для данного примера можно проверить, насколько надежен метод ПБО. Для этого мы обработали все модельные «экспериментальные» данные с помощью МНК. В ходе поиска одновременно оценивались все 161 параметра в сумме (7.6). При этом использовалась очень большая модель, показанная на Рис. 7.10. В ней для экономии места пропущены значительные фрагменты текста, которые, впрочем, легко восстановить по аналогии с соседними строками.

```
'Модель для одновременного оценивания всех параметров
y=P*A+Q*B+R*C
A=exp(-k1*t)
B=k1*1/(k1-k2)*[exp(-k2*t)-exp(-k1*t)]
C=1+1/(k1-k2)*[k2*exp(-k1*t)-k1*exp(-k2*t)]
P=P1+P2+P3+P4+P5+P6+P7+P8+P9+P10+P11
Q=Q1+Q2+Q3+Q4+Q5+Q6+Q7+Q8+Q9+Q10+Q11
R=R1+R2+R3+R4+R5+R6+R7+R8+R9+R10+R11
    P1=p1*imp(x-1)+p2*imp(x-2)+p3*imp(x-3)+p4*imp(x-4)+p5*imp(x-5)
    . . .
    P11=p51*imp(x-51)+p52*imp(x-52)+p53*imp(x-53)
    Q1=q1*imp(x-1)+q2*imp(x-2)+q3*imp(x-3)+q4*imp(x-4)+q5*imp(x-5)
    . . .
    Q11=q51*imp(x-51)+q52*imp(x-52)+q53*imp(x-53)
    R1=r1*imp(x-1)+r2*imp(x-2)+r3*imp(x-3)+r4*imp(x-4)+r5*imp(x-5)
    . . .
    R11=r51*imp(x-51)+r52*imp(x-52)+r53*imp(x-53)
    p1=?
    q1=?
    r1=?
    . . .
    p53=?
    q53=?
    r53=?
```

Рис. 7.10 Модель в системе Fitter для одновременной обработки всех спектральных данных.

Результаты проведенного теста показаны на Рис. 7.11.

Точки в центре эллипсов **1** и **2** показывают значения оценок, полученных с помощью методов ПБО и МНК, соответственно. Они расположены очень близко, так же, как и соответствующие доверительные области, представленные эллипсами. Этот расчет подтверждает результат теоремы, доказанной ранее в разделе 2.3, состоящий в том, что ПБО оценки асимптотически сходятся к МНК оценкам.

Интересно выяснить, можем ли мы «доверять» этим доверительным областям, построенным с помощью метода стохастической аппроксимации, описанным в разделе 3.1. Разумеется, это зависит от того, насколько сильна нелинейность модели, использованной в модельном примере.

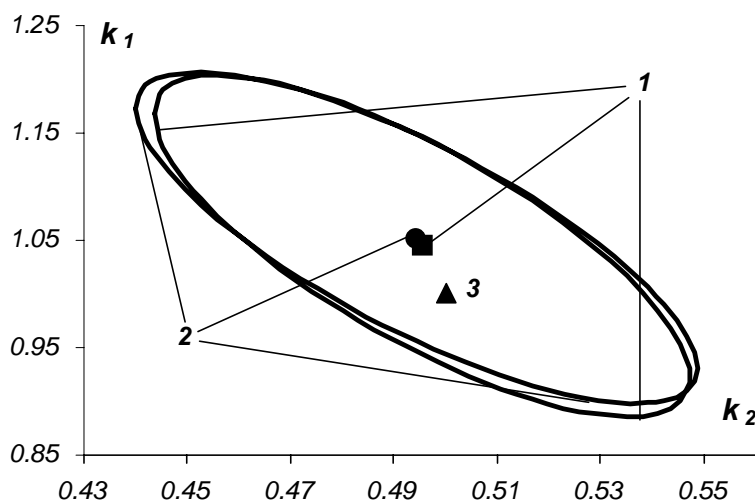


Рис. 7.11 Результаты оценивания, представленные 0.95 -доверительными эллипсами. ПБО (1, ■), МНК (2, ●) и «истинные» значения (3, ▲).

Как было показано в разделе 3.4, мерой нелинейности может служить коэффициент нелинейности γ , вычисляемый методом Монте-Карло. Для линейной (или похожей на линейную) модели коэффициент γ должен быть меньше чем 1 . Чем дальше этот коэффициент от единицы, тем модель более нелинейная. Вычисления, проведенные для рассматриваемого примера, дали значение $\gamma = 1$. Тем самым показано, что эта модель очень близка к линейной и, что построенные доверительные области надежны.

7.5. Реальный пример

Этот пример происходит из работ [112, 116]. В них рассматривается набор данных, состоящих из ИК спектров, снятых в ходе эпоксидизации 2,5-ди-*терт*-бутил-1,4-бензохинона. Реальный экспериментальный процесс описывается двухстадийной реакцией (7.3). Реагенты, методика измерений и другие детали эксперимента описаны подробно в [112]. Исходные экспериментальные данные можно найти в [117]. Там приведены 240 спектров, снятых на интервале 800-1100 нм с шагом 1.0 нм. Время реакции составляет 1200 с, а измерения проводились через каждые 5 с.

Исходные данные были подготовлены к анализу с соответствии с процедурой, описанной в [112]. Приведем ее краткое описание. Четвертый спектр (при $t=20$ с) используется как базовый и вычитается из всех других спектров. Затем для новых спектров вычисляется вторая производная с применением фильтра Савицкого-Голэя [118] с окном из 15 точек. Наконец, для расчетов использовался только узкий диапазон длин волн 860-880 нм. Эти данные представлены на Рис. 7.12.

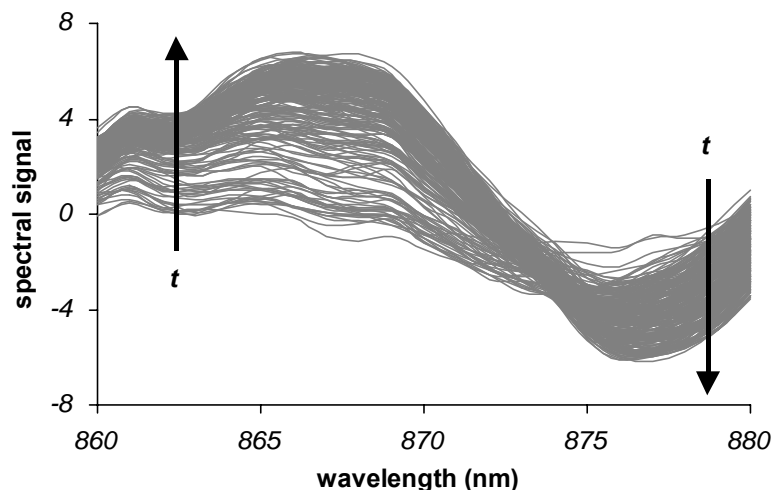


Рис. 7.12 Спектры, полученные в реальном примере

Они были обработаны процедурой ПБО с прямым порядком длин волн. Первые четыре длины волны (860, 861, 862 и 863) использовались на начальной, иницирующей стадии. Результаты оценивания показаны на Рис. 7.13, где изображено как меняются оценки кинетических параметров и их точность по ходу последовательной процедуры.

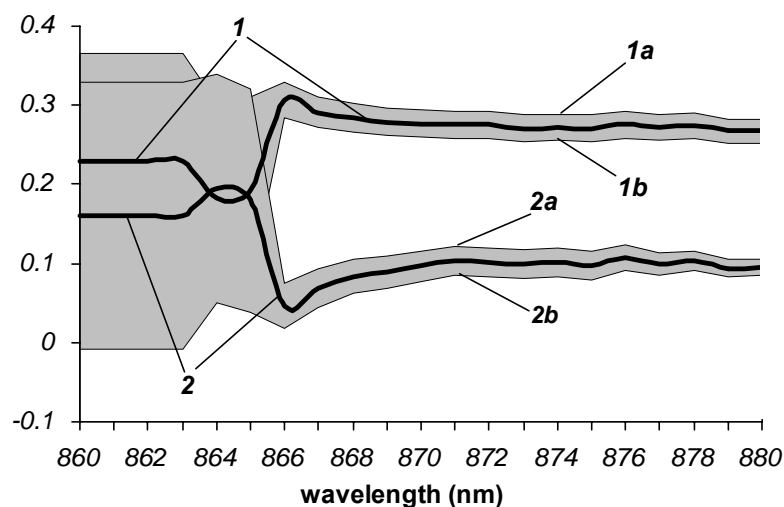


Рис. 7.13 Промежуточные результаты последовательного оценивания кинетических констант в реальном примере. Толстые кривые представляют оценки k_1 (1) и k_2 (2); тонкие кривые (1a, 1b, 2a, 2b) показывают границы областей стандартных отклонений.

Устройство этого графика аналогично Рис. 7.7 а. Окончательные значения оценок констант, которые были достигнуты на последнем шагу для длины волны 880 нм, составили: $k_1 = 0.267 \pm 0.015$ (мин⁻¹), $k_2 = 0.095 \pm 0.010$ (мин⁻¹). Здесь указаны также и стандартные отклонения. Коэффициент корреляции $r = -0.18$.

Интересно сравнить эти результаты с аналогичными оценками, полученными с использованием других, традиционных методов, на том же массиве экспериментальных данных.

Все эти методы подробно описаны в [116], поэтому здесь приводится только краткая информация о них.

Метод взвешенной подгонки кривых (weighted curve resolution – WCR) [119] объединяет формальный подход, в котором используется сингулярное разложение матрицы данных Y , и кинетический подход для вычисления матрицы концентраций C .

Обобщенный метод аннигиляции ранга (generalized rank annihilation method – GRAM) [120] – это формальный подход, в котором используется простое уравнение

$$\frac{e^{-kt}}{e^{-k(t+s)}} = e^{ks}$$

показывающее, что константа скорости реакции первого порядка может быть определена из отношения исходной и сдвинутой экспоненты. Наконец, используется формальный метод, названный в работе [112] LM-PAR. Он улучшает оценки полученные методом GRAM, за счет использования алгоритма Левенберга-Маркварда [67, 68] (см. раздел 4.1) и процедуры PARAFAC [121]. Нужно отметить, что и GRAM и LM-PAR могут быть применены только для моделирования кинетики (псевдо) первого порядка.

Оценки, полученные этими методами, приведены в Табл. 7.1

Табл. 7.1 Результаты оценки кинетических параметров в реальных данных различными методами. Приведены также среднеквадратичные отклонения оценок и коэффициент корреляции.

| Метод | Оценка k_1 мин ⁻¹ | СКО k_1 мин ⁻¹ | Оценка k_2 мин ⁻¹ | СКО k_2 мин ⁻¹ | Коэффициент корреляции |
|--------|-----------------------------------|--------------------------------|-----------------------------------|--------------------------------|---------------------------|
| ПБО | 0.27 | 0.02 | 0.10 | 0.01 | -0.18 |
| WCR | 0.26 | 0.03 | 0.07 | 0.03 | -0.61 |
| GRAM | 0.28 | 0.03 | 0.10 | 0.06 | -0.47 |
| LM-PAR | 0.25 | 0.04 | 0.09 | 0.04 | -0.79 |

Те же результаты показаны на Рис. 7.14, где каждый метод представлен 0.95-доверительным эллипсом, построенным по данным Табл. 7.1.

Сравнивая размеры и формы эллипсов, показанных на этом рисунке, можно легко сопоставить разные методы. Так видно, что эллипс **I** имеет наименьший размер и наиболее близкую к кругу форму. Это означает, что метод последовательного байесовского оценивания (ПБО) дает самые маленькие значения ошибок оценивания и корреляцию среди всех сравниваемых методов.

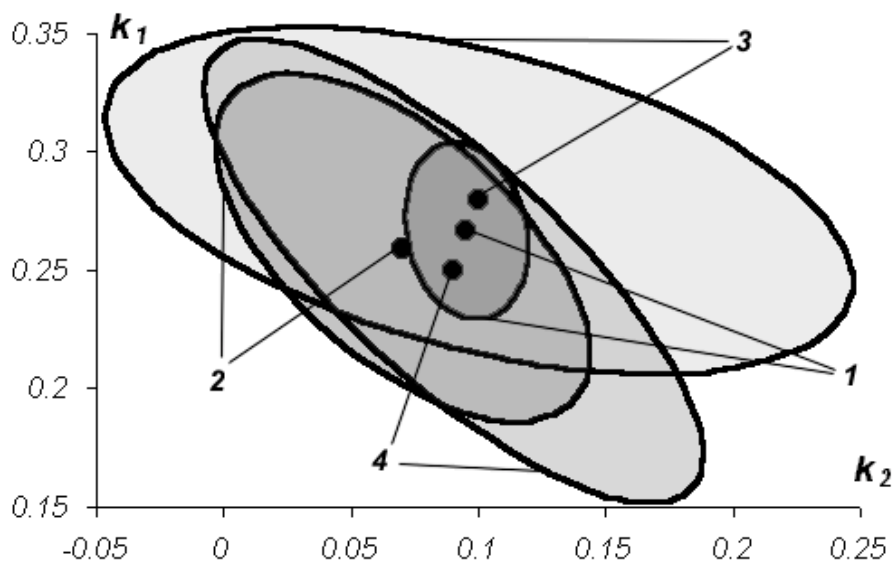


Рис. 7.14 Результаты, полученные для реального примера различными методами. Каждый метод представлен 0.95-доверительным эллипсом. ПБО (1), WCR (2), GRAM (3), LM-PAR (4).

7.6. Результаты главы 7

В этой главе мы рассмотрели, как метод последовательного байесовского оценивания (ПБО) может быть применен для определения констант скоростей химической реакции по спектральным данным для случая, когда индивидуальные спектры реагентов неизвестны. Все изложение велось для примера последовательных реакций первого порядка, однако, важно отметить, что этот подход имеет самую общую природу и может быть применен также и для других кинетических моделей. Достоверность метода была продемонстрирована на модельном и реальном примерах. Модельный пример показал, что предлагаемая методика эффективна как для малого числа наблюдений, так и для сильно перекрывающихся спектров. Также было показано, что метод ПБО дает результаты, близкие к значениям, получаемым традиционным МНК. Из реального примера можно сделать вывод, что метод ПБО дает самые маленькие погрешности оценивания по сравнению с некоторыми другими, традиционными подходами.

Метод ПБО – это довольно быстрый способ оценивания, но его скорость неизвестна априори. Однако расчеты в обоих примерах заняли всего несколько минут. Важно отметить, что этот метод является Байесовским только по форме, но не по сути. Это означает, что для его применения не требуется привлекать никакой *субъективной* априорной информации. Каждый элемент используемой априорной информации однозначно строится по результатам предыдущего шага процедуры, и только форма применения этой информации диктуется теоремой Байеса. Никаких дополнительных предположений (числа главных

компонент, сдвига по времени, первого порядка реакции) не требуется для использования этого метода.

8. Прогнозирование старения эластомерных материалов

Прогнозирование эксплуатационной устойчивости материалов – одна из сложнейших задач математического моделирования [124]. Под эксплуатацией здесь имеется в виду все возможные варианты использования изделия, содержащего исследуемый материал: хранение, применение. Сроки хранения (службы) многих материалов свыше 10 лет и, как правило, статистика отказов в условиях эксплуатации неизвестна. Наиболее распространенным методом прогнозирования является метод ускоренных испытаний (УИ), когда материал (изделие, образец) подвергается воздействию более жестких, нежели в реальности, внешних условий x (факторов старения), вызывающих ускоренное изменение его свойств. По результатам таких испытаний строится математическая модель, описывающая изменение практически важных свойств y со временем t и в зависимости от величин факторов старения x ,

$$y=f(t, x, a).$$

Эта модель зависит от неизвестных параметров a , которые оцениваются по данным УИ методом максимума правдоподобия (см. главу 1). При этом оценки параметров \hat{a} определяются как значения, при которых целевая функция имеет минимум. После того, как найдены оценки всех неизвестных параметров модели, можно провести ее экстраполяцию на условия эксплуатации. Если известны критические значения измеряемых откликов, то нетрудно определить и срок, в течение которого это значение будет достигнуто [38, 39].

Так, в принципе, выглядит процедура прогнозирования эксплуатационной устойчивости материалов. Однако в ней содержится несколько сложных моментов, которые заслуживают специального рассмотрения.

Как правило, при старении материала меняются все его практически-важные свойства и в ходе УИ контролируются несколько откликов. Для каждого такого отклика необходимо построить свою модель, описывающую его изменение. Поэтому для оценивания модели приходится оперировать не с одной, а с несколькими целевыми функциями – по одной на каждый отклик. Дело осложняется тем, что в моделях, описывающих разные отклики, могут быть общие параметры. Наличие общих параметров – это естественное дело, т.к. старение это единый процесс, обуславливающий изменение всех свойств материала. Поэтому, такие параметры как энергии активации и предэкспоненты констант скоростей химических и физических процессов, должны одновременно входить во все уравнения модели. В этом случае мы не можем оценивать параметры, минимизируя каждую целевую функ-

цию отдельно – ведь тогда для каждого отклика мы получим свои (и вообще говоря, разные) оценки общих параметров. Традиционный статистический подход к такой проблеме состоит в том, что все целевые функции нужно сложить, причем каждую со своим весом-множителем, и искать минимум этой полной суммы. Необходимость введения весов диктуется тем обстоятельством, что ошибки измерения каждого отклика разные и их необходимо уравновесить при сложении, иначе показатель, измеряемый с точностью 5%, «забьет» показатель, измеряемый с точностью 30%. Как правило, величины ошибок измерения нам не известны, поэтому определение весов сводится к большой итерационной процедуре. Как альтернатива этому подходу был предложен метод последовательного байесовского оценивания [30], позволяющий обрабатывать отклики последовательно, по одному (см. главу 2). Этот метод широко используется для интерпретации экспериментальных данных [32, 42, 43, 44]. В этой главе мы применим его для прогнозирования срока службы резин [46].

В практических задачах часто случается, что минимум целевой функции определяется очень плохо. С геометрической точки зрения это выглядит как вырождение поверхности, когда она становится похожа на овраг без явно выраженной минимальной точки. Эта ситуация называется мультиколлинеарностью. В разделе 4.4 эта проблема была подробно рассмотрена и показано, что имеются два случая мультиколлинеарности – «строгая» и «нестрогая». В первом случае в модели присутствуют некоторые внутренние (может быть неявные) связи между параметрами, которые не позволяют определить все параметры в принципе. Второй случай мультиколлинеарности вызывается плохим планом эксперимента и может быть, в принципе, разрешен его изменением. Критерий «правильности» плана эксперимента выразить довольно сложно. Он зависит от ошибок эксперимента, от точности вычислений и может быть улучшен введением шкалирования и масштабирования данных и модели. Первый случай мультиколлинеарности чаще всего встречается при содержательном моделировании кинетики химических реакции, тогда как второй случай имеет место, прежде всего, при формальном описании наблюдаемых процессов. Прогнозирование эксплуатационной устойчивости материалов – это как раз такой случай.

Все перечисленные выше трудности кажутся все же незначительными на фоне основной проблемы прогнозирования – экстраполяции за пределы области наблюдений. Чем мы можем гарантировать правильность этой экстраполяции? Что может подтвердить правильность модели выбранной для формального описания процессов старения? Классические методы проверки правильности моделирования – использование тестового массива и

перекрестная самопроверка – здесь не работают, т.к. они пригодны лишь для задач интерполяции. Строго говоря, в нашем случае никакие математические методы в принципе не могут дать нужный ответ. С другой стороны, очевидно, что и содержательное физико-химическое моделирование с использованием сложных кинетических схем для описания процессов деградации материала, не сможет дать этой гарантии. Любое усложнение модели в условиях ограниченного по времени и по содержанию эксперимента не только не улучшит надежность прогноза, но и приведет к переопределенности модели и, как следствие, к строгой мультиколлинеарности всей задачи. Конечно, старение материала – это сложный процесс, в котором происходят множество химических и физических превращений, характеризуемых большим набором параметров. Однако их можно ранжировать, упорядочить по степени влияния на процесс старения. В большинстве практических задач можно ожидать, что старение определяется только одним-двумя, редко тремя параметрами. Для их правильного выбора и определения важно соблюдать следующие условия. Во-первых, эти параметры должны присутствовать в математических описаниях всех измеряемых параметров – т.е. они должны быть общими. Во-вторых, план УИ должен быть выстроен так, чтобы можно было заметить и идентифицировать влияние этих параметров – иными словами, условия эксперимента не должны слишком отдаляться от условий эксплуатации. Здесь можно заметить прямую аналогию с проекционными методами, применяемыми при линейном моделировании многомерных данных [122, 123] – выбор главных компонент. Однако есть и существенное отличие – доказательство правильности модели проводится через построение внутренне согласованной, непротиворечивой модели. В разделе 8.2 мы покажем, как это делается для нашего примера.

Описанная здесь процедура прогнозирования требует проведения довольно сложных математических операций с данными и моделями. Метод последовательного байесовского оценивания (ПБО), описанный в главе 2 позволяет выполнить все эти операции и получить требуемый результат.

8.1. Эксперимент

Возможности данного подхода иллюстрируются примером прогнозирования термоокислительного старения шинных резин. Образцы материала – пластины размером 200мм×200мм×4мм – были приготовлены американской компанией Uniroyal Chemical, специально для этого эксперимента. Нам известны следующие характеристики материала. Образцы представляли саженаполненный серный вулканизат смеси натурального и бутадиенового каучуков. Вулканизация: температура 160 С, время 16 мин. Назначение: про-

тектор шины. Из пластин вырубались стандартные образцы, имеющие форму лопаток, которые подвергались ускоренному старению в соответствии с процедурой, описанной ниже. В ходе старения измерялись свойства, приведенные в Табл. 8.1.

Табл. 8.1 Свойства, измеряемые при ускоренном старении.

| № | Свойство | Обозначение | Ед. измер. |
|-----|--|-----------------|------------|
| 1-5 | Значения условного модуля (напряжения) при относительном удлинении $ELN=1,2,3,4,5$ | $MOD (ELN,t,T)$ | КПа |
| 6 | Прочность (разрывное напряжение) | $TEN (t,T)$ | КПа |
| 7 | Относительное удлинение при разрыве | $ELB (t,T)$ | 1 |

Измерения проводились на разрывной машине Instron. Время измерялось в часах, а температура в градусах Цельсия. Относительное удлинение (ELN и ELB) измерялось в условных единицах, которые соответствуют традиционным процентам: 1- 100%, 2 – 200%, и т.д.

Ускоренные испытания – это искусственное старение материала при нескольких (обычно трех) постоянных температурах. План УИ имеет большое значение для получения достоверного прогноза. При его составлении необходимо решить две главные задачи. Во-первых, в ходе УИ должна быть достигнута достаточная глубина старения за короткий срок. Очевидно, что изменение свойств при ускоренном старении должно быть не менее, а лучше – больше, чем критическое значение свойства, характеризующее отказ. Это соображение заставляет увеличивать температуру испытаний. Во-вторых, при УИ необходимо воспроизвести механизм старения в условиях эксплуатации, что принуждает снижать температуру. В работе [125] описывается процедура эволюционного планирования эксперимента (ЭПУИ), которая помогает решить эти задачи. Эта процедура состоит из трех стадий (см. Рис. 8.1).

Сначала вводится информация о материале и о режимах вулканизации. По этим данным программа устанавливает температуру T_{max} предварительных испытаний и два минимальных времени съема. Температура T_{max} выбирается на 15 - 30 градусов ниже температуры вулканизации T_{cur} .

На второй стадии по этому плану проводятся испытания, результаты которых заносят в таблицу. Если эти данные удовлетворяют критериям полноты и достаточно быстрого убывания, то процедура позволяет перейти к третьей стадии. В противном случае предлагается провести дополнительные эксперименты по дополненному плану.

На третьей стадии, используя результаты предварительных испытаний, можно вычислить время достижения необходимой глубины старения для любой заданной температуры, которая ниже чем T_{max} .

На Рис. 8.1 показан внешний вид этой процедуры.

Задача #1C NR/BR Passenger tire tread

| Информация о материале | | Предварительный план | |
|---|-------------------------------------|----------------------|------------------------------------|
| Темп-ра вулканизации | <input type="text" value="160"/> С | Температура | <input type="text" value="140"/> С |
| Время вулканизации | <input type="text" value="16"/> мин | Время 1 | <input type="text" value="1"/> час |
| С серой | <input checked="" type="checkbox"/> | Время 2 | <input type="text" value="2"/> час |
| Разрывное удлинение (планируемое значение) | <input type="text" value="200"/> % | Время 3 | <input type="text" value="?"/> час |

Изменить температуру

1-я Стадия

Укажите здесь информацию о материале и программа вычислит предварительный план

Экспериментальные данные (ГОСТ 9.713-86)

| Температура | | 140 С | | | |
|-------------|--------------------------------------|-------|-----|-----|-----|
| время, час | Относительное разрывное удлинение, % | | | | |
| 0.0 | 492 | 480 | 520 | 528 | 532 |
| 2.0 | 260 | 252 | 260 | 260 | 240 |
| 4.0 | 180 | 180 | 180 | 192 | |

2-я Стадия

Проведите УИ в соответствии с предварительным планом и введите результаты в таблицу.

План ускоренных испытаний

| | | | |
|-------------|------------------------------------|-----------------|--------------------------------|
| Температура | <input type="text" value="110"/> С | Число измерений | <input type="text" value="5"/> |
| время, час | 0 5.5 11 27 42 | | |

3-я Стадия

Введите желаемую температуру старения и программа вычислит моменты измерений

| Темп, С | 0 | 2.1 | 4.1 | 7.7 | 12 |
|---------|---|-----|-----|-----|----|
| 140 | 0 | | | | |
| 125 | 0 | 2.1 | 4.1 | 7.7 | 12 |
| 110 | 0 | 5.5 | 11 | 27 | 42 |

Рис. 8.1 Программа эволюционного планирования эксперимента (ЭПУИ)

В Табл. 8.2 приведен план эксперимента, предложенный программой ЭПУИ и план, реализованный в эксперименте.

Табл. 8.2 План ускоренных испытаний

| | Температура (С) | | Время (час) | | |
|-------------------------------------|-----------------|-----|-------------|------|------|
| Предложено ЭПУИ | 140 | 0.9 | 1.7 | 2.8 | 3.9 |
| | 125 | 2.1 | 4.1 | 7.7 | 12.0 |
| | 110 | 5.5 | 11.0 | 27.0 | 42.0 |
| Проведено в эксперименте | 140 | 2.0 | 4.0 | 6.0 | 9.0 |
| | 125 | 3.0 | 6.0 | 14.0 | 23.0 |
| | 110 | 8.0 | 16.0 | 38.0 | 60.0 |

Расчетная продолжительность старения оказалась меньше, чем время, отведенное для экспериментов, поэтому реальные времена съёмов были выбраны несколько больше, чем предлагалось. Результаты некоторых экспериментов приведены на Рис. 8.3

8.2. Модель

По-видимому, наилучшими данными для прогноза механических свойств резин были бы деформационные кривые $STR(ELN)$ – диаграммы “напряжение (STR) как функция удлинения (ELN)”, полученные для каждого образца. Преимущества такого подхода очевидны, т.к. в нем в единой компактной форме строится «траектория старения» материала [44], отражающая связь между изменениями различных свойств в ходе процесса. Существование единой, не зависящей от температуры испытаний, траектории старения означает, что все наблюдаемые изменения свойств обусловлены протеканием единого процесса. В простейшем случае это означает наличие во всех моделях нескольких общих кинетических параметров

Если построить модель деформационной кривой $STR(ELN, t, T)$, справедливую для любых значений удлинения ELN , времени t и температуры T , то модель для каждого условного модуля непосредственно следует из нее, например –

$$MOD(2, t, T) = STR(2, t, T). \quad (8.1)$$

Если построить также модель разрывного удлинения $ELB(t, T)$, то из этих двух зависимостей вытекает модель разрывного напряжения

$$TEN(t, T) = STR(ELB(t, T), t, T). \quad (8.2)$$

Кроме того, имея деформационные кривые, полученные для каждого образца в отдельности, можно установить, чем один образец отличается от другого и, тем самым разделить ошибки измерения и неоднородность образцов [39]. Однако такие измерения не проводи-

лись, и мы были вынуждены восстанавливать деформационные кривые по имеющимся экспериментальным данным.

Для описания эксперимента использовались две базовые модели:

- деформационной кривой

$$STR(ELN, t, T) = [b_0 + b_1 e^{-K_1 t} + b_2 e^{-K_2 t}] ELN + b_3 (1 - e^{-ELN b_4}). \quad (8.3)$$

- разрывного удлинения

$$ELB(t, T) = a_0 + a_1 e^{-K_1 t} + a_2 e^{-K_2 t} \quad (8.4)$$

Из которых, как было указано выше (формулы (8.1) и (8.2)), вытекают вспомогательные модели

- разрывного напряжения

$$TEN(t, T) = [b_0 + b_1 e^{-K_1 t} + b_2 e^{-K_2 t}] ELB(t, T) + b_3 (1 - e^{-ELB(t, T) b_4}) \quad (8.5)$$

- условного модуля

$$MOD(ELN, t, T) = [b_0 + b_1 e^{-K_1 t} + b_2 e^{-K_2 t}] ELN + b_3 (1 - e^{-ELN b_4}). \quad (8.6)$$

На Рис. 8.2 показана модель (8.5) для описания разрывного напряжения, в том виде, в котором она используется в системе Fitter.

```
'Модель разрывного напряжения
TEN=HMOD*ELB+b3*[1-exp(-b4*ELB)]
HMOD=b0+b1*exp(-K1*t)+b2*exp(-K2*t)
K1=exp[-(k1+E1*X)]
K2=exp[-(k2+E2*X)]
ELB=a0+a1*exp(-K1*t)+a2*exp(-K2*t)
X=1000/(T+273)-Xm
Xm="Xm"
a0=?
a1=?
a2=?
b0=?
b1=?
b2=?
b3=?
b4=?
k1=?
E1=?
k2=?
E2=?
```

Рис. 8.2 Модель разрывного напряжения в системе Fitter

Функции (8.3)- (8.6) зависят от двух общих (кинетических) параметров – K_1 и K_2 и от нескольких частных (формальных) параметров. В свою очередь, общие параметры зависят от температуры по закону Аррениуса, который мы запишем в следующем виде

$$K_i = e^{-k_i - E_i X}, \quad i = 1, 2$$

где

$$X = \frac{1000}{T + 273} - X_0, \quad X_0 = \frac{1}{3} \sum_{i=1}^3 \left(\frac{1000}{T_i + 273} \right) \quad T_1 = 140, T_2 = 125, T_3 = 110.$$

Такая форма представления закона Аррениуса очень удобна для проведения вычислений. Если сравнить ее с традиционной формой записи,

$$K = g \exp\left(-\frac{F}{R(T + 273)}\right), \quad F = 1000RE, \quad g = e^{-k + EX_0},$$

то видно, что величины E (энергия активации) и k (предэкспонента) значительно ближе к единице, чем «натуральные» параметры g и F . (Смотри также Табл. 8.3). Эта замена переменных позволяет значительно улучшить структуру модели, уменьшить внутренние корреляции и, тем самым, помогает справиться с проблемой нестрогой мультиколлинеарности. Подробно этот прием описан в разделе 4.4.

Таким образом, для описания экспериментальных данных используются четыре модели с восемью частными и четырьмя общими параметрами. Оценка всех этих параметров проводится методом последовательного байесовского оценивания.

8.3. Обработка данных УИ

В рассматриваемом примере данные обрабатывались в следующей последовательности:

$$ELB \xrightarrow{1} STR \xrightarrow{2} ELB \xrightarrow{3} STR \quad (8.7)$$

Вся процедура начинается с разрывного удлинения, причем на первом шаге использовалась упрощенная модель с одним кинетическим параметром

$$ELB(t, T) = a_0 + a_1 e^{-K_1 t},$$

т.к. оценить полную модель невозможно – слишком мало данных. По итогам этой обработки строится апостериорная информация (2.21)-(2.24), которая затем пересчитывается в априорную информацию (2.33)-(2.37). Эта информация используется для подгонки деформационных кривых на втором шагу в схеме (8.7). Параметры k_i и E_i являются общими, а a_0 и a_1 – частными. Очевидно, что ошибка измерения разрывного удлинения отличается от ошибки измерения напряжения, поэтому на этом этапе мы используем априорную информацию 2-го рода, так, как это объяснено в разделе 2.1.

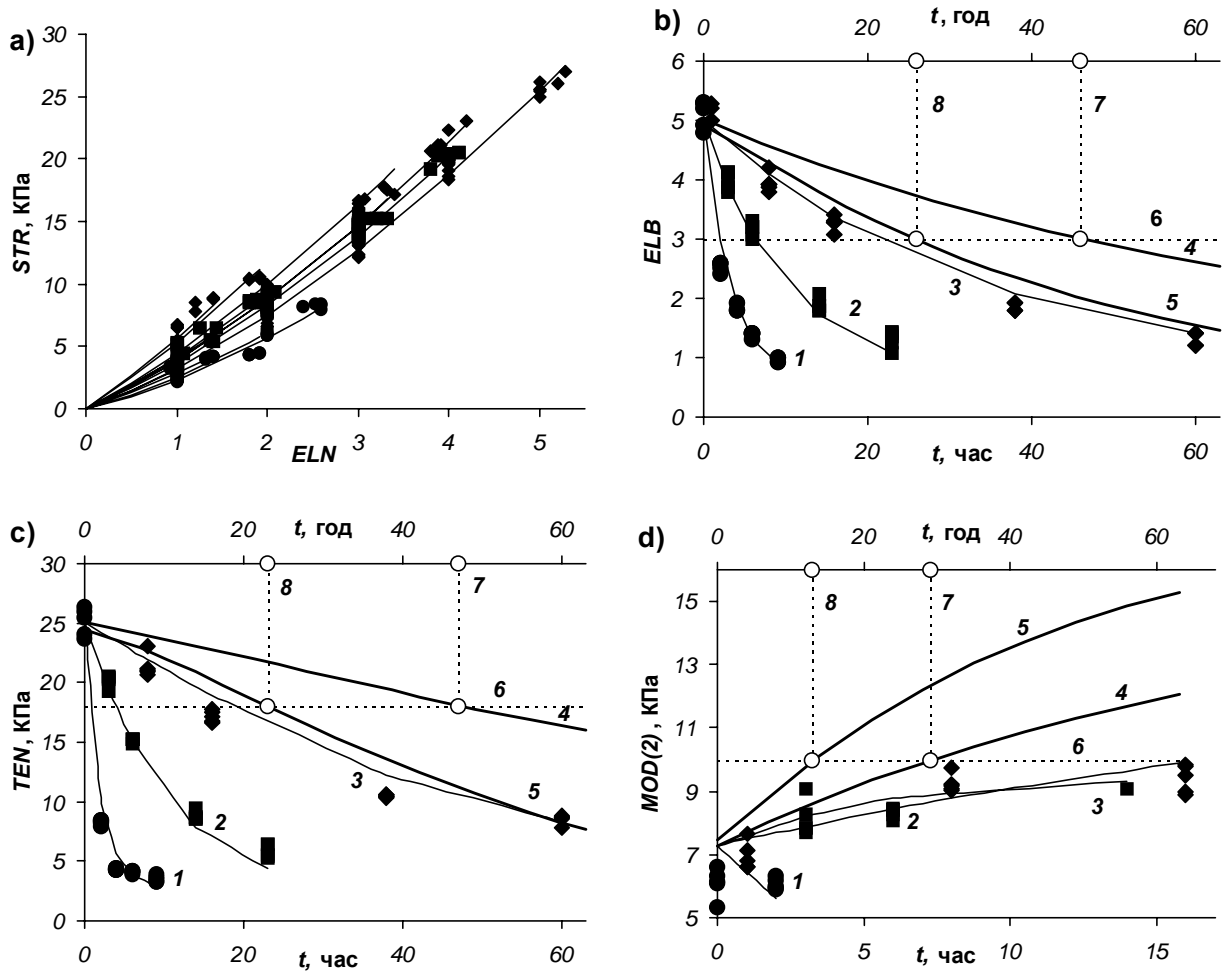


Рис. 8.3 Данные ускоренных испытаний и результаты экстраполяции показателей на температуру $T=20^{\circ}\text{C}$ а) деформационные кривые, б) разрывное удлинение, в) разрывная прочность д) условный модуль. 1 (●) – $T=140^{\circ}\text{C}$, 2 (■) – $T=125^{\circ}\text{C}$, 3 (◆) – $T=110^{\circ}\text{C}$, 4 – среднее значение, 5 – 0.95 доверительное значение, 6 – критический уровень, 7 – средний срок, 8 – доверительный срок. Кривые 5, 6 и значения 7, 8 на рисунках б) - д) относятся к верхней оси абсцисс.

Из деформационных кривых уже удастся найти оценки всех четырех общих параметров и передать их как априорную информацию дальше. Соответствующие матрицы – априорная и апостериорная показаны на Рис. 8.4.

После этого этапа мы вновь возвращаемся к данным разрывного удлинения ELB (третий шаг схемы (8.7)) с тем, чтобы уточнить параметры a_0 , a_1 , k_2 , E_2 и оценить параметр a_2 . При этом параметры k_1 и E_1 исключаются из процедуры оценивания, т.е. сохраняют свои значения, найденные на предыдущем шаге в соответствии с методом обратного последовательного байесовского оценивания (ОПБО), описанным в разделе 2.4.

| Prior Information after ELB for STR | | | | | | | | | | | |
|-------------------------------------|-------|--------|---|---|---|---|--------|--------|---|---|---------|
| Name | Value | Matrix | | | | | | | | | Exclude |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| k1 | 1.402 | 0 | 0 | 0 | 0 | 0 | 386.68 | 27.167 | 0 | 0 | |
| E1 | 12.53 | 0 | 0 | 0 | 0 | 0 | 27.167 | 15.053 | 0 | 0 | |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |

| Posterior Information after ELB+STR for ELB | | | | | | | | | | | |
|---|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| Name | Value | Matrix | | | | | | | | | Exclude |
| b0 | 11.56 | 3007.5 | 2162 | 2487.2 | 341.04 | -1E+05 | -2452 | -194.4 | 1276.6 | 62.36 | |
| b1 | -4.22 | 2162 | 1714.9 | 1921.7 | 240.67 | -82409 | -1460 | -129.3 | 685.17 | 36.247 | |
| b2 | 3.95 | 2487.2 | 1921.7 | 2234.3 | 278.2 | -95803 | -1797 | -180.7 | 801.34 | 55.857 | |
| b3 | -62.2 | 341.04 | 240.67 | 278.2 | 38.908 | -13717 | -286.5 | -22.33 | 151.18 | 7.261 | |
| b4 | 0.141 | -1E+05 | -82409 | -95803 | -13717 | 5E+06 | 103699 | 7926.9 | -55517 | -2618 | |
| k1 | 1.655 | -2452 | -1460 | -1797 | -286.5 | 103699 | 2974.8 | 210.73 | -1568 | -68.45 | |
| E1 | 12.32 | -194.4 | -129.3 | -180.7 | -22.33 | 7926.9 | 210.73 | 34.78 | -68.45 | -8.83 | |
| k2 | 1.512 | 1276.6 | 685.17 | 801.34 | 151.18 | -55517 | -1568 | -68.45 | 1132.7 | 32.383 | |
| E2 | 25.78 | 62.36 | 36.247 | 55.857 | 7.261 | -2618 | -68.45 | -8.83 | 32.383 | 4.1908 | |

Рис. 8.4 Пример априорной и апостериорной информационных матриц

По итогам этого шага окончательно определяются значения общих (кинетических) параметров и остается только уточнить значения частных параметров b_0, \dots, b_4 . Для этого на четвертом шагу вновь подгоняются деформационные кривые, но уже с неизменными значениями общих параметров. Здесь также применяется процедура ОПБО.

При такой технике оценивания, когда одни и те же данные используются несколько раз, необходимо следить за тем, чтобы второй и более раз пересчитывались только частные параметры, а общие параметры оставались неизменными. В Табл. 8.3 приведена схема такого оценивания. В ней итоговые значения параметров выделены жирным шрифтом. Так, например, видно, что пара кинетических параметров k_1 и E_1 окончательно определяется на втором шаге процедуры (**STR-1**) и далее уже не меняется, а частные параметры a_0, a_1, a_2 окончательно оцениваются на третьем шаге (**ELB-2**). Кроме того, в последних двух колонках даны значения «физических», традиционных значений предэкспонент (час^{-1}) и энергий активации (КДж/моль) кинетических параметров.

Таким образом, мы определяем все значения неизвестных параметров в моделях (8.3)–(8.6) и можем выполнить прогнозирование, т.е. экстраполяцию на условия эксплуатации.

Табл. 8.3 Схема последовательного оценивания параметров

| | <i>ELB-1</i> | <i>STR-1</i> | <i>ELB-2</i> | <i>STR-2</i> | Физические величины |
|----------------|-----------------------|--------------|--------------|--------------|--|
| частные | <i>a</i> ₀ | 0.97 | | 0.64 | |
| | <i>a</i> ₁ | 4.19 | | 4.31 | |
| | <i>a</i> ₂ | | | 0.06 | |
| | <i>b</i> ₀ | | 11.56 | | 11.41 |
| | <i>b</i> ₁ | | -4.22 | | -4.14 |
| | <i>b</i> ₂ | | 3.95 | | 3.87 |
| | <i>b</i> ₃ | | -62.20 | | -59.71 |
| | <i>b</i> ₄ | | 0.14 | | 0.14 |
| общие | <i>k</i> ₁ | 1.40 | 1.65 | | $2.70 \times 10^{+12}$ час ⁻¹ <i>g</i> ₁ |
| | <i>E</i> ₁ | 12.53 | 12.32 | | 102.32 КДж/моль <i>F</i> ₁ |
| | <i>k</i> ₂ | | 1.51 | 1.50 | $1.96 \times 10^{+27}$ час ⁻¹ <i>g</i> ₂ |
| | <i>E</i> ₂ | | 25.78 | 26.16 | 217.44 КДж/моль <i>F</i> ₂ |

Подробности всех расчетов можно найти в [126], где приведен файл *RUBBER.XLS*, содержащий все исходные данные и порядок расчетов.

8.4. Прогнозирование

Прогнозирование проводилось экстраполяцией моделей на температуру $T=20C$ при заданных критических уровнях для всех показателей свойств (См. Табл. 8.4). Величины этих уровней (вторая строка Табл. 8.4) были определены разработчиками материала – компанией Uniroyal Chemical. Для каждого показателя определялись два срока – средний и доверительный. Первый вычислялся как время достижения критического уровня для модели, взятой при средних значениях параметров (6-я строка Табл. 8.4). Вторым сроком (5-я строка) определялся с учетом неопределенности в оценках параметров. Для этого в каждой модели вычислялся односторонний 0.95 доверительный интервал (для условных модулей $MOD(n)$ – верхний, а для остальных – нижний) и срок устанавливался по пересечению этого интервала с критическим значением.

Табл. 8.4 Результаты экстраполяции моделей на температуру $T=20C$

| Показатель свойств | <i>ELB</i> | <i>TEN</i> | <i>MOD(1)</i> | <i>MOD(2)</i> | <i>MOD(3)</i> | <i>MOD(4)</i> | <i>MOD(5)</i> |
|--------------------------|------------|------------|---------------|---------------|---------------|---------------|---------------|
| Начальное значение | 5.01 | 25.05 | 3.10 | 7.29 | 12.40 | 18.33 | 24.95 |
| Предельное значение | 0.71 | 2.04 | 7.25 | 15.57 | 24.83 | 34.89 | 45.66 |
| Критическое значение | 3.00 | 18.00 | 5.00 | 10.00 | 17.00 | 25.00 | 32.00 |
| Изменение показателя | 47% | 31% | 46% | 33% | 37% | 40% | 34% |
| Доверительный срок (год) | 25.00 | 23.00 | 20.00 | 14.00 | 16.00 | 18.00 | 15.00 |
| Средний срок (год) | 46.00 | 47.00 | 45.00 | 29.00 | 34.00 | 38.00 | 31.00 |

Соответствующие данные приведены также и на Рис. 8.3 **b)-d)**, где показаны средние значения показателей (кривые **4**), доверительные границы (кривые **5**), критический уровень (прямая **6**), а также соответствующие сроки – точки **7** и **8**. Разумеется все доверительные сроки меньше средних. Кроме того, в Табл. 8.4 представлены начальные и предельные значения всех показателей (первая и вторая строки). Начальное значение равно величине соответствующего показателя при $t=0$

$$y_0 = y(T=20C, t=0).$$

Предельное значение всех показателей вычисляется в одной и той же точке –

$$y_{lim} = y(T=20C, t=1000 \text{ лет}).$$

Выбор этой точки в качестве предельной объясняется следующими соображениями. Величина разрывного удлинения *ELB* монотонно падает и ее предельное значение равно значению на бесконечности $ELB(\infty) \approx ELB(1000 \text{ лет})$. Величины условных модулей *MOD(n)* достигают максимума, а показатель разрывной прочности *TEN(t)* имеет перегиб в этой точке. В четвертой строке Табл. 8.4 представлены величины изменения всех показателей, которые вычисляются как

$$\left| \frac{y_0 - y_{crit}}{y_{lim} - y_{crit}} \right|.$$

Интересно сопоставить величины сроков службы (строки 5 и 6) с соответствующими изменениями показателя (строка 4). Видно, что более длительные сроки службы (*ELB* и *TEN*) связаны с высокими показателями изменения свойств. Возникает предположение о том, что рассогласование в оценках сроков службы, полученных для разных показателей,

может быть вызвано неверным заданием критических уровней. Для ее проверки мы построили графики изменения показателей свойств (коэффициенты старения) в зависимости от времени.

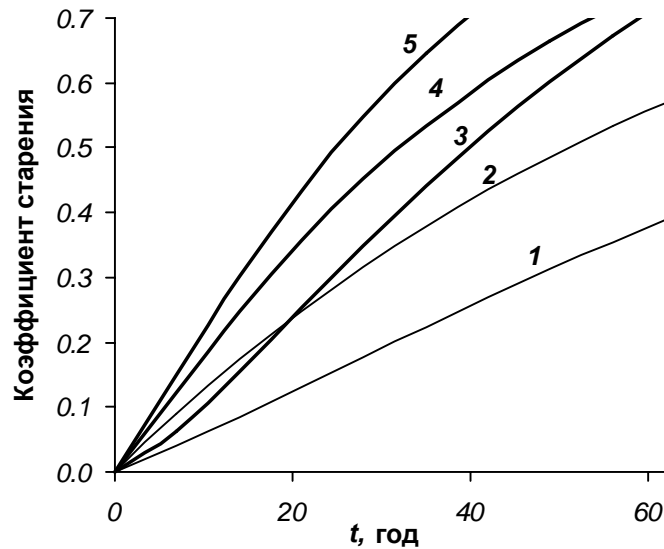


Рис. 8.5 Коэффициенты старения для различных показателей свойств и их доверительных значений: **1** – разрывная прочность в среднем, **2** – разрывное удлинение и все модули (в среднем), **3** – разрывная прочность доверительно, **4** – разрывное удлинение доверительно, **5** – все условные модули доверительно.

На Рис. 8.5 показаны кривые изменения коэффициентов старения

$$\text{коэффициент старения} = \frac{y(t) - y_{crit}}{y_{lim} - y_{crit}}$$

для всех показателей свойств (как средних, так и их доверительных значений), вычисленных для практически важного отрезка времени. Все эти кривые выходят из нуля и стремятся к единице на бесконечности. Из графика видно, что на представленном отрезке времени коэффициенты старения всех модулей совпадают. Это обстоятельство, очевидно, вытекает из уравнений для этих показателей. Гораздо интереснее совпадение среднего коэффициента старения по разрывному удлинению с коэффициентами всех средних модулей (кривая **2**). Этот факт не следует из вида моделей, а является особенностью данной системы. Коэффициент старения для показателя разрывной прочности (кривые **1** и **3**) идет ниже всех других коэффициентов. Это свидетельствует о том, что этот показатель слабее других изменяется при старении.

8.5. Результаты главы 8

Использование метода последовательного оценивания параметров при прогнозировании эксплуатационной устойчивости дает несомненные преимущества. Он позволяет строить

сложные, внутренне согласованные модели, содержащие общие параметры. Мы полагаем, что сам факт существования такой модели, ее адекватность эксперименту является доказательством возможности и правильности экстраполяции. Вот ход наших рассуждений.

Центральным моментом при построении модели термоокислительного старения является вопрос о количестве общих кинетических параметров, зависящих от температуры. Почему в рассматриваемом примере мы взяли только два таких параметра? Можно показать, что наличие параметра K_2 необходимо в модели деформационной кривой (8.3) и излишне в модели разрывного удлинения (8.4). В тоже время, введение третьего параметра K_3 не улучшает описание кривых STR/ELN , а лишь создает переопределенность – строгую мультиколлинеарность, о которой шла речь во введении. Хорошо известно [127], что избыточное описание данных неустойчиво и ведет к ошибкам в прогнозе. «Лишний» параметр K_2 в модели разрывного удлинения не противоречит этому принципу, т.к. он хорошо определяется из деформационной кривой. По сути дела, набор уравнений (8.3)–(8.6) является единой моделью многомерного отклика, и оценки общих параметров принадлежат всем этим откликам. Кстати, совпадение коэффициентов старения для всех условных модулей с коэффициентом старения ELB связано как раз с несущественностью параметра K_2 при прогнозе этих показателей. В тоже время при правильной обработке кривой разрывного удлинения необходимо все же учитывать наличие в полной модели параметра K_2 . Если этого не сделать, то оценки параметров k_I и E_I останутся равными значениям, полученным на первом шаге последовательного оценивания (первый столбец в Табл. 8.3). Тогда при прогнозе мы получим завышенные сроки: средний 52 года, доверительный – 36 лет.

Второй момент, который нужно учитывать при построении моделей старения – это их форма и зависимость от частных параметров. Нам представляется, что этот вопрос значительно менее важен, чем выбор числа общих параметров. Если ускоренные испытания были спланированы правильно, так, чтобы все показатели свойств перешли за свои критические значения, то экстраполяцию по времени t в моделях делать не нужно. А раз так, то их форма, т.е. зависимость от t не существенна. Например, в уравнении для разрывного удлинения (8.4) можно вместо экспоненты взять гиперболу, но это практически ничего не изменит в результатах прогноза – доверительный срок 25 изменится на 23, а средний 46 на 43. Правильный план эксперимента проще всего выстроить эволюционно – последовательно уточняя время следующего съема по уже имеющимся данным. Это можно сделать как с помощью программы Fitter, так и с помощью процедуры ЭПУИ [125], специально предназначенной для этой цели.

Таким образом, мы можем заключить, что число общих кинетических параметров в каждой задаче можно определить, исходя из очевидных соображений адекватности и идентифицируемости. Эти величины имеют физический смысл – они отражают объективные процессы, происходящие в материале при термоокислительной деструкции. Частные параметры, характеризующие форму кривых изменения свойств, физического смысла не имеют и могут сильно меняться в зависимости от выбранного математического описания, что, однако, не существенно для результатов прогноза.

Оценивая численные результаты прогноза, полученные в рассмотренном примере, мы можем заключить, что они, по-видимому, близки к эмпирическим оценкам сроков хранения. Разногласие в оценках коэффициента старения по разным показателям (см. Рис. 8.5), очевидно, связано с тем, что единый процесс старения данного материала по разному влияет на различные показатели его свойств. Этого и следовало ожидать.

В заключение хочется подчеркнуть, что такой анализ модели старения можно было провести, только используя метод последовательного байесовского оценивания. Другие методы идентификации моделей не позволяют провести оценивания параметров многооткликового гетероскедастичного массива данных.

9. Моделирование диффузии

Диффузионные модели широко используются в химической физике [130]. Особое значение они имеют для исследования свойств и старения стеклообразных полимеров, каковыми являются все конструкционные термопласты [131]. Первой и необходимой стадией, обеспечивающей и в значительной мере определяющей все дальнейшие эффекты старения материала, является сорбция влаги материалом. При этом наблюдаются различные эффекты изменения, как свойств, так и структуры материала. Эти эффекты могут быть обратимы (набухание, заполнение пор) и необратимы (химические и физические превращения, гидролиз). С другой стороны, любые структурные изменения, происходящие в материале, обратимые или необратимые, оказывают влияние на кинетику сорбции. Таким образом, диффузия в полимерах заслуживает изучения с двух точек зрения. Во-первых, как явление, приводящее к старению, а во-вторых, как инструмент, позволяющий исследовать структуру и строение материала.

Моделирование диффузионных процессов заставляет строить довольно сложные описания [129]. Здесь мы сталкиваемся с ситуацией, когда модель, как правило, представляется бесконечным рядом. Коэффициенты этого ряда зависят от времени t и устроены таким образом, что ряд хорошо сходится только при больших значениях t . В некоторых случаях нужно удерживать в ряду сотни или даже тысячи членов, чтобы правильно вычислить значения модели вблизи начала координат. В тоже время, стоит несколько отойти от нуля, как для приемлемого описания достаточно лишь 4-5 членов ряда. Для решения этой проблемы предлагается вначале использовать иное – асимптотическое описание, которое хорошо приближает модель в нуле, а затем переходить на выражение в форме конечного ряда, хорошо работающее при больших t .

В этой главе будет показано, как такой подход используется для моделирования нормальных и аномальных диффузионных процессов, а также кинетики цикла «увлажнение–сушка». Подробности можно найти в [126], где приведен файл *DIFFUS.XLS*, содержащий все модели и примеры их использования.

9.1. Нормальная диффузия

Любая модель, описывающая процесс диффузии, опирается на классическое уравнение теплопроводности. Для постоянного коэффициента диффузии D оно имеет вид [132]

$$\frac{\partial c}{\partial t} = D\Delta c \quad (9.1)$$

В этом уравнении величина $c(x, t)$ – это концентрации сорбента в точке x в момент времени t . Помимо уравнения (9.1) модель диффузии должна также включать начальное

$$c(x, 0) = \varphi(x) \quad (9.2)$$

и граничное

$$c(x, t) \Big|_{x \in \partial B} = \mu(t) \quad (9.3)$$

условия. В уравнениях (9.2) и (9.3) ∂B – это граница тела B , $\varphi(x)$ и $\mu(t)$ – известные функции.

Как правило, в эксперименте наблюдается величина

$$M(t) = \int_B c(x, t) dx \quad (9.4)$$

определяющая изменение массы тела B к моменту времени t .

Во многих случаях влиянием объемных факторов при диффузии можно пренебречь и свести задачу к одномерной. Тогда уравнение (9.1) превращается в

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} \quad (9.5)$$

Рассмотрим случай, когда тело B является достаточно большой плоскопараллельной пластиной толщиной l . Тогда мы можем пренебречь диффузией через узкие торцы и учитывать только одномерный поток через две большие грани. Обычными условиями для такой системы являются следующие уравнения

$$\begin{aligned} c(x, 0) &= \varphi(x) \\ c(0, t) &= \mu_1(t) \quad c(l, t) = \mu_2(t) \end{aligned} \quad (9.6)$$

Решение уравнения (9.5) с условиями (9.6) имеет вид

$$c(x, t) = u(x, t) - \int_0^t \int_0^l G_l(x, \xi, t - \tau) u'_l(\xi, \tau) d\xi d\tau + \int_0^l G_l(x, \xi, t) [\varphi(\xi) - u(\xi, 0)] d\xi \quad (9.7)$$

где

$$G_l(x, \xi, t) = \frac{2}{l} \sum_{n=1}^{\infty} e^{-\left(\frac{\pi n}{l}\right)^2 Dt} \sin \frac{\pi n}{l} x \cdot \sin \frac{\pi n}{l} \xi \quad (9.8)$$

- это функция источника для диффузии в отрезок, а

$$u(x,t) = \mu_1(t) - \frac{x}{l} [\mu_2(t) - \mu_1(t)], \quad u'_t(x,t) = \frac{\partial u(x,t)}{\partial t} \quad (9.9)$$

- это вспомогательная функция.

В некоторых случаях влиянием второго конца отрезка можно пренебречь и рассматривать отрезок как полуось $[0, +\infty]$. Формально это можно рассматривать как предельный переход в уравнении (9.7) при $\frac{Dt}{l^2} \rightarrow 0$, однако удобнее решить уравнение (9.5) с новыми ус-

ловиями

$$\begin{aligned} c(x, 0) &= c_0 \\ c(0, t) &= \mu(t) \end{aligned} \quad (9.10)$$

Решение имеет вид

$$c(x,t) = c_0 \int_0^\infty [G_2(x, \xi, t) - G_2(x, -\xi, t)] d\xi + 2D \int_0^t \frac{\partial G_2}{\partial \xi}(x, 0, t - \tau) \mu(\tau) d\tau \quad (9.11)$$

где

$$G_2(x, \xi, t) = \frac{l}{2\sqrt{\pi Dt}} e^{-\frac{(x-\xi)^2}{4Dt}} \quad (9.12)$$

- это функция источника на оси. Первое слагаемое в формуле (9.11) можно представить в виде

$$c_1(x,t) = \frac{c_0}{2} \left[l + \operatorname{erf} \left(\frac{x}{2\sqrt{Dt}} \right) \right] \quad (9.13)$$

где

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt \quad (9.14)$$

- это интеграл ошибок. Для постоянного граничного условия (9.10) второе слагаемое в формуле (9.11) можно представить как

$$c_2(x,t) = \mu_0 \left[l - \operatorname{erf} \left(\frac{x}{2\sqrt{Dt}} \right) \right] \quad (9.15)$$

Если величина начальной концентрации c_0 отлична от нуля, то исходная масса сорбента в образце $M(0)$ формально равна бесконечности, однако величина $M(t) - M(0)$ конечна и равна

$$m(t) = M(t) - M(0) = (2\mu_0 - c_0) \sqrt{\frac{Dt}{\pi}} \quad (9.16)$$

Рассмотрим кинетику сорбции в отрезок (9.4) для простейшего варианта условий

$$\begin{aligned} c(x, 0) &= c_0 \\ c(0, t) &= c(l, t) = \mu_0 \end{aligned} \quad (9.17)$$

Тогда уравнение (9.9) принимает вид

$$u(x, t) = \mu_0, \quad u'_t(x, t) = 0 \quad (9.18)$$

а уравнение (9.7)

$$c(x, t) = \mu_0 + (c_0 - \mu_0) \int_0^l G_l(x, \xi, t) d\xi \quad (9.19)$$

Используя уравнение (9.8) для функции источника G_l получаем

$$c(x, t) = \mu_0 + (c_0 - \mu_0) \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} e^{-\left[\frac{(2k+1)\pi}{l}\right]^2 Dt} \sin \frac{(2k+1)\pi}{l} x \quad (9.20)$$

Очевидно, что для условий в виде (9.17) имеет место

$$M(0) = lc_0 \quad M(+\infty) = l\mu_0 \quad (9.21)$$

Интегрируя выражение (9.20) по x , получаем формулу для кинетики сорбции

$$M(t) = lc_0 + (l\mu_0 - lc_0) \left\{ 1 - 8 \sum_{k=0}^{\infty} \frac{1}{n_k^2} e^{-n_k^2 \frac{D}{l^2} t} \right\} \quad (9.22)$$

где

$$n_k = \pi(2k+1) \quad (9.23)$$

Известно [132], что ряд в (9.22) хорошо сходится только для достаточно больших t . Для расчета кинетики сорбции в начале процесса необходимо удерживать в уравнении большое количество членов. Поэтому, приведем другую формулу для расчета величины $M(t)$, удобную для малых t . Очевидно, что в начале процесса диффузии можно пренебречь взаимным влиянием противоположных граней и воспользоваться результатом (9.16), полученным для полубесконечной оси. С учетом того, что имеется две поверхности, через которые идет диффузия, получаем

$$M(t) = M(0) + 2(2\mu_0 - c_0) \sqrt{\frac{Dt}{\pi}} \quad (9.24)$$

Рассмотрим, когда можно использовать полученную асимптотику. Обозначим

$$d = \frac{D}{l^2}, \quad M_0 = l\mu_0, \quad C_0 = lc_0 \quad (9.25)$$

тогда уравнения (9.22) и (9.24) можно записать в виде удобном для вычислений

$$M(t) = \begin{cases} C_0 + 2(2M_0 - C_0)\sqrt{\frac{dt}{\pi}}, & t < \tau \\ M_0 + 8(C_0 - M_0) \sum_{k=0}^{\infty} \frac{e^{-n_k^2 dt}}{n_k^2}, & t \geq \tau \end{cases} \quad (9.26)$$

Очевидно, что на начальном участке $t < \tau$ асимптотика (9.24) будет меньше, чем решение (9.22) с конечным число членов в ряду.

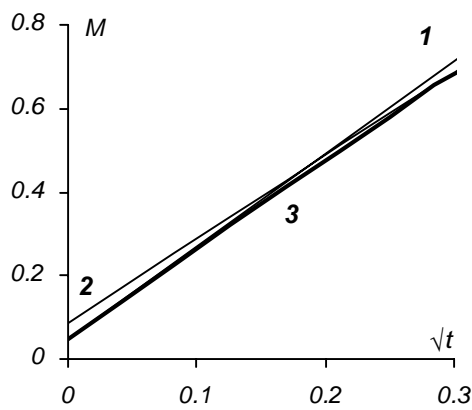


Рис. 9.1 Начальный участок кинетики нормальной диффузии

На Рис. 9.1 показаны кривые, построенные в координатах $(M(t), \sqrt{t})$ при следующих значениях параметров модели: $d=1, M_l=1, M_0=0.05$. Кривая 1 соответствует асимптотике (9.24), кривая 2 – уравнению (9.22) с четырьмя членами в ряду, а кривая 3 – это результирующее уравнение (9.26). Видно, что на начальном участке кривая 2 дает неверный, завышенный результат, а на конечном участке, наоборот, кривая 1 проходит выше кривой 2. Таким образом, момент времени τ в формуле (9.26) следует выбирать

из условия совпадения значений $M(t)$ в уравнениях (9.22) и (9.24). Заметим, что для этого не обязательно явно вычислять само значение τ .

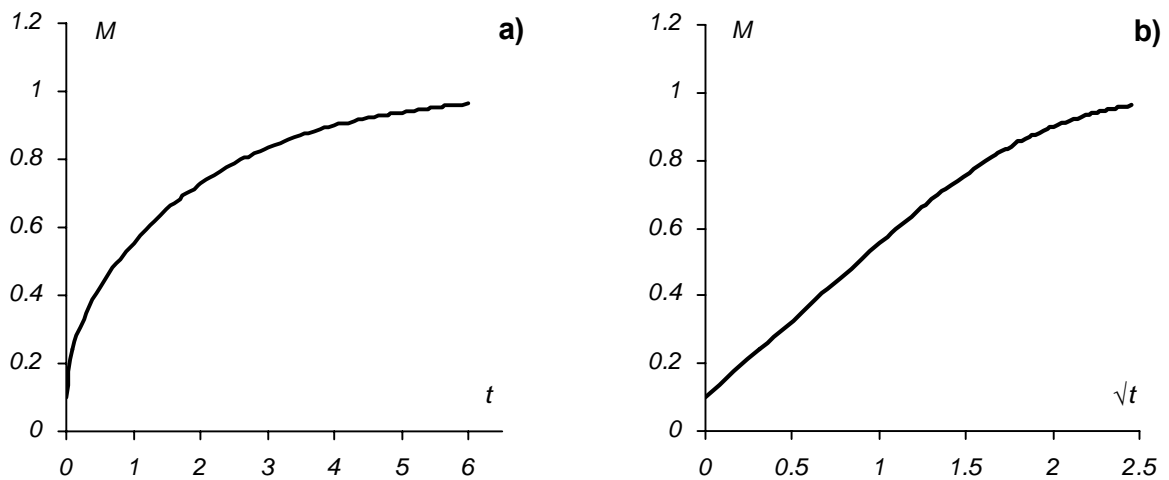


Рис. 9.2 Кинетические кривые нормальной диффузии в координатах (M, t) а) и (M, \sqrt{t}) б) для $d=0.05, M_0=1, C_0=0.1$.

Из уравнения (9.26) следует, что кинетические кривые в координатах $\left(\frac{M(t)}{M(\infty)}, \sqrt{t/l} \right)$

1. монотонны
2. не имеют точек перегиба
3. не зависят от толщины образца l

Эти формальные признаки характеризуют *нормальную* диффузию [130], а при их нарушении диффузия называется *аномальной*. Типичные кинетические кривые нормальной диффузии показаны на Рис. 9.2.

Рассмотрим, как моделируется нормальная диффузия с помощью программы Fitter. На Рис. 9.3 показан текстовое поле с моделью, используемой для расчетов кинетики (9.26).

```

'Normal model
M=m1*hev(USEm1)+m2*[hev(-USEm1)+imp(-USEm1)]
'condition to apply asimptotic case m1
  USEm1=m2-m1
'constants and intermediates
  P2=PI*PI
  P12=(PI)^(-0.5)
  x=d*t
'asymptotic case at t<tau
  m1=C0+2*P12*(2*M0-C0)*(x)^0.5
'series case at t>tau
  m2=M0+(C0-M0)*8*(S0+S1+S2+S3+S4)
  n0=P2
  S0=exp[-n0*x]/n0
  n1=P2*9
  S1=exp[-n1*x]/n1
  n2=P2*25
  S2=exp[-n2*x]/n2
  n3=P2*49
  S3=exp[-n3*x]/n3
  n4=P2*81
  S4=exp[-n4*x]/n4
'unknown parameters
  d=?
  M0=?
  C0=?
    
```

Рис. 9.3 Модель нормальной диффузии в виде, используемом в программе Fitter.

Разберем ее устройство подробно. Первая строка (а также, третья, пятая, девятая, одиннадцатая и двадцать третья) – это комментарий. Собственно модель начинается со второй строки. В ней использованы стандартные функции Fitter: $hev(x)$ и $imp(x)$. Они определяются следующим образом (см. раздел 5.3)

$$hev(x) = \begin{cases} 0, & x \geq 0 \\ 1, & x < 0 \end{cases} \quad imp(x) = \begin{cases} 0, & x \neq 0 \\ 1, & x = 0 \end{cases}$$

Видно, что при $USE_{m1} > 0$ величина $M=m_1$, а при $USE_{m1} \leq 0$ величина $M=m_2$, что соответствует уравнению (9.26). Величина USE_{m1} вычисляется в четвертой строке, которая задает условие применения асимптотики. Переменные $p_2 = \pi^2$ и $p_{12} = \pi^{-0.5}$ – это константы. В десятой строке записано уравнение для асимптотики при малых временах (m_1), а в двенадцатой строке – формула для кинетики при больших временах (m_2). В сумме оставлены только четыре члена, что достаточно для удовлетворительной оценки. Заключительные строки модели содержат указание на неизвестные параметры, подлежащие оценке.

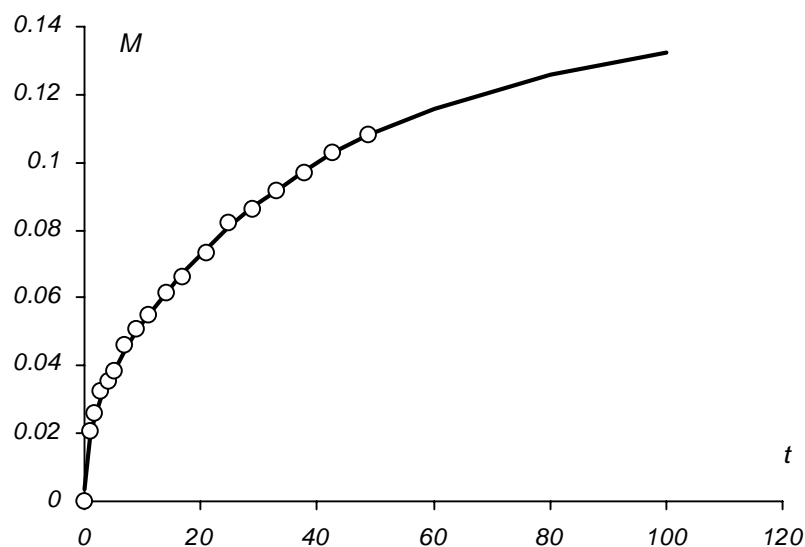


Рис. 9.4 Увлажнение органопластика Органит 7Т (условия выдержки: температура 60С, относительная влажность 100%), t - время в сутках, M - прирост массы.

На Рис. 9.4 показана кривая прироста массы органопластика Органит 7Т [134], которая хорошо описывается моделью нормальной фиковской сорбции (Рис. 9.3).

9.2. Аномальная диффузия

В реальных системах кинетика сорбции редко идет нормально. Некоторые примеры аномальных сорбционных кривых показаны на Рис. 9.5. Здесь представлена кинетическая кривая нормальной сорбции (**0**), и 4 типичных кривых аномальной сорбции (**1-4**).

Кривые 1-3. Это так называемая псевдонормальная кинетика сорбции (Рис. 9.5 **1** и **3**). Для нее характерна фиковская кинетика на начальном участке процесса и очень быстрая (**1**) или очень медленная (**3**) конечная стадия установления равновесия. Для ее объяснения обычно предполагают либо концентрационную зависимость коэффициентов диффузии [129], либо сорбент рассматривается как двухфазная система, состоящая из дисперсной фазы и дисперсионной среды, которые обладают различными парциальными коэффициентами диффузии [133].

Кривая 2. Это, так называемый, экстремальный тип аномальной кинетики сорбции (Рис. 9.5 2). Как правило, этот тип кинетики возникает в системах, сорбент которых первоначально находился в неравновесном состоянии, а после его пластификации молекулами сорбата система структурно релаксирует, постепенно приближаясь к новому термодинамически более устойчивому состоянию. При этом избыток сорбата, заполнивший сорбент на предыдущих стадиях процесса, может “выталкиваться” из образца [135]. Однако возможны и иные причины экстремальной кинетики сорбции [138].

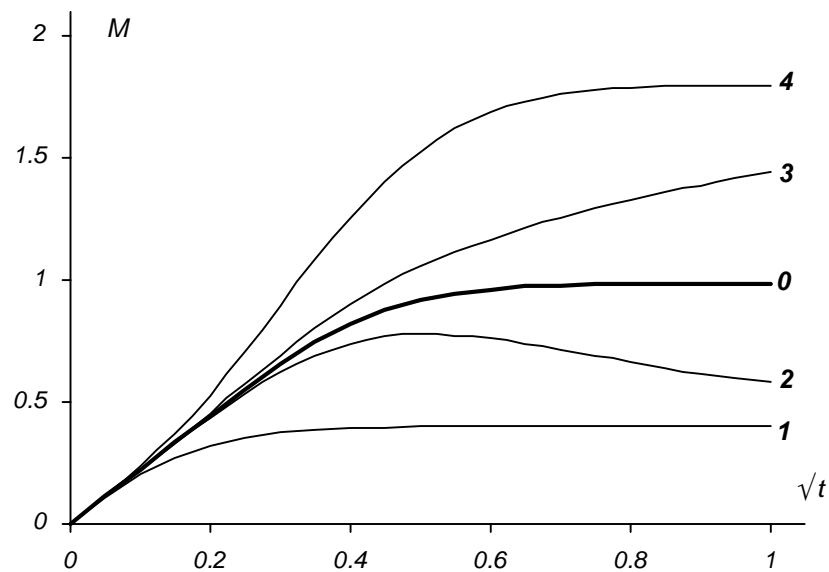


Рис. 9.5 Кинетические кривые нормальной (0) и аномальной диффузии (1-4).

Кривая 4. S-образная кинетика установления сорбционного равновесия (Рис. 9.5 4). Обычно этот тип кинетики наблюдается в стеклообразных сильно набухающих системах. С формальной точки зрения его связывают с релаксационными процессами установления граничной концентрации на поверхности.

Далее будут рассмотрены две наиболее сложные модели, используемые для описания аномальной диффузии — *релаксационная* и *конвекционная*.

9.3. Релаксационная модель аномальной диффузии

В релаксационной модели предполагаются переменные граничные условия вида [142]

$$\mu(t) = \mu_1 + (\mu_0 - \mu_1)e^{-rt}, \quad r \geq 0 \quad (9.27)$$

где

μ_0 — начальная растворимость

μ_1 — предельная растворимость

r — константа релаксации

Мы покажем, что эта модель позволяет описать эффекты, присущие аномальной диффузии. Физический смысл предлагаемого условия (9.27) заключается в том, что оно позволяет учесть влияние на кинетику сорбции одновременно протекающих процессов структурной релаксации. При этом переменная величина предельной сорбции выступает как характеристика состояния надмолекулярной структуры материала, которая отражает перестройки в структуре образцов, стимулированные поглощенным сорбентом. Эти изменения могут быть связаны, например, с образованием в образцах дополнительного свободного объема.

Мы будем рассматривать релаксационную диффузию в отрезок при следующих условиях

$$\begin{aligned} c(x, 0) &= c_0 \\ c(0, t) = c(l, t) &= \mu(t) \\ f(x, t) &= 0 \end{aligned} \quad (9.28)$$

Обозначим $C_0=lc_0$, $M_0=l\mu_0$, $M_1=l\mu_1$.

Тогда, используя соотношения (9.4) и (9.7), можно показать, что

$$M(t) = C_0 + (M_0 - C_0)Y_0(t) - r(M_0 - M_1) \int_0^t e^{-r(t-\tau)} Y_0(\tau) d\tau \quad (9.29)$$

где

$$Y_0(t) = 1 - 8 \sum_{k=0}^{\infty} \frac{e^{-n_k^2 dt}}{n_k^2} \quad (9.30)$$

– кинетика сорбции при постоянном граничном условии $\mu(t)=1$.

Интегрируя в (9.29) с учетом (9.30) получаем уравнение для кинетики релаксационной сорбции

$$\begin{aligned} M(t) &= M_1 + (M_0 - M_1) e^{-rt} - 8 \sum_{k=0}^{\infty} S_k \\ S_k &= \frac{(V_0 n_k^2 d - V_1 r) e^{-n_k^2 dt} + r(M_1 - M_0) e^{-rt}}{n_k^2 (n_k^2 d - r)} \\ V_0 &= M_0 - C_0; \quad V_1 = M_1 - C_0 \end{aligned} \quad (9.31)$$

Так же, как и уравнение (9.22), ряд (9.31) хорошо сходится только для больших значений t , поэтому построим его асимптотику для $dt \ll 1$. Используя формулу (9.26) при $C_0=0$, $M_0=1$

$$Y_0(t) \approx 4\pi^{-1/2} \sqrt{dt} \quad (9.32)$$

получим

$$M(t) \approx C_0 + 4\pi^{-1/2} \sqrt{dt} [M_0 - C_0 + (M_1 - M_0)\beta(rt)] \quad (9.33)$$

где функция $\beta(x)$ выражается через вырожденную гипергеометрическую функцию [143] $\Phi(a, b, x)$

$$\beta(x) = \frac{2}{3} x e^{-x} \Phi\left(\frac{3}{2}, \frac{5}{2}, x\right), \quad (9.34)$$

Однако удобнее использовать приближенное выражение

$$\beta(x) = 1 - e^{-z}, \quad z = \frac{a_1 x + a_2 x^2 + a_3 x^3}{1 + b_1 x + b_2 x^2 + b_3 x^3} \quad (9.35)$$

где

$$\begin{aligned} a_1 &= 0.66666448235932 & b_1 &= -0.0601632363536 \\ a_2 &= -0.0845402528610 & b_2 &= 0.0135016609125 \\ a_3 &= 0.0088582230275 & b_3 &= 0.0002360095898 \end{aligned}$$

Такое приближение дает относительную точность вычисления $\beta(x)$ не хуже чем указано в Табл. 9.1

Табл. 9.1 Точность вычисления функции $\beta(x)$.

| Область | Точность |
|-----------|-----------|
| $ x < 1$ | $1.1E-05$ |
| $ x < 2$ | $1.4E-05$ |
| $ x < 3$ | $2.0E-04$ |
| $ x < 4$ | $1.1E-03$ |

Окончательно, уравнение сорбции имеет вид

$$M(t) = \begin{cases} C_0 + 4\pi^{-1/2} \sqrt{dt} [M_0 - C_0 + (M_1 - M_0)\beta(rt)], & t < \tau \\ M_1 + (M_0 - M_1)e^{-rt} - 8 \sum_{k=0}^{\infty} S_k, & t \geq \tau \end{cases} \quad (9.36)$$

Здесь функция $\beta(x)$ определена соотношением (9.35), а величины S_k в уравнении (9.31). Здесь мы также будем использовать асимптотику (9.33) для вычисления кинетики сорбции до тех пор, пока она меньше, чем решение (9.31), которое пригодно для больших времен. Заметим, что в этом случае сложно определить момент времени τ , когда нарушается это условие.

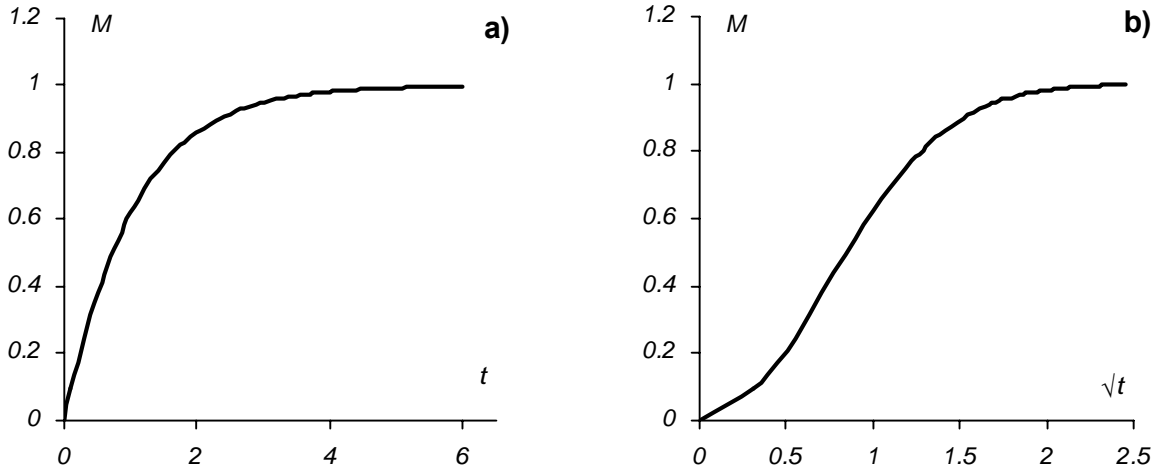


Рис. 9.6 Кривые релаксационной диффузии в координатах (M, t) **a)** и (M, \sqrt{t}) **b)** $d=0.07, M_0=0.5, M_I=1, r=1$.

На Рис. 9.6 показаны типичные кривые аномальной диффузии, в которых имеется точка перегиба.

Рассмотрим, описывает ли модель релаксационной диффузии зависимость от толщины образца. На Рис. 9.7 показаны кривые релаксационной сорбции в координатах

$\left(\frac{M}{M_I}, \frac{\sqrt{t}}{l} \right)$ для различных значений толщины образца $l=1, 2, 3, 4, 5$.

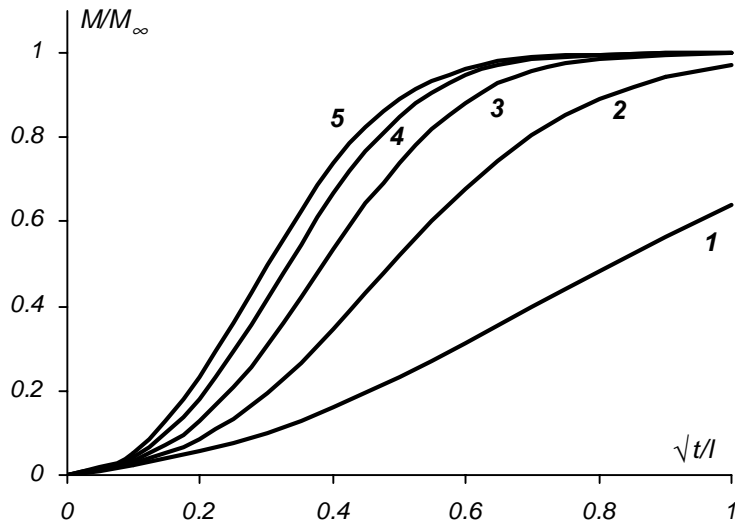


Рис. 9.7 Кинетические кривые релаксационной диффузии в координатах $(M/M_\infty, \sqrt{t}/l)$ при $D=0.05, \mu_0=0.1, \mu_l=l, r=1$ для различных значений толщины образца $l=1, 2, 3, 4, 5$.

Из рисунка видно, что для разных значений параметра l кривые отличаются, что характерно для аномальной диффузии.

В случае, когда предельная растворимость в уравнении (9.27) меньше чем начальная, т.е. $\mu_l < \mu_0$, то кривая сорбции будет, очевидно, проходить через максимум. Типичный случай

показан на Рис. 9.5 (кривая 2). Возвращаясь к этому рисунку, мы видим, что кривая 2 соответствует случаю медленного (по сравнению с характерным временем диффузии $\pi^{-2}d^{-1}$) падения граничного условия, кривая 1 – быстрому падению. Кривые 3 и 4 отражают случаи медленного и быстрого роста граничного условия соответственно.

Табл. 9.2 Параметры релаксационной модели.

| Кривая | d | M_0 | r | M_1 |
|--------|-----|-------|------|-------|
| 0 | 1.0 | 1.0 | 0.0 | 1.0 |
| 1 | 1.0 | 1.0 | 25.0 | 0.4 |
| 2 | 1.0 | 1.0 | 2.0 | 0.5 |
| 3 | 1.0 | 1.0 | 1.5 | 1.6 |
| 4 | 1.0 | 1.0 | 10.0 | 1.8 |

В Табл. 9.2 приведены значения параметров в релаксационной модели (9.36), которые описывают все эти кривые. Кривая 1 соответствует случаю быстрого (по сравнению с характерным временем диффузии) падения граничного условия (9.27), кривая 2 – медленному падению. Кривые 3 и 4 соответствуют случаю медленного и быстрого роста граничного условия. Таким образом, мы видим, что с помощью релаксационной модели можно количественно описать все особенности, присущие аномальной диффузии.

Релаксационная модель в виде, пригодном для использования в программе Fitter, представлена на Рис. 9.8. По своей структуре она подобна модели для нормальной диффузии, представленной на Рис. 9.3. Отметим некоторые ее особенности.

При $r=0$ модель переходит в нормальную модель, при этом параметр M_1 не определяется. При $r = n_k^2 d$, $k=0,1,2,3,4$ в модели возникает переполнение, связанное с неопределенностью типа $0/0$. Это обстоятельство может затруднить оценку параметров. Если заранее известно, что, например, $r = n_0^2 d = \pi^2 d$, то использовать модель в виде, показанном на Рис. 9.8, нельзя. Ее необходимо модифицировать, исключив один из неизвестных параметров d или r из всех уравнений. При этом выражение для S_0 следует записать по-другому – $S_0 = (M_0 - M_1) \cdot \exp(-x)$. Подробнее проблема разрешения неопределенностей в системе Fitter будет рассмотрена в следующем разделе.

Формально модель диффузии с релаксацией можно рассматривать и для случая $r < 0$. Точное решение (9.31) при этом сохранится, а асимптотическое решение в виде (9.33) будет достаточно хорошо описывать начальный участок кинетики вплоть до $rt < 4$ (см. Табл. 9.1)

Однако, при отрицательном значении коэффициента релаксации r предельное значение сорбции M_I будет, согласно уравнению (9.27), стремиться к бесконечности, что, по-видимому, лишено физического смысла.

```

' Relaxation model
M=m1*hev(USEm1)+m2*[hev(-USEm1)+imp(-USEm1)]
'condition to apply asymptotic case m1
  USEm1=m2-m1
'constants and intermediates
  P2=PI*PI
  P12=(PI)^(-0.5)
  x=r*t
  V0=M0-C0
  V1=M1-C0
'asymptotic case at t<tau
  m1=C0+4*P12*(d*t)^0.5*[M0-C0+(M1-M0)*beta]
  beta=1-exp(-z)
  z=(a1*x+a2*x*x+a3*x*x*x)/(1+b1*x+b2*x*x+b3*x*x*x)
  a1 = 0.6666644823593
  b1 = -0.0601632363536
  a2 = -0.0845402528610
  b2 = 0.0135016609125
  a3 = 0.0088582230275
  b3 = 0.0002360095898
'series case at t>tau
  m2=M1+(M0-M1)*exp(-x)-8*S
  R=r*(M1-M0)*exp(-x)
  S=S0+S1+S2+S3+S4
  n0=P2
  S0=[(V0*n0*d-V1*r)*exp(-n0*d*t)+R]/(n0*d-r)/n0
  n1=P2*9
  S1=[(V0*n1*d-V1*r)*exp(-n1*d*t)+R]/(n1*d-r)/n1
  n2=P2*25
  S2=[(V0*n2*d-V1*r)*exp(-n2*d*t)+R]/(n2*d-r)/n2
  n3=P2*49
  S3=[(V0*n3*d-V1*r)*exp(-n3*d*t)+R]/(n3*d-r)/n3
  n4=P2*81
  S4=[(V0*n4*d-V1*r)*exp(-n4*d*t)+R]/(n4*d-r)/n4
'unknown parameters
  d=?
  M0=?
  M1=?
  C0=?
  r=?

```

Рис. 9.8 Модель релаксационной диффузии в виде, используемом в программе Fitter.

Тем не менее, тот факт, что модель, представленная на Рис. 9.8, сохраняет непрерывность и точность даже при отрицательных значениях коэффициента r , имеет большое значение для правильного оценивания параметров модели. Ведь в ходе поиска процедура минимизации целевой функции может приводить к таким промежуточным величинам параметров

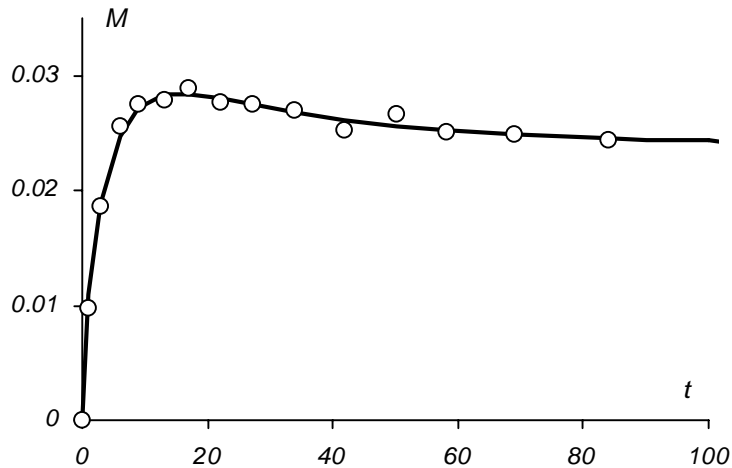


Рис. 9.9 Увлажнение стеклопластика авиационного назначения (условия выдержки: температура 60°C, относительная влажность 100%), t - время в сутках, M - прирост массы.

В заключение этого раздела продемонстрируем реальное практическое использование описанной модели. На Рис. 9.9 представлена кинетика сорбции воды стеклопластиком [136], которая хорошо описывается моделью, показанной на Рис. 9.8 при следующих значениях параметров

$$d=0.024 \quad M_0=0.032 \quad M_l=0.024 \quad C_0=0.0 \quad r=0.033$$

9.4. Конвекционная модель аномальной диффузии

Для учета конвекционных явлений в сильно набухающих стеклообразных полимерах в уравнение (9.5) вводят дополнительный член, ответственный за движение сорбента как целого

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} - v \left| \frac{\partial c}{\partial x} \right| \quad (9.37)$$

Мы будем решать уравнение (9.37) при условиях

$$\begin{aligned} c(x, 0) &= 0 \\ c(0) = c(l, t) &= \mu_0 \end{aligned} \quad (9.38)$$

Мы начнем исследование уравнения (9.37) с решения его стационарного варианта

$$\begin{aligned} D \frac{\partial^2 c_0}{\partial x^2} - v \left| \frac{\partial c_0}{\partial x} \right| &= 0 \\ c_0(0) = c_0(l) &= \mu_0 \end{aligned} \quad (9.39)$$

Легко видеть, что его решение имеет вид

$$c_0(x) = \mu_0 \frac{1 + \exp\left(\frac{v}{D}\left|x - \frac{l}{2}\right|\right)}{1 + \exp\left(\frac{vl}{2D}\right)} \quad (9.40)$$

После интегрирования выражения (9.40) по x на интервале $[0, l]$ получаем выражение для предельной равновесной сорбции

$$M_\infty = \frac{M_0}{1 + \exp(w)} \left[\frac{\exp(w) - 1}{w} + 1 \right] \quad (9.41)$$

здесь M_0 определено в уравнении (9.25), а величина w определяется как

$$w = \frac{vl}{2D} \quad (9.42)$$

Очевидно, что решение уравнения (9.37) с условиями в форме (9.38) должно стремиться к величине (9.41) при $t \rightarrow \infty$

$$\lim_{t \rightarrow \infty} M(t) = M_\infty$$

График равновесной конвекционной сорбции в координатах $(M_\infty/M_0, w)$ представлен на Рис. 9.10

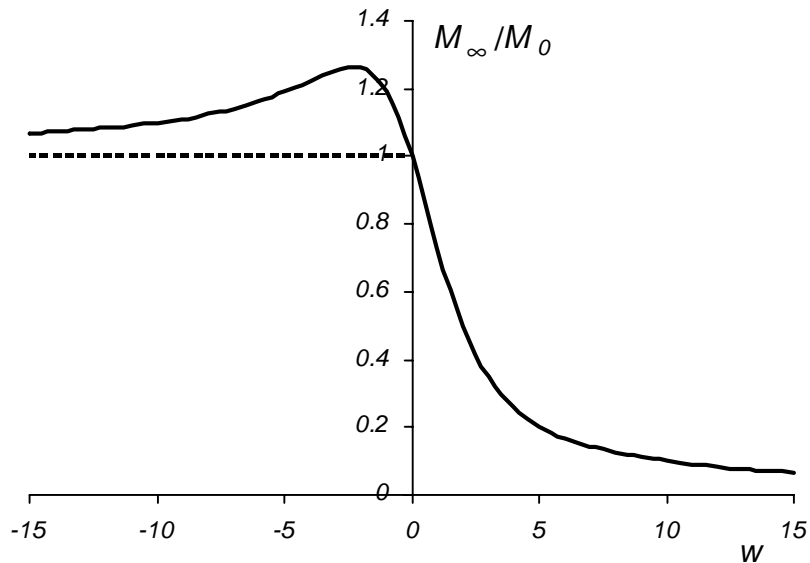


Рис. 9.10 Зависимость предельной равновесной сорбции в конвекционной модели от параметра модели.

Видно, что имеются два предельных случая. Если $w \rightarrow +\infty$, что соответствует большой положительной скорости конвекции, то равновесная сорбция стремится к нулю. В этом случае сорбент «выносятся» из материала быстрее, чем идет диффузия. Если же $w \rightarrow -\infty$, что соответствует большой отрицательной скорости конвекции, то равновесная сорбция стремится к M_0 . В этом случае конвекция помогает «заполнить» образец быстрее, чем идет

диффузия. В промежуточных случаях величина равновесной сорбции может быть как меньше, так и больше чем M_0 , что зависит от направления конвекции. Максимальная равновесная сорбция $M_\infty = 1.264M_0$ достигается при $w = -2.3$.

Вернемся теперь к исходному уравнению конвекционной модели сорбции (9.37). Будем искать его решение в виде

$$c(x, t) = c_0(x) - \exp(\gamma|x| - dw^2t) \rho(x, t) \quad (9.43)$$

где функция $c_0(x)$ – это равновесная концентрация, т.е. функция, представленная уравнением (9.40), величина w определена в уравнении (9.43), а параметр γ

$$\gamma = \frac{v}{2D} = \frac{w}{l} \quad (9.44)$$

Видно, что функция $\rho(x, t)$ удовлетворяет системе

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= D \frac{\partial^2 \rho}{\partial x^2} \\ \rho(x, 0) &= \rho_0(x) \\ \rho(0, t) &= \rho(l, t) = 0 \end{aligned} \quad (9.45)$$

Где

$$\rho_0(x) = \frac{2\mu_0}{1 + e^w} \operatorname{ch}\left(\gamma\left|x - \frac{l}{2}\right|\right) \quad (9.46)$$

Из уравнений (9.7) и (9.8) следует, что решение системы (9.45) имеет вид

$$\rho(x, t) = \frac{2}{l} \sum_{n=1}^{\infty} e^{-(\pi n)^2 dt} \sin \frac{\pi n}{l} x \cdot \int_0^l \rho_0(\xi) \sin \frac{\pi n}{l} \xi d\xi \quad (9.47)$$

Вычисляя интеграл в соотношении (9.47) легко видеть, что (также как и при нормальной сорбции) он отличен от нуля только при $n=2k+1$, и равен

$$\int_0^l \rho_0(\xi) \sin \frac{\pi n}{l} \xi d\xi = \frac{4l\mu_0 n_k \operatorname{ch}\left(\frac{w}{2}\right)}{(1 + e^w)(w^2 + n_k^2)} \quad (9.48)$$

поэтому,

$$\rho(x, t) = \frac{8\mu_0 \operatorname{ch}\left(\frac{w}{2}\right)}{1 + e^w} \sum_{k=0}^{\infty} \frac{n_k}{w^2 + n_k^2} e^{-n_k^2 dt} \sin\left(n_k \frac{x}{l}\right) \quad (9.49)$$

Теперь, используя уравнения (9.43), можно найти кинетику конвекционной сорбции, проинтегрировав выражение (9.49) по x на интервале $[0, l]$.

$$M(t) = M_{\infty} - e^{-w^2 t} \frac{\delta\mu_0 \operatorname{ch}\left(\frac{w}{2}\right)}{1 + e^w} \sum_{k=0}^{\infty} \frac{n_k}{w^2 + n_k^2} e^{-n_k^2 t} \int_0^l e^{\gamma \left|x - \frac{l}{2}\right|} \sin\left(n_k \frac{x}{l}\right) dx \quad (9.50)$$

Здесь величина предельной сорбции M_{∞} определена уравнением (9.42). Беря интеграл, окончательно получаем

$$M(t) = M_{\infty} - \frac{16M_0 \operatorname{ch}(0.5w)}{1 + e^w} \sum_{k=0}^{\infty} \frac{n_k \left(n_k e^{0.5w} + (-1)^{k+1} w \right)}{(w^2 + n_k^2)^2} e^{-(n_k^2 + w^2)t} \quad (9.51)$$

Для того чтобы найти асимптотику выражения (9.51) при малых t , рассмотрим модель конвекционной диффузии в полуось -

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} + v \frac{\partial c}{\partial x} \quad (9.52)$$

Мы будем решать это уравнение (9.52) при условиях

$$\begin{aligned} c(x, 0) &= 0 \\ c(0, t) &= \mu_0 \end{aligned} \quad (9.53)$$

Отметим, что здесь переменная x изменяется на интервале $[0, +\infty]$. Представим решение уравнения (9.52) при условиях (9.53) как

$$c(x, t) = \mu_0 \exp(\gamma x - \lambda t) \rho(x, t) \quad (9.54)$$

где

$$\lambda = dw^2,$$

а величина γ определена в (9.44). Тогда функция $\rho(x, t)$ будет удовлетворять системе

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= D \frac{\partial^2 \rho}{\partial x^2} \\ \rho(x, 0) &= 0 \\ \rho(0, t) &= e^{-\lambda t} \end{aligned} \quad (9.55)$$

Из уравнения (9.11) следует, что

$$c(x, t) = 2\mu_0 D e^{\gamma x} \int_0^t \frac{\partial G_2}{\partial \xi}(x, 0, \tau) e^{-\lambda \tau} d\tau \quad (9.56)$$

где функция G_2 определена в уравнении (9.12).

Тогда скорость кинетики сорбции можно представить как

$$\frac{dM(t)}{dt} = 2\mu_0 D e^{-\lambda t} \int_0^{\infty} \frac{\partial G_2}{\partial \xi}(x, 0, t) e^{\gamma x} dx \quad (9.57)$$

Проинтегрировав в (9.57) по частям и разлагая в ряд по малым λt , получаем

$$\frac{dM(t)}{dt} = 2\mu_0 D \left\{ \frac{e^{-\lambda t}}{2\sqrt{\pi Dt}} + \frac{\gamma}{2} [1 + \operatorname{erf}(\gamma\sqrt{Dt})] \right\} \approx M_0 d \left[|w| + \frac{1}{\sqrt{\pi}} \left(\frac{1}{\sqrt{dt}} + 2w^2 \sqrt{dt} \right) \right] \quad (9.58)$$

Поэтому

$$M(t) \approx M_0 \left[|w| dt + \frac{2\sqrt{dt}}{\sqrt{\pi}} \left(1 + \frac{2}{3} w^2 dt \right) \right] \quad (9.59)$$

Окончательно модель для конвекционной диффузии можно записать в виде

$$M(t) = M_0 \begin{cases} 2|w|dt + \frac{4\sqrt{dt}}{\sqrt{\pi}} \left(1 + \frac{2}{3} w^2 dt \right), & t < \tau \\ \varphi(w) - \delta\psi(w) \sum_{k=0}^{\infty} S_k, & t \geq \tau \end{cases} \quad (9.60)$$

где

$$\begin{aligned} \varphi(w) &= \frac{e^w - 1 + w}{w(e^w + 1)}, & \psi(w) &= \frac{2 \operatorname{ch}(w)}{e^w + 1} \\ S_k &= \frac{n_k \left(n_k e^{0.5w} + (-1)^{k+1} w \right)}{(w^2 + n_k^2)^2} e^{-(n_k^2 + w^2) dt} \end{aligned} \quad (9.61)$$

На Рис. 9.11 показаны типичные кривые конвекционной диффузии в координатах (M, \sqrt{t}) .

Форма этих кривых напоминает кинетику релаксационной сорбции (см. Рис. 9.10) и они действительно хорошо могут быть приближены моделью (9.36). Например, приведенные на Рис. 9.11 кривые хорошо приближаются релаксационной моделью при следующих значениях параметров

$$\text{Кривая (a)} \quad d=0.2138 \quad r=0.9496 \quad M_1=1.2541 \quad M_0=0.5428$$

$$\text{Кривая (b)} \quad d=0.0924 \quad r=2.8096 \quad M_1=0.3493 \quad M_0=0.6548$$

Точность приближения (среднее отклонение равно 0.00015) столь хороша, что кривые на графиках (a и b) не различимы. При $w > 0$ сорбционная кривая не имеет точки перегиба и может быть приближена с помощью модели нормальной диффузии (9.26).

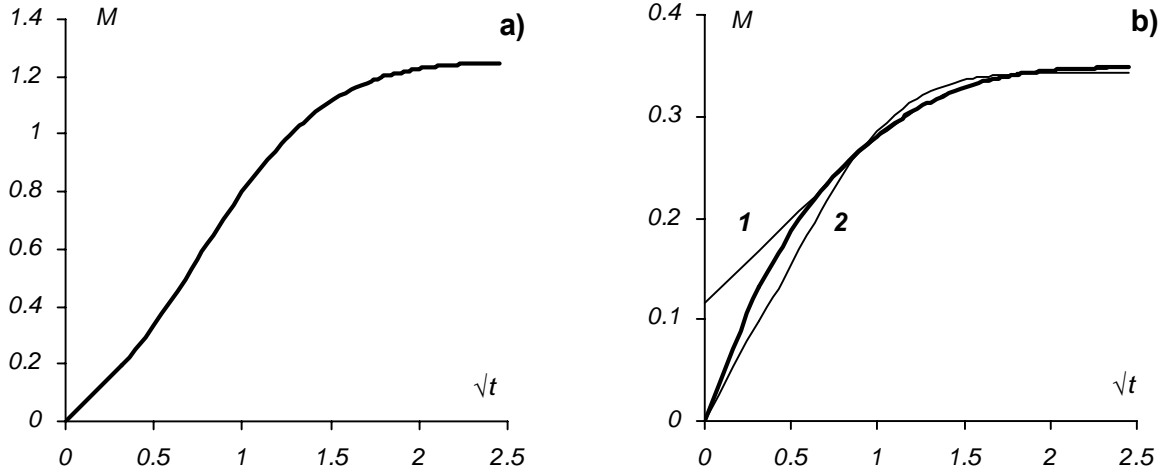


Рис. 9.11 Кинетика конвекционной диффузии для $d=0.05$, $M_0=1$ и $w=-3$ (а) и $w=+3$ (б).

На Рис. 9.11 б) это описание показано тонкими кривыми 1 и 2. Они получены при следующих значениях параметров модели

| | | | |
|------------|------------|--------------|--------------|
| Кривая (1) | $d=0.1017$ | $M_0=0.3486$ | $C_0=0.1154$ |
| Кривая (2) | $d=0.1603$ | $M_0=0.3434$ | $C_0=0.0$ |

Начальный участок кривой 1 и средний участок кривой 2 сильно отличаются от конвекционной кривой. Это показывает, что основное отличие конвекционной модели от нормальной при $w > 0$ – это более быстрый рост сорбционной кривой на начальном участке. При $w < 0$ (левый график а) сорбционная кривая имеет точку перегиба. В конвекционной модели также имеется зависимость от толщины образца, аналогичная той, которая присуща релаксационной модели. Графики такой зависимости вполне аналогичны показанному выше на Рис. 9.7, поэтому не приводятся.

Вернемся опять к типичным аномальным сорбционным кривым, представленным на Рис. 9.5.

Табл. 9.3 Параметры конвекционной модели.

| Кривая | d | M_0 | w |
|--------|-----|-------|------|
| 0 | 1.0 | 1.0 | 0.0 |
| 1 | 1.9 | 0.9 | 2.2 |
| 2 | - | - | - |
| 3 | 0.4 | 1.8 | 0.8 |
| 4 | 0.2 | 0.7 | -1.8 |

В Табл. 9.3 даны значения параметров конвекционной модели (9.60), которые позволяют построить эти кривые. Здесь важно отметить, что кривая (2) на Рис. 9.5, которая характерна тем, что она проходит через максимум, не может быть описана конвекционной моде-

лю. Остальные кривые приближаются с очень хорошей точностью. Таким образом, мы видим, что с помощью релаксационной модели можно количественно описать все особенности, присущие аномальной диффузии, за исключением свойства 1 – не монотонность.

На Рис. 9.12 представлена конвекционная модель в виде, используемом в программе Fitter. Рассмотрим некоторые ее особенности. В разделе, озаглавленном “асимптотика при $t < \tau$ ” используется выражение $\text{abs}W = (w * w)^{0.5}$, которое является ничем иным, как модулем величины w .

В разделе “разрешение неопределенности типа $0/0$ ”, использован прием, который необходимо разобрать подробно. Как видно, функция $\varphi(w)$ в формуле (9.61) имеет особенность при $w=0$. Эта неопределенность легко может быть раскрыта разложением экспоненты в ряд, что дает значение $\varphi(0)=1$. Однако программа не умеет работать с такими особенностями и при попытке использовать модель при $w=0$ дает ошибку вычисления – деление на ноль. Поэтому вычислять функцию $\varphi(w)$ нужно по следующему правилу

$$\varphi(w) = \frac{\varphi_1(w)}{\varphi_2(w)} \quad (9.62)$$

(что соответствует формуле $\text{ffi} = \text{ffi1} / \text{ffi2}$ в Fitter) где

$$\varphi_1(w) = \begin{cases} e^w - 1 + w, & w \neq 0 \\ 1, & w = 0 \end{cases} \quad \varphi_2(w) = \begin{cases} w(e^w + 1), & w \neq 0 \\ 1, & w = 0 \end{cases} \quad (9.63)$$

Для вычисления функций φ_1 и φ_2 используется следующая вспомогательная величина

$$\text{NOzero} = \text{hev}(w) + \text{hev}(-w).$$

Очевидно, что она равна единице везде, кроме точки $w=0$, где она равно нулю. Поэтому, выражения

$$\text{ffi1} = (\text{EX}w - 1 + w) * \text{NOzero} + \text{imp}(w)$$

$$\text{ffi2} = w * (1 + \text{EX}w) * \text{NOzero} + \text{imp}(w)$$

точно соответствуют формулам (9.63). Такой прием является стандартным при работе с Fitter и позволяет использовать модели, в которых присутствуют разрешимые неопределенности, называемые также разрывами 1-го рода [144].

Ряд в формуле (9.60) является знакопеременным, поэтому он сходится гораздо хуже, чем в уравнениях нормальной (9.26) и релаксационной (9.36) диффузии. По этой причине в сумме оставлены уже пять членов, что достаточно для удовлетворительной оценки.

```

'          Convectonal model
M=m1*hev(USEm1)+m2*[hev(-USEm1)+imp(-USEm1)]
'condition to apply asymptotic case m1
    USEm1=m2-m1
'constants and intermediates
    P12=(PI)^(-0.5)
    x=d*t
    psi=2*CHw2/(EXw+1)
    CHw2=0.5*(EXw2+1/EXw2)
    EXw=exp(w)
    EXw2=exp(0.5*w)
'aspytotic at t<tau
    m1=2*M0*[absW*x+2*P12*(x)^0.5*(1+2/3*w*w*x)]
    absW=(w*w)^0.5
'uncertainty 0/0 solution
    ffi=ffil/ffi2
    ffil=(EXw-1+w)*NOzero+imp(w)
    ffi2=w*(1+EXw)*NOzero+imp(w)
    NOzero=hev(w)+hev(-w)
'series case at t>tau
    m2=M0*(ffi-8*psi*S)
    S=S0+S1+S2+S3+S4+S5
    n0=PI
    nw0=n0*n0+w*w
    s0=-1
    S0=n0*(n0*EXw2+s0*w)*exp(-nw0*x)/(nw0*nw0)
    n1=PI*3
    nw1=n1*n1+w*w
    s1=1
    S1=n1*(n1*EXw2+s1*w)*exp(-nw1*x)/(nw1*nw1)
    n2=PI*5
    nw2=n2*n2+w*w
    s2=-1
    S2=n2*(n2*EXw2+s2*w)*exp(-nw2*x)/(nw2*nw2)
    n3=PI*7
    nw3=n3*n3+w*w
    s3=1
    S3=n3*(n3*EXw2+s3*w)*exp(-nw3*x)/(nw3*nw3)
    n4=PI*9
    nw4=n4*n4+w*w
    s4=-1
    S4=n4*(n4*EXw2+s4*w)*exp(-nw4*x)/(nw4*nw4)
    n5=PI*11
    nw5=n5*n5+w*w
    s5=1
    S5=n5*(n5*EXw2+s5*w)*exp(-nw5*x)/(nw5*nw5)
'unknown parameters
    d=?
    M0=?
    w=?

```

Рис. 9.12 Модель конвекционной диффузии в виде, используемом в программе Fitter.

Продемонстрируем практический пример применения модели конвекционной диффузии [137]. На Рис. 9.13 представлена кинетика сорбции воды через торцы алюмостеклопласти-

ка СИАЛ-1Н. Она хорошо описывается моделью, показанной на Рис. 9.12 при следующих значениях параметров

$$d=0.0002 \quad M_0=0.0103 \quad w=-16.6474$$

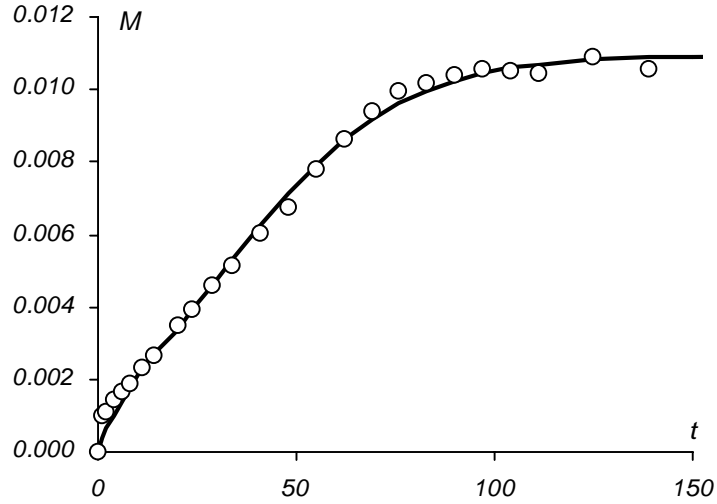


Рис. 9.13 Увлажнение алюмо стеклопластика СИАЛ-1Н (условия выдержки: температура 60°C, относительная влажность 100%, сорбция через торцы), t - время в сутках, M - прирост массы.

9.5. Кинетика цикла «увлажнение–сушка»

Одним из распространенных приемов исследования сорбционных свойств композиционных материалов является эксперимент «увлажнение–сушка». При этом сначала образец какое-то время находится во влажной среде (как правило, в воде), а затем его высушивают. В ходе эксперимента контролируется изменение массы образца.

Для моделирования такого эксперимента естественно рассматривать следующие условия в дифференциальном уравнении (9.5).

$$\begin{aligned} c(x,0) &= c_0 \\ c(0,t) &= c(l,t) = \mu(t) \\ \mu(t) &= \begin{cases} \mu_1 + (\mu_0 - \mu_1)e^{-rt}, & t < t_1 \\ 0, & t > t_1 \end{cases} \end{aligned} \quad (9.64)$$

где t_1 – это время перехода от увлажнения к сушке. Очевидно, что на участке увлажнения (при $t < t_1$) кинетика будет определяться уравнением (9.36), которое удобно записать как

$$\begin{aligned} M(t) &= M_1 + (M_0 - M_1)e^{-rt} - 8 \sum_{k=0}^{\infty} \frac{U_k}{n_k^2} S_k \\ U_k &= \frac{(V_0 n_k^2 d - V_1 r) e^{-n_k^2 d t} + r(M_1 - M_0) e^{-rt}}{(n_k^2 d - r)} \end{aligned} \quad (9.65)$$

Для определения кинетики десорбции нужно решить уравнение (9.5) с нулевыми граничными условиями и начальным значением концентрации $c(x, t_1)$ равным распределению концентрации в последний момент сорбционного цикла. Из уравнения (9.7) следует, что кинетика десорбции имеет вид

$$M(t) = 8 \sum_{k=0}^{\infty} \frac{a_k}{n_k^2} e^{-n_k^2 d(t-t_1)}, \quad t > t_1 \quad (9.66)$$

где коэффициенты a_k зависят только от t_1 . Из условия непрерывности кинетики следует, что

$$a_k = M_1 + (M_0 - M_1)e^{-rt_1} - U_k(t_1) \quad (9.67)$$

Поэтому

$$M(t) = 8 \sum_{k=0}^{\infty} \frac{M_1 + (M_0 - M_1)e^{-rt_1} - U_k(t_1)}{n_k^2} e^{-n_k^2 d(t-t_1)}, \quad t > t_1 \quad (9.68)$$

Очевидно, что асимптотика выражения (9.68) при $t \rightarrow t_1 + 0$ имеет вид

$$M(t) = M(t_1) - \frac{4}{\sqrt{\pi}} \left[M_1 + (M_0 - M_1)e^{-rt_1} \right] \sqrt{d(t-t_1)}, \quad t > t_1 \quad (9.69)$$

Таким образом, мы получаем окончательное выражение для вычисления кинетики цикла «увлажнение–сушка».

$$M(t) = \begin{cases} C_0 + 4\pi^{-1/2} \sqrt{dt} [M_0 - C_0 + (M_1 - M_0)\beta(rt)], & 0 \leq t < \tau \\ K(t) - 8 \sum_{k=0}^{\infty} \frac{U_k(t)}{n_k^2}, & \tau \leq t < t_1 \\ K(t_1) \left[1 - \frac{4}{\sqrt{\pi}} \sqrt{d(t-t_1)} \right] - 8 \sum_{k=0}^{\infty} \frac{U_k(t_1)}{n_k^2}, & t_1 \leq t < t_1 + \tau \\ 8 \sum_{k=0}^{\infty} \frac{K(t_1) - U_k(t_1)}{n_k^2} e^{-n_k^2 d(t-t_1)}, & t_1 + \tau \leq t < \infty \end{cases} \quad (9.70)$$

$$K(t) = M_1 + (M_0 - M_1)e^{-rt}$$

$$U_k(t) = \frac{(V_0 n_k^2 d - V_1 r) e^{-n_k^2 dt} + r(M_1 - M_0) e^{-rt}}{n_k^2 d - r}$$

Оно состоит из четырех частей, каждая из которых используется на своем участке. Здесь функция $\beta(x)$ определена в уравнении (9.35).

На Рис. 9.14 показаны две типичные кинетические кривые, получаемые в цикле «увлажнение-сушка» в координатах (M, t) . На левом графике (а) представлен цикл, в котором сорбция подчиняется нормальной диффузии (см. раздел 9.1). На правом графике (б) пока-

зан цикл, в котором сорбция описывается релаксационной моделью (см. раздел 9.3). В обоих вариантах переход от «увлажнения» к «сушке» происходит в момент времени $t_l=4$.

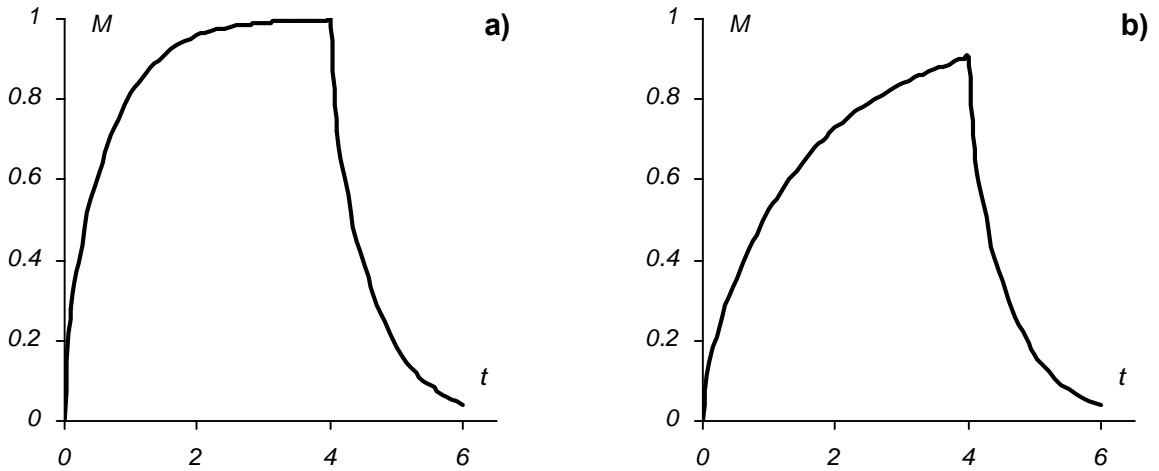


Рис. 9.14 Кинетические кривые цикла «увлажнение-сушка» при $d=0.15$ и $t_l=4$ для нормальной сорбции $M_0=1, M_l=C_0=r=0$ (а) и релаксационной сорбции $M_0=0.5, M_l=1, C_0=0, r=0.5$ (б).

Рассмотрим, как модель (9.70) реализуется в системе Fitter. На Рис. 9.15 показан соответствующий текстовое поле. Разберем его особенности.

Во второй строке приведена основная формула модели

$$M = \text{Sor} * \text{hev}(t_l - t) + \text{Des} * [\text{hev}(t - t_l) + \text{imp}(t - t_l)]$$

в которой переменная Sor отвечает за сорбционный, а переменная Des отвечает за десорбционный участки кинетики. Конструкция с использованием функций $\text{hev}(t - t_l)$ и $\text{imp}(t - t_l)$ обеспечивает переход с одного участка на другой в момент t_l .

Для вычисления величин Sor и Des применяются выражения

$$\text{Sor} = \text{Sor1} * \text{hev}(\text{USESor1}) + \text{Sor2} * [\text{hev}(-\text{USESor1}) + \text{imp}(-\text{USESor1})]$$

$$\text{Des} = \text{Des1} * \text{hev}(\text{USEDdes1}) + \text{Des2} * [\text{hev}(-\text{USEDdes1}) + \text{imp}(-\text{USEDdes1})]$$

которые аналогичны используемым в моделях, показанных на Рис. 9.3, Рис. 9.8 и на Рис. 9.12. Можно только отметить, что в выражении для условия применимости асимптотики десорбции

$$\text{USEDdes1} = \text{Des1} - \text{Des2}$$

изменен знак, т.к. кинетика десорбции – это падающая кривая.

Моделирование диффузии

```

        "Wetting-drying" model
M=Sor*hev(t1-t)+Des*[hev(t-t1)+imp(t-t1)]
' "wetting" kinetics
    Sor=Sor1*hev(USESor1)+Sor2*[hev(-USESor1)+imp(-USESor1)]
' "drying" kinetics
    Des=Des1*hev(USEDes1)+Des2*[hev(-USEDes1)+imp(-USEDes1)]
'conditions to apply asymptotic cases
    USESor1=Sor2-Sor1
    USEDDes1=Des1-Des2
'constants and intermediates
    t3=(t-t1)*hev(t-t1)
    t4=t*hev(t1-t)+t1*[hev(t-t1)+imp(t-t1)]
    P2=PI*PI
    P12=(PI)^(-0.5)
    R=r*(M1-M0)*exp(-r*t4)
    K=M1+(M0-M1)*exp(-r*t4)
    V0=M0-C0
    V1=M1-C0
'asymptotic case of sorption at 0<t<tau
    Sor1=C0+4*P12*(d*t)^0.5*[M0-C0+(M1-M0)*beta]
    beta=1-exp(-z)
    x=r*t
    z=(a1*x+a2*x*x+a3*x*x*x)/(1+b1*x+b2*x*x+b3*x*x*x)
    a1=0.6666539250029
    a2=0.0121051017749
    a3=0.0099225322428
    b1=0.0848006232519
    b2=0.0246634591223
    b3=0.0017549947958
'series case of sorption at tau<t<t1
    Sor2=K-8*S1
    S1=U01/n0+U11/n1+U21/n2+U31/n3+U41/n4
    n0=P2
    U01=[(V0*n0*d-V1*r)*exp(-n0*d*t4)+R]/(n0*d-r)
    n1=P2*9
    U11=[(V0*n1*d-V1*r)*exp(-n1*d*t4)+R]/(n1*d-r)
    n2=P2*25
    U21=[(V0*n2*d-V1*r)*exp(-n2*d*t4)+R]/(n2*d-r)
    n3=P2*49
    U31=[(V0*n3*d-V1*r)*exp(-n3*d*t4)+R]/(n3*d-r)
    n4=P2*81
    U41=[(V0*n4*d-V1*r)*exp(-n4*d*t4)+R]/(n4*d-r)
'asymptotic case of desorption at t1<t<t1+tau
    Des1=K*[1-4*P12*(d*t3)^0.5]-8*S1
'series case of desorption at t1+tau<t
    Des2=8*S2
    S2=U02/n0+U12/n1+U22/n2+U32/n3+U42/n4
    U02=(K-U01)*exp(-n0*d*t3)
    U12=(K-U11)*exp(-n1*d*t3)
    U22=(K-U21)*exp(-n2*d*t3)
    U32=(K-U31)*exp(-n3*d*t3)
    U42=(K-U41)*exp(-n4*d*t3)
'unknown parameters
    d=?
    M0=?
    M1=?
    C0=?
    r=?
    t1=?

```

Рис. 9.15 Модель кинетики цикла «увлажнение-сушка» в виде, используемом в программе Fitter.

Заметим, что промежуточные переменные

$$t_3 = (t - t_1) * \text{hev}(t - t_1)$$

$$t_4 = t * \text{hev}(t_1 - t) + t_1 * [\text{hev}(t - t_1) + \text{imp}(t - t_1)]$$

используемые в модели, позволяют упростить вычисление величин $K(t)$ и $U_k(t)$ в формуле (9.70).

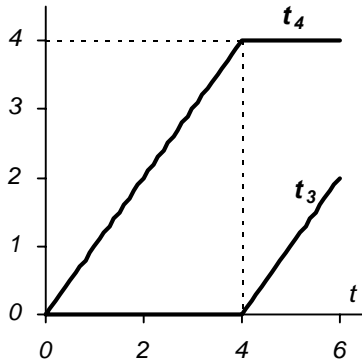


Рис. 9.16 Промежуточные переменные, используемые в модели «увлажнение-сушка» для $t_1=4$.

Их смысл ясен из Рис. 9.16. Видно, что

$$t_3=0, t_4=t \text{ при } t < t_1 \text{ и что}$$

$$t_3=t-t_1, t_4=t_1 \text{ при } t > t_1.$$

На Рис. 9.17 показаны различные асимптотики, используемые в модели. Кривые $1s$ и $2s$ показывают асимптотики для кинетики сорбции при малом и большом t соответственно. Кривые $1d$ и $2d$ демонстрируют асимптотики для кинетики начала процесса десорбции при $t = t_1$ и для его конца.

Рассмотрим реальный пример применения модели «увлажнение-сушка». На Рис. 9.18 представлены данные по термовлажностному старению гибридного композиционного материала со стеклообразной матрицей [145]. Кинетика сорбции влаги соответствует случаю кривой 3 на Рис. 9.5 – медленный рост граничного условия.

На Рис. 9.18 кроме кинетики сорбции-десорбции (кривая 1), которая описывается следующими значениями параметров

$$d=0.030, \quad M_0=2.261, \quad M_1=4.290, \quad r=0.025, \quad t_1=25,$$

изображена также кривая (2) изменения граничного условия

$$M_1 + (M_0 - M_1) \exp(-rt)$$

и кривые кинетики нормальной сорбции (9.26) построенные для следующих значений параметров

$$d=0.030, \quad M_0=2.261, \quad C_0=0,$$

кривая (3) и

$$d=0.030, \quad M_0=4.290, \quad C_0=0,$$

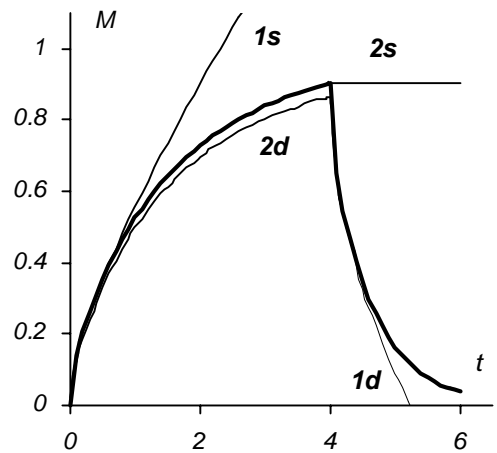


Рис. 9.17 Различные асимптотики, используемые в модели «увлажнение-сушка».

кривая (4).

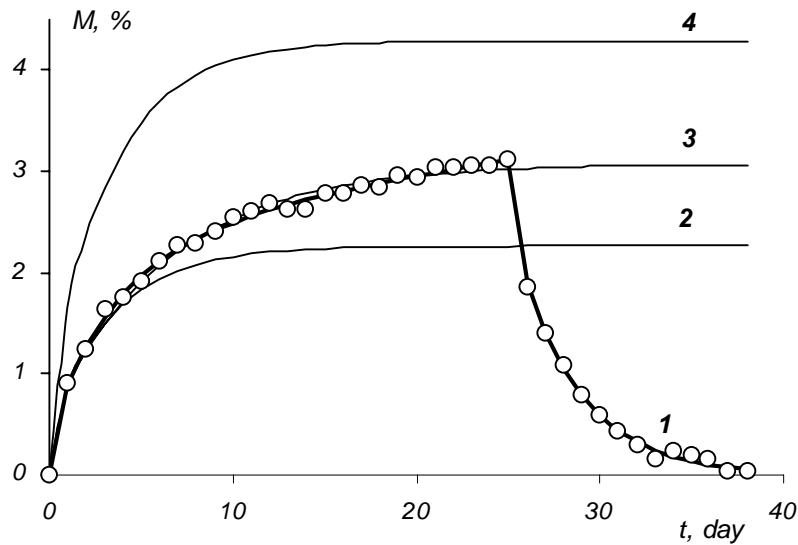


Рис. 9.18 Кинетика изменения влагосодержания в цикле "увлажнение-сушка" гибридного композиционного материала при температуре 70 С.

Повторяя циклы «увлажнение-сушка» с одним и тем же образцом, можно исследовать, как меняются значения начальной M_0 и предельной M_l сорбции в зависимости от номера цикла. В работе [145] показано, что во втором цикле значение начальной сорбции будет равно значению сорбции, достигнутой в первом цикле, т.е. $M_0=3.120$. Кроме того установлено, что наблюдается хорошая корреляция прочностных свойств образцов со значением предельной сорбции M_l , определенной для этого цикла, причем это не зависит от того в каких условиях (температура и время) проходила сорбция. Этот факт доказывает, что величина предельной сорбции M_l действительно характеризует необратимые структурные изменения, происходящие в материале под действием влаги и температуры.

На Рис. 9.19 показана кинетика цикла «увлажнение-сушка» для гибридного композиционного материала [138]. Эти данные интересны тем, что в них имеется ненулевое начальное значение. Это обстоятельство, однако, свидетельствует вовсе не о наличии влаги в образце в исходном состоянии, а о деструктивных процессах, проходящих в материале в ходе эксперимента.

Если перенести начало координат в точку соответствующую исходному состоянию образца (правая ось ординат на рисунке) то, начиная с момента времени $t=93$ дня, значения сорбционной кривой будут отрицательными. Это означает, что исходный образец потерял в ходе эксперимента больше сорбата, чем приобрел сорбента.

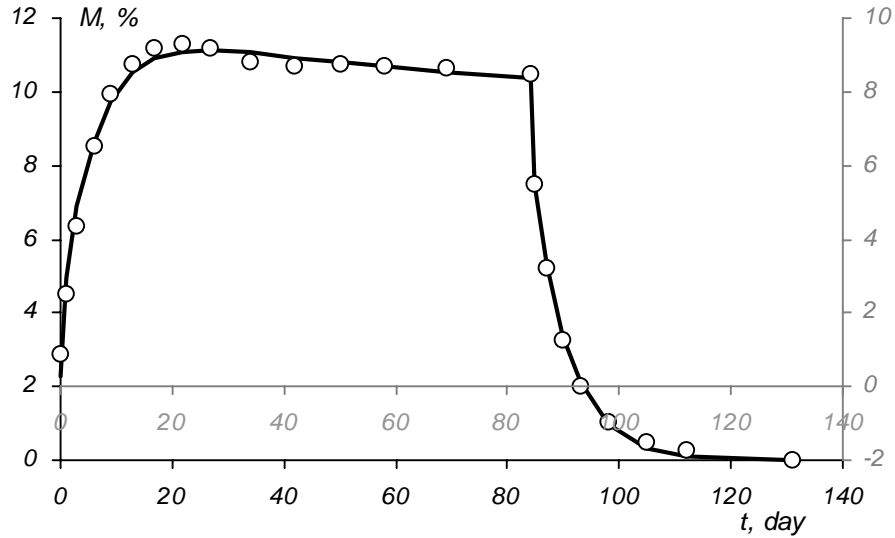


Рис. 9.19 Кинетика изменения влагосодержания в цикле "увлажнение-сушка" гибридного композиционного материала при температуре 70 С.

9.6. Другие модели сорбции

Для описания кинетики аномальной сорбции используется много разных моделей. В частности, модель Беренса и Хопфенберга [140], которая предполагает линейную суперпозицию вкладов относительного приращения массы за счет фиковской диффузии и релаксации первого порядка и каждому механизму приписывает фиксированную фракцию общего равновесного количества сорбированного растворителя. Такие модели имеют простую структуру, которая, как правило, состоит из стандартной нормальной модели диффузии (9.26) и несложных «добавок» к ней. Типичным примером является модель [141]

$$M(t) = (1 - \alpha)M_F + \alpha M_R$$

$$M_R = M_0 \frac{1 - \exp(-k\beta t)}{1 + (\beta - 1)\exp(-k\beta t)}; \quad \beta = \frac{d_\infty}{d_0} \quad (9.71)$$

комбинирующая кинетику (9.26) нормальной диффузии, обозначенную в (9.71) M_F , и релаксационный член M_R . Очевидно, выражение (9.71) может легко быть смоделировано в системе Fitter на базе модели, показанной на Рис. 9.3.

В литературе часто встречается модель сорбции, в которую добавлена химическая реакция первого порядка [129, 130], интерпретируемая как гидролиз сорбата

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} - kc, \quad k \geq 0$$

$$c(x, 0) = c_0$$

$$c(0, t) = c(l, t) = \mu_0 \quad (9.72)$$

Популярность этой модели объясняется, по-видимому, ее простотой. Действительно, заменой $c = u e^{-kt}$, дифференциальное уравнение в (9.72) можно преобразовать к стандартному виду (9.5). Однако, многие исследователи забывают учесть, что такая замена преобразует и граничные условия в системе (9.72). Они будут иметь вид

$$u(0, t) = u(l, t) = \mu_0 e^{kt}$$

Отсюда, легко видеть, что кинетика сорбции с химической реакцией, описываемая системой (9.72), сводится к релаксационной диффузии с граничным условием (9.27) при $r = -k$ и $\mu_l = 0$. Поэтому уравнение для кинетики сорбции имеет вид

$$M(t) = \begin{cases} e^{-kt} \left\{ C_0 + 4\pi^{-1/2} \sqrt{dt} [M_0(1 - \beta(-kt)) - C_0] \right\}, & t < \tau \\ M_0 - 8 \sum_{k=0}^{\infty} S_k, & t \geq \tau \end{cases} \quad (9.73)$$

$$S_k = \frac{[(M_0 - C_0)n_k^2 d - kC_0] e^{-(n_k^2 d + k)t} + kM_0}{n_k^2 (n_k^2 d + k)}$$

Здесь функция $\beta(x)$ определена соотношением (9.35). На Рис. 9.20 показаны несколько сорбционных кривых в координатах (M, \sqrt{t}) , построенных по этой модели для различных значений константы скорости реакции гидролиза $k = -4, -2, 0, 2, 4$.

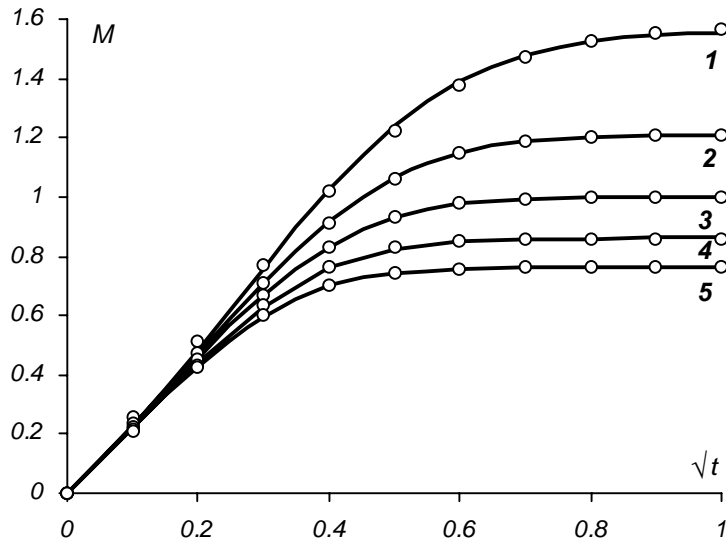


Рис. 9.20 Кинетические кривые диффузии с реакцией в координатах (M, \sqrt{t}) при $D=1$, $\mu_0=1$, для различных значений константы скорости реакции $k=-4, -2, 0, 2, 4$.

Из рисунка видно, что форма сорбционных кривых близка к нормальной, фиковской. Это обстоятельство подтверждается расчетом, результаты которого представлены на Рис. 9.20

точками. Эти значения были получены после подгонки кривых моделью нормальной сорбции.

Табл. 9.4 Параметры нормальной модели, использованные для подгонки кривых сорбции с реакцией.

| Кривая | k | d | M_0 |
|--------|------|-------|-------|
| 1 | -4.0 | 0.526 | 1.572 |
| 2 | -2.0 | 0.756 | 1.212 |
| 3 | 0.0 | 1.000 | 1.000 |
| 4 | 2.0 | 1.259 | 0.860 |
| 5 | 4.0 | 1.531 | 0.760 |

В Табл. 9.4 приведены значения полученных оценок параметров. Точность приближения очень хорошая – средняя ошибка приближения равна $9.7E-3$, что в прочем и так видно из представленного рисунка Рис. 9.20.

Решив стационарный вариант системы (9.72), можно определить, что предельное значение равновесной сорбции представляется выражением

$$M_{\infty} = \begin{cases} 2M_0 \sqrt{-\frac{d}{k}} \operatorname{tg} \left(\frac{1}{2} \sqrt{-\frac{k}{d}} \right) & k < 0 \\ M_0, & k = 0 \\ 2M_0 \sqrt{\frac{d}{k}} \operatorname{th} \left(\frac{1}{2} \sqrt{\frac{k}{d}} \right) & k > 0 \end{cases} \quad (9.74)$$

На Рис. 9.21 представлен график этой зависимости.

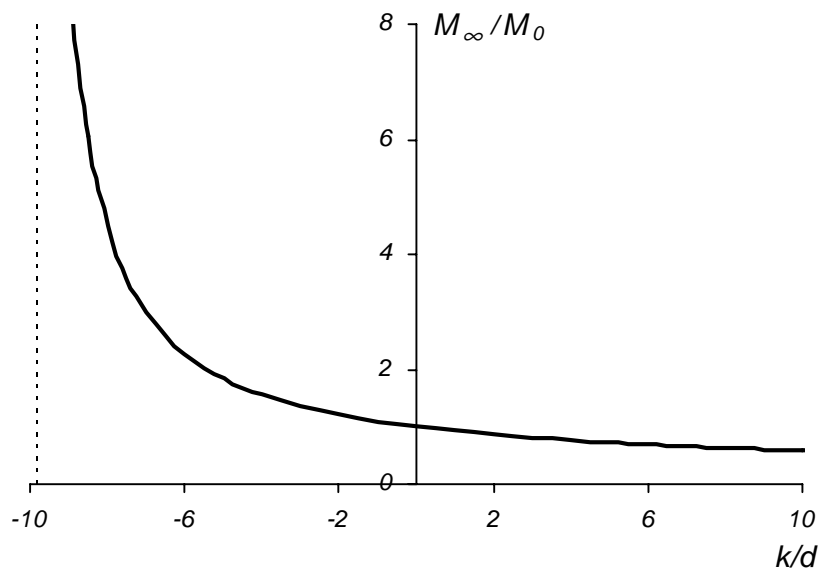


Рис. 9.21 Зависимость предельной равновесной сорбции от константы скорости реакции.

Видно, что предельная сорбция стремится к бесконечности при $k/d \rightarrow -\pi^2$. Это соответствует ситуации, при которой сорбция не может скомпенсировать прирост массы в результате химической реакции. Впрочем, вариант модели (9.72) при $k < 0$, физического смысла, по-видимому, не имеет.

Таким образом, мы можем сделать вывод о том, что использование системы (9.72) – диффузия с химической реакцией – для моделирования сорбции не оправдано. Этот вывод основан, прежде всего, на том, что форма получаемых кинетических кривых соответствует модели нормальной сорбции и хорошо ей приближается.

9.7. Результаты главы 9

Были рассмотрены несколько моделей, пригодных для описания аномальной диффузии. По-видимому, *релаксационная* модель является наиболее удобной для применения. Она позволяет описать все аномалии, которые являются типичными для нефиковской диффузии. Кроме того, эта модель содержит в себе достаточно серьезное физическое обоснование, которое может быть использовано для интерпретации экспериментальных данных. Анализ значений параметров этой модели позволяют давать некоторое разумное объяснение поведению материала.

Следующей по значимости является *конвекционная* модель, которая почти также удобна, как и релаксационная модель. Конвекционная модель также имеет некоторый физический смысл, но она не в состоянии описать сорбционные кривые с максимумом, что является типичным эффектом в аномальной сорбции. С формальной точки зрения этот недостаток происходит из-за нехватки свободных параметров в конвекционной модели – их там только три против пяти в релаксационной модели. Однако этой модели следует всегда отдавать предпочтение в случаях, когда есть достаточные основания предполагать наличие конвекционного эффекта.

Модель «увлажнение-сушка» имеет особое значение для моделирования диффузии в композиционных материалах. Она позволяет описать данные, получаемые в специальном плане эксперимента, который, как известно, является наиболее чувствительным при исследовании свойств полимеров. Эта модель объединяет нормальную и релаксационные модели диффузии и поэтому она является очень гибким и устойчивым описанием.

Другие рассмотренные модели значительно хуже подходят для описания аномальной диффузии. В частности, комбинация диффузии с химической реакцией оказалась довольно слабым методом для описания эффектов нефиковской сорбции.

Все эти модели рассчитывались и подгонялись с применением системы Fitter. Соответствующие текстовые поля с моделями были приведены в этой главе. Было показано, что даже наиболее сложные, многокомпонентные описания могут с успехом быть использованы в этой программе. Главным преимуществом разработанного подхода является то, что каждое большое или малое уточнение модели может легко быть сделано непосредственно в текстовом поле модели, используя естественный способ записи уравнений.

10. Обработка кривых титрования

Важными объектами математического моделирования в аналитической химии являются *титриметрические* процессы, отличающиеся многообразием химических реакций и регистрируемых сигналов [146]. Титриметрия – один из старейших методов, применяемый в химическом анализе уже более двухсот лет. Однако развитие современных экспериментальных методов (атомная абсорбция, хроматографии, масс-спектрометрии и др.) привело к пересмотру отношения ко многим традиционным подходам, в частности и к титриметрическому. Это объясняется тем, что, не смотря на значительную универсальность, селективность и точность, титриметрический процесс в его традиционном исполнении плохо поддавался автоматизации. В то же время реальные потребности не позволили полностью отказаться от титриметрии, и метод остался в арсенале аналитических служб со своим кругом задач, мало доступным другим методам, например: определение термодинамических характеристик химических равновесий: констант ионизации, автопротолиза и устойчивости, произведений растворимости, стандартных и окислительно-восстановительных потенциалов, стехиометрии соединений и проч.

Использование современных математических методов вдохнуло свежую струю в титриметрический анализ. Обработка кривых титрования на основе многопараметрической подгонки моделей описана авторами многих публикаций. При этом использовалось разложение функции титрования в ряд Тейлора в окрестности начальных приближений параметров до членов первого порядка с последующим расчетом поправок линейным МНК [147-148]. В книге [149] приведен полный текст программы NONLIN, реализующей данный подход к решению обратных задач на языке ФОРТРАН-10. Эта программа в модифицированном виде часто используется в целях сопоставления с программами, реализующими алгоритмы линейной анаморфозы. Нелинейный регрессионный анализ данных титрования минимизацией суммы квадратов невязок также неоднократно освещен, например, в работах [150-151].

Уравнения кривых титрования, выражающих зависимость измеряемого аналитического сигнала y (потенциала E , pH , диффузионного тока I , оптической плотности D и др.) от переменной состояния x (объема добавленного титранта V , количества электричества Q , пошедшего на его электрохимическое генерирование) нередко весьма сложны и не могут быть записаны в явной, относительно y , форме. Это в определенной мере затрудняет использование моделей кривых титрования для решения обратной задачи, т.е. для оценива-

ния параметров по измеренным точкам кривой. Если целью титрования является определение концентрации, то в число искомых параметров, наряду с эквивалентным объемом титранта V_I или количеством электричества Q_I , затраченным на его генерирование в точке эквивалентности, необходимо включить константу равновесия химической реакции K или произведение растворимости, константу ионизации, а также дополнительно реальный потенциал системы E_I при логарифмическом (потенциометрическом) титровании и коэффициент чувствительности детектируемого иона k при амперометрическом, спектрофотометрическом и им подобным линейных титрованиях. Таким образом, по данным титрования необходимо оценить как минимум два, а чаще три неизвестных параметра.

В основе традиционного подхода [151-158] к обработке кривых титрования лежит преобразование данных в некоторых новых координатах, специально подбираемых для каждой задачи (см. L -метод в разделе 3.1). При этом преобразованные данные уже являются линейными относительно точки эквивалентности и константы реакции, которые определяются простейшим линейным МНК. Если в уравнение для кривой входит в качестве неизвестного параметра реальный потенциал системы или коэффициент чувствительности, то он находится подбором. Другие параметры титрования, например, начальный объем титруемого раствора, молярная концентрация титранта, предполагаются известными.

Такой способ обработки данных представляется несколько устаревшим. Основная проблема – это оценка третьего, а, тем более, и четвертого и т.д. параметра системы. Предположение о том, что все прочие параметры системы всегда известны точно, вызывает сомнения в случаях, когда растворы реагента малоустойчивы или их стандартизация затруднена. Необходимость подбирать для каждой задачи свои специальные преобразования, также сильно затрудняет всю процедуру. Кроме того, имеются и более сложные задачи, например, кислотно-основное титрование в смешанных водно-органических и неводных растворителях. Математические описания кривых титрования в таких средах весьма сложны и не поддаются линеаризации заменой переменных

Использование подходов и методов, описанных в этой работе, позволяет существенно упростить процедуру обработки титрометрических кривых. В этом случае модель не требует линеаризации и может использоваться в своем естественном виде. При этом можно извлекать из данных не 2-3 параметра, а существенно больше информации.

По-видимому, интересно провести сравнение результатов, получаемых с помощью системы Fitter с оценками, полученными другими способами. При этом сравнение точности

различных результатов является довольно сложным. Величины среднеквадратичных отклонений, приводимые в цитируемых ниже работах, вызывают большое сомнение, связанное с процедурой линеаризации. Поэтому, предлагается следующий объективный критерий сравнения качества оценок.

Пусть $\hat{a}_1, \hat{a}_2, \dots$ – различные оценки, построенные по одним и тем же данным, а S_1, S_2, \dots – соответствующие суммы квадратов отклонений (1.22). Пусть S_0 – минимальное значение среди этих сумм, т.е. $S_0 = \min\{S_1, S_2, \dots\}$. Тогда для каждой оценки можно вычислить тестовые значения статистики Фишера (1.54), используемой для проверки гипотезы об альтернативных значениях параметров (см. раздел 1.4).

$$\xi_i = \frac{N_f(S_i - S_0)}{pS_0} \quad (10.1)$$

где N_f – число степеней свободы (1.44), а p – это число неизвестных параметров (1.7). Кроме того, можно установить некоторое критическое значение (1.55), которое определяется через квантиль F -распределения (1.52)

$$t = F_{p, N_f}(1 - \alpha) \quad (10.2)$$

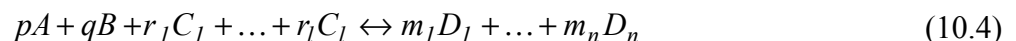
Определим теперь величину, называемую U -критерием

$$U_i = \frac{\xi_i}{t}, \quad (10.3)$$

которую можно вычислить для каждой сравниваемой оценки на выбранном уровне значимости (обычно $\alpha=0.05$). Чем меньше величина U , тем лучше оценка. Те оценки, для которых $U > 1$, являются неприемлемыми, т.к. они не выдерживают критерия альтернативных параметров, описанного в разделе 1.4. Рассмотрим теперь различные примеры, детали которых можно найти в [126], где приведен файл *TITR.XLS*, содержащий исходные данные и детали обработки.

10.1. Линейное титрование

В работе [152] рассматривается уравнение кривой линейного титрования, основанного на одностадийной реакции типа



где A, B – титруемая и титрующая частицы, C_i – другие компоненты реакции, концентрации которых предполагаются постоянными, D_i – продукты реакции, p, q, r_i, m_i – стехиометрические коэффициенты. Если в ходе титрования измеряется зависимость силы тока $I = k_A[A]$ от объема титранта V , то уравнение кривой титрования имеет вид

$$\frac{Vb^{n-1}R_A^k}{(1-bR_A)^n} = V_1R_A^k \left(\frac{b}{1-bR_A} \right)^{n-1} + G, \quad b = 1 + \frac{V}{V_0}, \quad R_A = \frac{I}{I_0} \quad (10.5)$$

где V , V_0 , и I_0 – предикторы, k , n – константы, V_1 , G – неизвестные параметры. Если ввести новые переменные

$$x = R_A^k \left(\frac{b}{1-bR_A} \right)^{n-1}, \quad y = \frac{Vb^{n-1}R_A^k}{(1-bR_A)^n} \quad (10.6)$$

то уравнение (10.5) преобразуется к виду

$$y = V_1x + G \quad (10.7)$$

которое уже легко оценивается. Именно таким способом получены оценки параметров V_1 , G в цитируемой работе.

Нам же удобно записать уравнение (10.5) в виде

$$b^{n-1} \left(\frac{I}{I_0} \right)^k \left[V - V_1 \left(1 - b \frac{I}{I_0} \right) \right] - G \left(1 - b \frac{I}{I_0} \right)^n = 0 \quad (10.8)$$

пригодном для применения в системе Fitter.

На Рис. 10.1 показано соответствующее текстовое поле с этой моделью.

```
'Model 1
0=b^(n-1)*(I/I0)^k*{V-V1*[1-b*(I/I0)]}-G*[1-b*(I/I0)]^n; 0<I<10
    b=1+V/V0
        k=1
        n=1
                v1=?
                G=?
```

Рис. 10.1 Модель для кривой титрования (10.8)

Она использовалась для определения точки эквивалентности и константы титрования на примере амперометрического титрования перхлората меди раствором комплексона III и перхлората железа серной кислотой [153].

Данные показаны на Рис. 10.2., а результаты оценивания приведены в Табл. 10.1.

Табл. 10.1 Результаты оценки параметров модели (10.8) Приведены также значения статистики Фишера (10.1) и U -критерия (10.3).

| Метод | V_1 | G | $S \times 10^{+5}$ | ξ | U |
|--------------|-------|-------|--------------------|-------|-------|
| Fitter | 1.244 | 0.189 | 0.442 | 0.000 | 0.000 |
| Линеаризация | 1.247 | 0.191 | 1.147 | 4.794 | 0.932 |

Здесь указаны оценки параметров, найденные системой Fitter и приведенные в цитируемой работе [152] (вторая строка), соответствующие суммы отклонений (четвертый столбец), значения статистик (10.1) (пятый столбец) и значения U -критерия (10.3) (последний столбец). Видно, что значения полученные методом линеаризации находятся на границе допустимых отклонений параметров, т.к. $U \approx 1$.

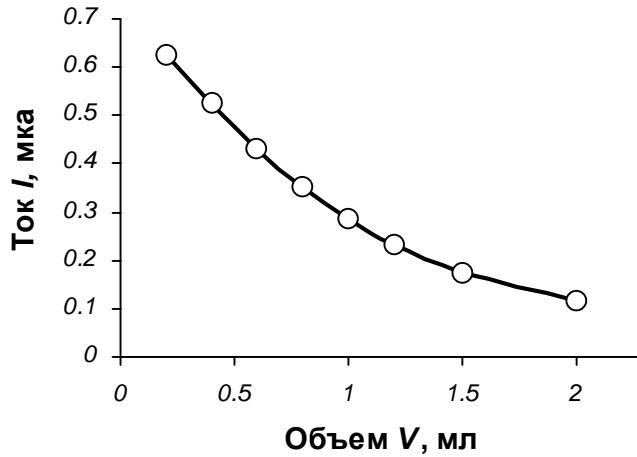


Рис. 10.2 Данные титрования для модели (10.8)

Аналогичная задача рассмотрена в работе [154], где проведена индикация по титрующему компоненту $I = k_B[B]$. Тогда уравнение кривой титрования имеет вид

$$R_B^k \left(\frac{V}{bV_0} - R_B \right)^{l-n} = \frac{V_I R_B^k}{bV_0} \left(\frac{V}{bV_0} - R_B \right)^{-n} - G, \quad b = I + \frac{V}{V_0}, \quad R_B = \frac{I}{Mk_B} \quad (10.9)$$

где V , V_0 , и I_0 – предикторы, k , n , M – константы, V_I , G , k_B – неизвестные параметры. Здесь уже имеются не два, а три неизвестных параметра, что затрудняет линеаризацию модели. Поэтому параметр k_B подбирается, а не оценивается.

Нам удобнее записать это уравнение в виде

$$0 = (R_B I)^k \left[R_B I + \frac{V_I - V}{V + V_0} \right] - G \left[\frac{V}{V + V_0} - R_B I \right]^n \quad (10.10)$$

На Рис. 10.3 показано текстовое поле с соответствующей моделью системы Fitter.

```
'Model 2
0=(Rb*I)^k*[Rb*I+(V1-V)/(V+V0)]-G*[V/(V+V0)-Rb*I]^n; 0<I<7
k=1
n=1
V1=?
G=?
Rb=?
```

Рис. 10.3 Модель для кривой титрования (10.10)

Эта модель использовалась для определения параметров титрования на примере амперометрического титрования комплексона III раствором нитрата кадмия [154]. Данные показаны на Рис. 10.4.

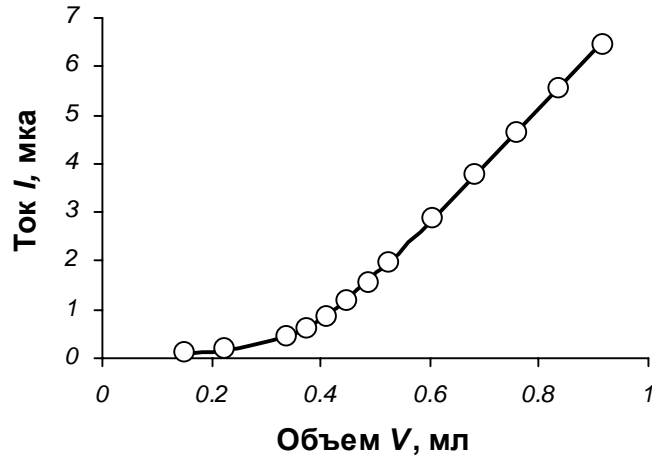


Рис. 10.4 Данные титрования для модели (10.10)

В работе [154] приведены оценки параметров модели (10.10), полученные различными способами. Суть этих методов довольно проста. Все они базируются на некоторых специальных преобразованиях, которые линеаризуют исходную модель. Различие состоит только в способах линеаризации и количестве данных, используемых для оценки. Эти результаты представлены в Табл. 10.2.

Табл. 10.2 Результаты оценки параметров модели (10.10) различными способами. Приведены также значения суммы квадратов отклонений и U -критерий.

| Метод оценки | V_I | $G \times 10^{+3}$ | $R_B \times 10^{+3}$ | $S \times 10^{+3}$ | U |
|------------------|-------|--------------------|----------------------|--------------------|-------|
| Fitter | 0.385 | 0.253 | 2.02 | 11.3 | 0.000 |
| Конфигурационный | 0.378 | 0.233 | 2.05 | 14.2 | 0.269 |
| Уточнение | 0.375 | 0.217 | 2.06 | 16.2 | 0.450 |
| Спрявление | 0.381 | 0.274 | 2.04 | 22.6 | 1.680 |
| Графический | 0.365 | 0.210 | 2.11 | 43.5 | 2.925 |

Первый столбец содержит название метода, второй, третий и четвертый – значения оценок параметров. В пятом столбце указано значение невязки, т.е. суммы квадратов отклонений (1.22), а в последнем столбце приведены значения U -критерия. Из этой таблицы видно, что последние два метода – «Спрявление» и «Графический» дают значения лежащие вне области адекватности модели. Также видно, что с помощью системы Fitter получаются самые точные оценки.

10.2. Потенциометрическое титрование

В работе [155] предлагается метод оценки точки эквивалентности и константы аналитической реакции при потенциометрическом титровании неактивного вещества электроактивным реагентом. Метод также основан на преобразовании кривой титрования в линейную регрессию, параметры которой определяются традиционным способом. Уравнение $E(V)$, выражающее изменение равновесной концентрации иона В в ходе титрования иона А с образованием трудно растворимого соединения имеет вид

$$\frac{(V + V_0)^{n-1} R^k}{[V - (V_0 - V)R]^n} [V_1 - V + (V_0 + V)R] + G = 0, \quad R = 10^{\frac{E-E_t}{m}} \quad (10.11)$$

где V и V_0 , – предикторы, k , n , m и – константы, V_1 , G , E_t – неизвестные параметры. Это уравнение линеаризуется в координатах

$$x = \frac{(V + V_0)^{n-1} R^k}{[V - (V_0 - V)R]^n}, \quad y = x[V - (V_0 + V)R] \quad (10.12)$$

Нам удобнее записать это уравнение в виде

$$0 = R^k (V + V_0)^n [V_1 - V + (V + V_0)R] + G(V + V_0) [V - (V + V_0)R]^n \quad (10.13)$$

где

$$R = \exp[h(E - E_t)], \quad h = \frac{\ln(10)}{m} \quad (10.14)$$

Здесь есть проблема, связанная с вычислением степеней отрицательных величин. С физической точки зрения выражение

$$[V - (V + V_0)R]^n$$

должно быть положительно, однако при подгонке параметров программа может "заблестить" и в физически некорректные области. Приходится изменить модель, чтобы спасти ситуацию и представить это выражение более сложным образом.

$$[V - (V + V_0)R]^n = \{[V - (V + V_0)R][V - (V + V_0)R]\}^{n/2}$$

На Рис. 10.5 показано текстовое поле с соответствующей моделью системы Fitter, в которой сделано еще несколько очевидных подстановок.

```
'Model 3
E=Et+log(R)/h
0=R^k*[V1-V+b*R]*b^n+G*b*{[V-b*R]*[V-b*R]}^(0.5*n); 0<R<1
b=V+V0
k=1
n=0
h=log(10)/59.1
G=?
V1=?
Et=?
```

Рис. 10.5 Модель для кривой титрования (10.11)

Эта модель использовалась для определения параметров титрования на примере осадительного титрования KIO_3 раствором $AgNO_3$. Данные показаны на Рис. 10.6.

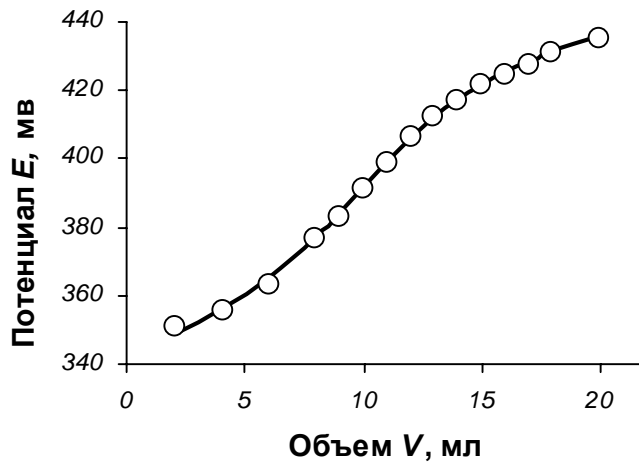


Рис. 10.6 Данные титрования для модели (10.11)

В работе [155] приведены оценки параметров модели(10.11), полученные способом, описанным выше и градиентным поиском при фиксированном $E_f=485$. Эти величины представлены в Табл. 10.3.

Табл. 10.3 Результаты оценки параметров модели (10.11) различными способами. Приведены также значение суммы квадратов отклонений и U -критерий.

| Метод оценки | V_1 | $G \times 10^{+3}$ | $E_f \times 10^{+3}$ | $S \times 10^{+3}$ | U |
|---------------------|--------|--------------------|-----------------------|--------------------|-------|
| Fitter | 10.299 | -0.787 | 485.95 | 9.612 | 0.000 |
| Фиксированный E_f | 10.08 | -0.785 | 485.00 ^{*)} | 11.940 | 0.437 |
| Спрявление | 9.985 | -0.762 | 484.70 ^{**)} | 14.416 | 0.697 |

^{*)} значение фиксировано; ^{**)} значение подобрано

Из таблицы видно, что с помощью системы Fitter получаются самые точные оценки. Кроме того, при использовании второго и третьего методов, первая точка в данных $V=2.02$, $E=351$ является выбросом.

В работе [156] рассматривается задача потенциометрического титрования при определении серебра в шлаках. Уравнение кривой имеет вид

$$\frac{VR}{b} = V_1 R \left(\frac{I}{b} - R \right) + G, \quad R = 10^{\frac{E-E_t}{m}}, \quad b = \frac{V+V_0}{V_0} \quad (10.15)$$

Очевидно, что это частный случай уравнения (10.11).

В работе [157] рассматривается титрование слабой одноосновной кислоты НА сильным основанием В. Уравнение кривой титрования $pH(V)$ имеет вид

$$a_H(V - V_2 + R) = K^M (V_1 - V + V_2 - R) \\ R = \frac{V + V_0}{C_B} \left(\frac{a_H}{g_1} - \frac{g_2}{a_H} \right), \quad a_H = 10^{-pH} \quad (10.16)$$

Нам удобнее записать его в виде

$$\left[V - V_2 + \left(\frac{V + V_0}{C_B} \right) \left(\frac{a_H}{g_1} - \frac{g_2}{a_H} \right) \right] [a_H \exp(a_{10} pK) + 1] - V_1 = 0 \\ pH = -a_{10} \ln(a_H) \quad (10.17)$$

где pH – отклик, V – предиктор, a_H – промежуточная переменная, V_0, C_B, g_1, g_2 и $a_{10} = \ln(10)$ – константы, V_1, V_2 и $pK = -\lg(K^M)$ – неизвестные параметры

На Рис. 10.7 показано текстовое поле с соответствующей моделью системы Fitter.

```
'Model 4
pH=-log(aH)/a10
0={V-V2+(V+V0)/Cb*[aH/g1-g2/aH]}*[aH*exp(a10*pK)+1]-V1; 1E-5<aH<0.05
V0=10
g1=0.83
g2=1.32E-14
Cb=0.0991
a10=log(10)
    pK=?
    V1=?
    V2=?
```

Рис. 10.7 Модель для кривой титрования (10.17)

Эта модель использовалась в примере титрования НАД раствором КОН на фоне 0.1 М (alk)₄NBr. Данные показаны на Рис. 10.8

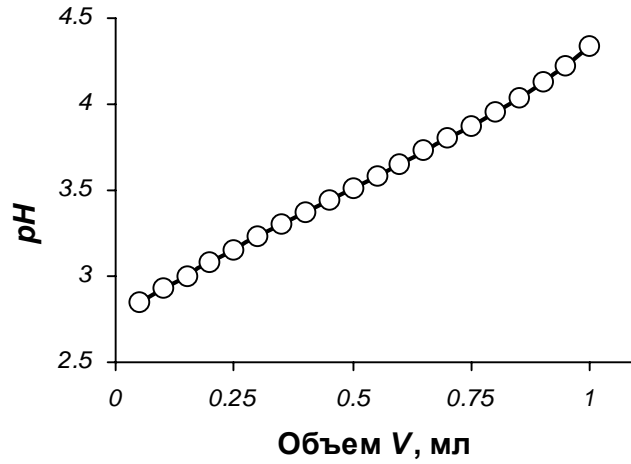


Рис. 10.8 Данные титрования для модели (10.17).

В работе [157] приведены оценки параметров модели (10.17), полученные с помощью линеаризации. Данные были также обработаны системой Fitter в двух вариантах. В первом случае константа g_1 считалась неизвестной и оценивалась, а во втором ее значение фиксировалось равным значению, приведенному в цитируемой работе. Эти величины представлены в Табл. 10.4.

Табл. 10.4 Результаты оценки параметров модели (10.17) различными способами. Приведены также значение суммы квадратов отклонений и U -критерий.

| Метод | pK | V_1 | V_2 | g_1 | $S \times 10^{+5}$ | U |
|------------|--------|--------|-----------------------|--------------------|--------------------|--------|
| Fitter 1 | 3.6828 | 1.1272 | 0.0829 | 0.8016 | 2.65 | 0.0000 |
| Fitter 2 | 3.6767 | 1.1332 | 0.0752 | 0.83 ^{*)} | 3.19 | 0.2953 |
| Спрявление | 3.6766 | 1.1322 | 0.0752 ^{**)} | 0.83 ^{*)} | 4.69 | 0.7699 |

*) значение фиксировано; **) значение подобрано

Из таблицы видно, что Fitter дает не только наилучшие оценки, но и позволяет извлечь из данных больше информации, чем это делается традиционным способом.

В заключение рассмотрим еще одну работу [158], где приводится модель кулонометрического титрования $E(Q)$

$$10^{\frac{zE}{\theta}} \left(Q - FV 10^{\frac{z(E-E_0)}{\theta}} \right) = 10^{\frac{zE}{\theta}} Q_e - P \cdot F \cdot V \cdot 10^{\frac{zE_0}{\theta}} \quad (10.18)$$

которое мы представим в виде

$$\exp(aE) \{ Q - Q_0 - FV \exp[a(E - E_0)] \} + PFV \exp(aE_0) = 0 \quad (10.19)$$

где E – отклик, Q – предиктор, V , F , и $a = z \ln(10) / \theta$ – константы, P , Q_0 и E_0 – неизвестные параметры.

На Рис. 10.9 показано текстовое поле с соответствующей моделью системы Fitter.

```
'Model 5
0=exp(a*E)*[Q-Q0-F*V*exp(a*(E-E0))] + N*P*V*F*exp(a*E0); 150<E<400
V=112
a=0.039207
F=96.48456
N=1e-10
P=?
Q0=?
E0=?
```

Рис. 10.9 Модель для кривой титрования (10.19)

Здесь добавлена нормировочная константа $N=10^{-10}$, которая облегчает поиск оценок в соответствии с общей концепцией, изложенной в разделе 4.4.

Эта модель использовалась в примере титрования NaCl электрогенерированным Ag^+ на фоне $0.1 \text{ M KNO}_3 + 0.5 \text{ M HNO}_3$. Данные показаны на Рис. 10.10.

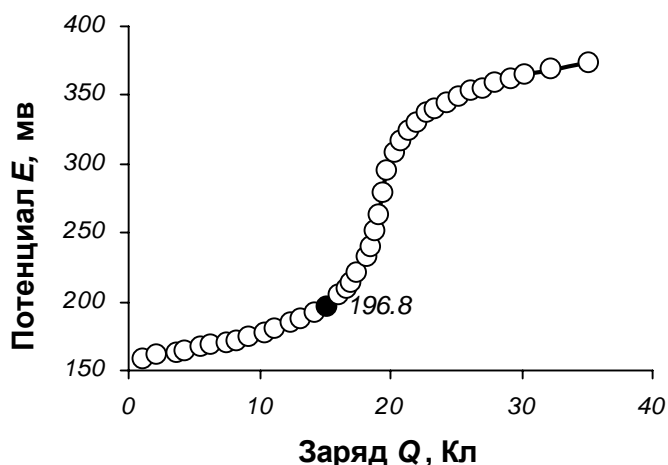


Рис. 10.10 Данные титрования для модели (10.19).

Здесь данные также были обработаны системой Fitter в двух вариантах. В первом случае константа a считалась неизвестной и оценивалась, а во втором ее значение фиксировалось равной величине, приведенной в цитируемой работе. В обоих случаях наблюдался выброс $E=196.8$ в точке, отмеченной на графике. Расчеты проводились без этого измерения. Результаты представлены в Табл. 10.5

Табл. 10.5 Результаты оценки параметров модели (10.19) различными способами. Приведены также значение суммы квадратов отклонений и U -критерий.

| Метод | P | Q_0 | E_0 | a | S | U |
|------------|-------|--------|-----------------------|----------------------|-------|-------|
| Fitter 1 | 5.785 | 19.115 | 541.41 | 0.0390 | 2.717 | 0.000 |
| Fitter 2 | 5.561 | 19.113 | 540.16 | 0.0392 ^{*)} | 3.351 | 0.977 |
| Спряmlение | 5.570 | 19.113 | 540.10 ^{**)} | 0.0392 ^{*)} | 3.398 | 1.049 |

*) значение фиксировано; **) значение подобрано

Из таблицы мы снова видим, что Fitter дает наилучшие оценки и позволяет извлечь из данных дополнительную информацию.

10.3. Результаты главы 10

В этой главе было рассмотрено приложение методов нелинейного регрессионного анализа к задачам обработки данных титрования. Показано, что реализуемый программой Fitter вычислительный алгоритм позволяет облегчить проблему оценивания параметров, обеспечить высокую точность результатов вычислений, отбраковать выпадающие измерения. На его базе можно разработать компьютерный лабораторный практикум для студентов-аналитиков, ориентированный на решение обратных задач титрования, что является весьма актуальной формой обучения. При этом отпадает необходимость расчета кривых титрования с использованием приближенных уравнений, чему в общем курсе по аналитической химии уделяется значительное внимание

В области обратных задач титриметрии имеется одна актуальная, но еще практически неисследованная область - кислотно-основное титрование в смешанных водно-органических и неводных растворителях. Математические описания кривых титрования в таких средах выполнены рядом авторов, главным образом отечественных. Они весьма сложны и не поддаются линеаризации заменой переменных. Для таких задач можно было бы использовать систему Fitter. Успешное решение таких задач было бы значительным вкладом в теорию и практику титриметрического метода анализа

11. Хемилюминесцентный метод оценки эффективности ингибиторов

В этой главе рассматривается задача оценки параметров, характеризующих эффективность действия ингибиторов радикального цепного окисления хемилюминесцентным методом. Ингибиторы цепного радикального окисления представляют собой акцепторы свободных радикалов, ведущих процесс окисления. Одной из особенностей таких ингибиторов является то, что в процессе принимают участие не только молекулы самих ингибиторов, но и некоторые продукты их превращения, образующиеся в результате взаимодействия ингибитора с активными свободными радикалами. Радикал ингибитора, образующийся в акте взаимодействия ингибитора с активным радикалом - это лишь самый простой пример такого продукта. С макрокинетической точки зрения это означает, что одна молекула ингибитора может выводить из процесса более одного активного радикала. В следствие этого эффективность действия ингибиторов характеризуется двумя параметрами качественно разной природы:

- антирадикальной активностью молекулы ингибитора, эффективностью ее взаимодействия с активными радикалами и
- ресурсом ингибитора, количеством активных радикалов, выводимых из процесса каждой молекулой ингибитора, можно сказать стехиометрией процесса ингибирования.

Оба эти параметра могут быть количественно оценены по данным о кинетике изменения концентрации свободных радикалов в ходе расходования ингибитора, введенного в процесс инициированного окисления. Уменьшение концентрации свободных радикалов при введении в систему ингибитора характеризует антирадикальную активность ингибитора, а кинетика увеличения концентрации свободных радикалов по мере расходования ингибитора определяет стехиометрию процесса ингибирования, количество свободных радикалов, в среднем акцептируемых одной молекулой ингибитора и всеми продуктами, из нее образующимися.

Могут быть использованы разные методы регистрации кинетики концентрации активных радикалов в ходе ингибированного инициированного окисления: Можно, например, методом ЭПР непосредственно измерять кинетику концентрации перекисных радикалов, ведущих процесс инициированного окисления. Также, можно регистрировать кинетику поглощения кислорода в ходе этого процесса, мгновенная скорость цепного окисления пропорциональна мгновенной концентрации активных радикалов, ведущих процесс окисления. Наконец, можно измерять интенсивность хемилюминесценции, сопровождающей

процесс ингибированного инициированного окисления, которая обычно пропорциональна квадрату концентрации перекисных радикалов. Это позволяет получить информацию о кинетике изменения концентрации этих радикалов.

Все эти методы относительно широко используются для оценки эффективности действия ингибиторов цепного окисления. При этом обычно оценивается взаимосвязь некоторых особенностей получаемых кинетических кривых с количественными характеристиками эффективности ингибиторов. Определяется величина наклона кинетической кривой накопления радикалов или интенсивности хемилюминесценции в точке перегиба, время достижения точки перегиба, строятся различные нелинейные анаморфозы получаемых зависимостей и т.п. На всех этапах таких построений происходят разного рода потери информации, становится затруднительной более или менее строгая оценка точности получаемых результатов, проверка адекватности развиваемых представлений.

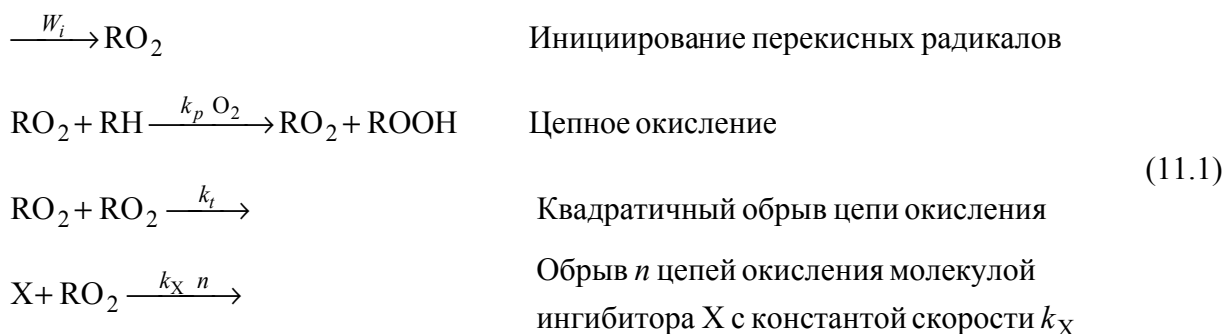
В этой главе, на примере обработки хемилюминесцентных измерений, демонстрируется строгий статистический подход к решению данной задачи, подход, который после принятия решения о кинетической схеме процесса по существу не требует никаких интеллектуальных усилий для оценки величины искомых параметров, характеризующих эффективность действия ингибиторов, для количественной характеристики точности сделанной оценки, для проверки адекватности развиваемых представлений используемой первичной информации.

Обработка таких экспериментальных данных вызывает определенные сложности, связанные, прежде всего, с неясностью, а также и с нелинейностью модели. Кроме того, часто бывает необходимо провести совместную обработку нескольких кинетических кривых, в которых априори может предполагаться наличие (или отсутствие) общих параметров. Применение системы Fitter позволяет легко решить эту задачу. При этом пользователь может легко модифицировать модель в зависимости от того, какие параметры системы являются общими, а какие частными. Однако в повседневной практической деятельности возникает необходимость полностью автоматизировать процесс построения модели с тем, чтобы перепоручить анализ данных персоналу, имеющему слабые навыки работы с компьютером. Это можно сделать, используя встроенные функции системы Fitter, описанные в разделе 5.12. В результате получается простой (с точки зрения пользователя) шаблон, в который заносятся (в ручную или автоматически) экспериментальные данные. Затем пользователь устанавливает несколько простых настроек и сразу же получает требуемый результат

В этой главе мы покажем, как это делается, рассмотрев основные детали построения такого шаблона, названного *ChemLum*. Подробности можно найти в [126], где приведен файл *CHEMLUM.XLS*, содержащий этот шаблон.

11.1. Модель

В качестве базовой модели описания процесса ингибированного иницированного окисления используется модель цепного окисления с квадратичным обрывом цепи, в которой присутствует ингибитор, обрывающий n цепей окисления. Кинетическая схема такого процесса может быть сведена к четырем актам химического превращения:



Этой кинетической схеме соответствует система обыкновенных дифференциальных уравнений, описывающих кинетику изменения концентрации перекисных радикалов, $[\text{RO}_2]$, в ходе расходования ингибитора.

$$\begin{aligned}
 \frac{d[\text{X}]}{dt} &= -k_X[\text{RO}_2][\text{X}] \\
 \frac{d[\text{RO}_2]}{dt} &= W_i - 2k_t[\text{RO}_2]^2 - nk_X[\text{RO}_2][\text{X}]
 \end{aligned} \tag{11.2}$$

В ходе процесса концентрация перекисных радикалов квазистационарна. Это означает, что в системе уравнений (11.2) правую часть второго уравнения можно приравнять нулю. Это можно делать в том случае, если за время жизни перекисного радикала можно пренебречь изменением концентрации ингибитора. Формально это означает, что концентрация перекисных радикалов много меньше концентрации ингибитора и, соответственно, скорость ее изменения много меньше скорости иницирования.

$$\begin{aligned}
 [\text{RO}_2] &\ll [\text{X}] \\
 \frac{d[\text{RO}_2]}{dt} &\ll W_i
 \end{aligned} \tag{11.3}$$

При выполнении условий (11.3) система уравнений (11.2) может быть приведена к виду

$$\begin{cases} \frac{d[X]}{dt} = - \left[1 - \left(\frac{[RO_2]}{[RO_2]_0} \right)^2 \right] \frac{W_i}{n} \\ W_i = 2k_t[RO_2]^2 + nk_X[RO_2][X] \end{cases} \quad (11.4)$$

где $[RO_2]_0 = \sqrt{W_i/k_t}$.

Акты, составляющие кинетическую схему (11.1), не вполне элементарны. Акту образования перекисного радикала окисляемого вещества предшествуют акты образования радикалов инициатора и его превращения в перекисный радикал окисляемого вещества. В актах цепного окисления участвует молекула кислорода, которая достаточно быстро превращает первичный алкильный радикал окисляемого вещества в перекисный. В акте квадратичного обрыва цепи кислород, наоборот, частично выделяется. Обрыв n цепей окисления молекулой ингибитора представляет собой сложную последовательность реакций радикала окисляемого вещества с продуктами превращения ингибитора.

Для того, чтобы кинетическую схему (11.1) можно было свести к (11.4) требуется выполнение ряда условий:

- радикалы инициатора должны полностью превращаться в радикалы окисляемого вещества, они должны практически не успевать взаимодействовать между собой и не участвовать в актах обрыва цепи;
- концентрация кислорода должна быть достаточно велика для того, чтобы алкильные свободные радикалы окисляемого вещества (образующиеся в акте взаимодействия перекисного радикала с окисляемым веществом) также не успевали бы взаимодействовать между собой и не участвовать в актах обрыва цепи;
- продукты взаимодействия перекисных радикалов с ингибитором должны обладать достаточно высокой способностью взаимодействовать с перекисными радикалами, только в этом случае сложный механизм образования этих продуктов и их взаимодействия с перекисными радикалами можно свести к формальному изменению стехиометрического параметра n .

В большинстве модельных систем, используемых для определения антирадикальной активности ингибиторов эти условия выполняются. Рассматриваемый в настоящей работе вариант программного обеспечения приспособлен для решения системы уравнений (11.4), однако, использование системы Fitter позволяет также обрабатывать данные, описываемые и более сложной кинетической схемой.

В эксперименте измеряется интенсивность хемилюминесценции $I(t)$ при разных начальных значениях концентрации ингибитора X_0 и при одной постоянной скорости инициирования W_i . Предполагая, что интенсивность хемилюминесценции I пропорциональна квадрату концентрации перекисных радикалов, получаем следующую систему

$$\begin{cases} \frac{dX}{dt} = -\left(I - \frac{I}{I_l}\right) \frac{W_i}{n}, & X(0) = X_0 \\ \left(I - \frac{I}{I_l}\right) \frac{W_i}{n} = KX \sqrt{\frac{I}{I_l} W_i} \end{cases} \quad (11.5)$$

моделирующую измеряемый сигнал I (отклик). В этой модели участвуют следующие известные величины (предикторы): t – время, X_0 – начальная концентрация ингибитора, W_i – скорость инициирования, а также неизвестные значения (параметры): K – эффективная кинетическая константа,

$$K = \frac{k_X}{\sqrt{k_t}}$$

n – число активных радикалов, уничтожаемых одной частицей ингибитора, I_l – предельная величина сигнала. Кроме того, в модели присутствует промежуточная величина $X(t)$ – текущая концентрация ингибитора.

Эффективность ингибитора характеризуется параметром n , а также эффективной константой K .

Решение системы (11.5) можно выразить как неявную функцию относительно сигнала I .

$$\frac{I - \sqrt{\frac{I}{I_l}}}{I + \sqrt{\frac{I}{I_l}}} - \frac{Z - I}{a} \exp\left(-\sqrt{\frac{I}{I_l}} - Kt + \frac{I}{Z - a}\right) = 0 \quad (11.6)$$

где

$$a = MKn \frac{X_0}{W_i}, \quad Z = \sqrt{a^2 + 1}, \quad M = 120$$

Обращает на себя внимание то, что форма кинетической кривой в таком процессе явно не зависит ни от начальной концентрации ингибитора X_0 , ни от числа активных радикалов, уничтожаемых одной частицей ингибитора, n . Параметр, определяющий активность ингибитора – K , лишь трансформирует эту универсальную зависимость, сжимает ее вдоль оси времени, а начальная концентрация ингибитора, X_0 , и параметр, определяющий эффек-

тивность его действия – n , сдвигают ее вдоль той же оси. При обработке только одной кинетической кривой эту модель можно представить в системе Fitter в следующем виде

```
'Модель хемиллюминесценции
0=(1-(I/I1)^0.5)/(1+(I/I1)^0.5)-(Z-1)/a*exp[-(I1/I)^0.5-K*t+1/(Z-a)]; Imin<I<Imax
Z=(a*a+1)^0.5
a=M*K*n*(X0/Wi)
M=1/120
Imin=1
Imax=100
K=?
n=?
I1=?
```

Рис. 11.1 Модель для описания хемиллюминесцентных данных

Однако, как правило, необходимо совместно обрабатывать несколько серий экспериментальных данных, полученных при разных значениях начальной концентрации ингибитора (см. Рис. 11.2). При этом из трех неизвестных параметров модели – K , n и I_1 , присутствующих в модели, один или два могут быть общими для всех серий.

Табл. 11.1 Варианты обработки k серий данных с общими параметрами моделью (11.6)

| Вариант обработки | Неизвестные параметры | | | Число параметров |
|-------------------------------|-----------------------|-------------------|-------------------|------------------|
| 1 Все параметры разные | K_1, \dots, K_k | n_1, \dots, n_k | I_1, \dots, I_k | $3k$ |
| 2 Параметры K – общие | K | n_1, \dots, n_k | I_1, \dots, I_k | $2k+1$ |
| 3 Параметры n – общие | K_1, \dots, K_k | n | I_1, \dots, I_k | $2k+1$ |
| 4 Параметры K и n – общие | K | n | I_1, \dots, I_k | $k+2$ |

В Табл. 11.1 показаны все четыре варианта оценивания неизвестных параметров при совместной обработке k серий данных с учетом возможной одинаковости первых двух параметров в разных сериях. Последний параметр, отвечающий за предельное значение люминесцентного сигнала, всегда является частным для каждой серии.

11.2. Устройство шаблона

При построении шаблона *ChemLum* мы исходили из следующих соображений. Во-первых, все промежуточные величины должны быть надежно спрятаны от пользователя. Во вторых, интерфейс должен быть максимально прост и очевиден. В идеале процедура обработки должна быть организована так, чтобы ей можно было управлять, не вводя ни одного символа с клавиатуры, а используя только компьютерную мышь.

В шаблоне имеется пять листов. Четыре из них: «Описание», «Инструкция», «Данные» и «Результаты» – это видимые листы программы Excel, а пятый, названный странным именем – «FitterHiddenSheet», является невидимым – “xlSheetVeryHidden”. Именно на

этом, последнем листе и располагается вся скрытая от пользователя, рабочая информация с которой оперирует система Fitter.

На листе «*Описание*» располагается текст, который примерно соответствует предыдущему разделу 11.1. Второй лист «*Инструкция*» содержит объяснения того, что можно и чего нельзя делать с этим шаблоном, что примерно соответствует данному разделу. Эти два листа защищены паролем; они не могут быть изменены пользователем.

Лист «*Данные*» предназначен для ввода экспериментальных данных. Если прибор не соединен к компьютеру, то эта процедура выполняется оператором вручную. На этом листе имеется заранее заготовленная таблица, фрагмент которой (верхний левый угол) показан на Рис. 11.2.

| № | 1 | | 2 | | 3 | | 4 | | 5 | |
|-------|-----------|-------|-----------|-------|-----------|-------|-----------|-------|-----------|-----|
| | t (min) | I | t (min) | I | t (min) | I | t (min) | I | t (min) | I |
| W_i | 0.0000002 | | 0.0000002 | | 0.0000002 | | 0.0000002 | | | |
| X_0 | 0.00001 | | 0.0000225 | | 0.00007 | | 0.0002 | | | |
| 1 | 0 | 0.49 | 0 | 0.195 | 0 | 0.12 | 0 | 0.025 | | |
| 2 | 1 | 0.61 | 2 | 0.325 | 5 | 0.18 | 20 | 0.05 | | |
| 3 | 2 | 0.735 | 4 | 0.465 | 10 | 0.31 | 30 | 0.1 | | |
| 4 | 3 | 0.85 | 6 | 0.625 | 13 | 0.46 | 35 | 0.15 | | |
| 5 | 4 | 0.96 | 7 | 0.73 | 15 | 0.6 | 40 | 0.25 | | |
| 6 | 5 | 1.06 | 8 | 0.81 | 18 | 0.78 | 45 | 0.48 | | |
| 7 | 6 | 1.15 | 9 | 0.92 | 20 | 0.98 | 47 | 0.63 | | |
| 8 | 7 | 1.23 | 10 | 1.03 | 22 | 1.165 | 49 | 0.78 | | |
| 9 | 8 | 1.3 | 11 | 1.13 | 25 | 1.4 | 51 | 0.93 | | |
| 10 | 9 | 1.35 | 12 | 1.23 | 30 | 1.56 | 54 | 1.1 | | |
| 11 | 10 | 1.4 | 13 | 1.31 | 35 | 1.64 | 57 | 1.23 | | |
| 12 | 12 | 1.48 | 14 | 1.38 | 40 | 1.65 | 62 | 1.32 | | |
| 13 | 14 | 1.515 | 16 | 1.485 | 43 | 1.68 | 67 | 1.39 | | |
| 14 | 16 | 1.53 | 18 | 1.56 | | | 72 | 1.42 | | |
| 15 | 20 | 1.57 | 20 | 1.59 | | | 77 | 1.43 | | |
| 16 | 25 | 1.58 | 22 | 1.63 | | | 92 | 1.47 | | |
| 17 | 29 | 1.59 | 24 | 1.63 | | | | | | |
| 18 | | | | | | | | | | |
| 19 | | | | | | | | | | |

Рис. 11.2 Таблица для ввода данных. Показаны три серии данных.

Каждая серия данных (кинетика) вводится в два столбца: время t (в минутах) и рядом соответствующая интенсивность I . В два верхних ряда над этой парой колонок вводятся соответствующие значения скорости инициирования W_i и начальной концентрации ингибитора X_0 . Следующая серия вводится рядом и т.д. Желательно, но необязательно, располагать время в порядке возрастания, иначе искажается форма расчетных кривых. Пробелы в перечне результатов измерения не допускаются, т.к. после пробела все следующие за ним данные не учитываются при расчете. Если какое-то значение (время или интенсивность)

написано перечеркнутым шрифтом, то эта точка игнорируется при обработке – соответствующий вес становится равным нулю.

Разумеется, такое устройство таблицы данных отличается от стандарта, предусмотренного в системе Fitter (см. Рис. 5.1). Поэтому в шаблоне предусмотрено, что данные из таблицы с листа «Данные» автоматически переписываются на скрытый лист «FitterHiddenSheet», в обычном формате Fitter, так, как это показано на Рис. 11.4.

На листе «Результаты» собраны как элементы, управляющие обработкой, так и элементы, представляющие пользователю полученные результаты. Вид этого листа показан на Рис. 11.3

Рассмотрим элементы управления, собранные в таблицу, озаглавленную **Блок управления**.

Ошибка измерения. Эта настройка задает тип погрешности измерения и соответствует настройке **Error** в первом окне Регистратора Настроек (Рис. 5.9). Значение по умолчанию «НЕТ» (не отмечено) означает абсолютную погрешность (1.11). Если настройка отмечена, то ошибка измерений считается относительной (1.12).

Проверка гипотез. С помощью этой настройки устанавливается уровень значимости при проверке гипотез (см. раздел 1.4), что соответствует настройке α во втором окне Регистратора Настроек (Рис. 5.11). Настройке «Строго» соответствует значимость $\alpha = 0.05$, «Средне» – $\alpha = 0.0225$ и «Мягко» – $\alpha = 0.001$. Проверяются две гипотезы – *тест выбросов* и *тест серий*.

Общие параметры. Под этим заголовком расположены две настройки: Общие К и Общие n, действие которых соответствует схеме установки параметров, приведенной в Табл. 10.3.

Оценить. Эта кнопка запускает процедуру обработки данных.

Запомнить стиль графика. С помощью этой кнопки можно запомнить текущие настройки, используемые при построении графика.

В таблицу, озаглавленную **Общие результаты**, выводится информация, полученная по результатам оценивания данных. В клетку **Ошибка**, помещается оценка дисперсии ошибки измерения, а в клетки **Выброс** и **Серии** выводятся результаты проверки соответствующих гипотез. Это могут быть следующие значения: «Есть», «Нет» и «Не знаю».

Хемиллюминесцентный метод оценки антирадикальной активности ингибиторов

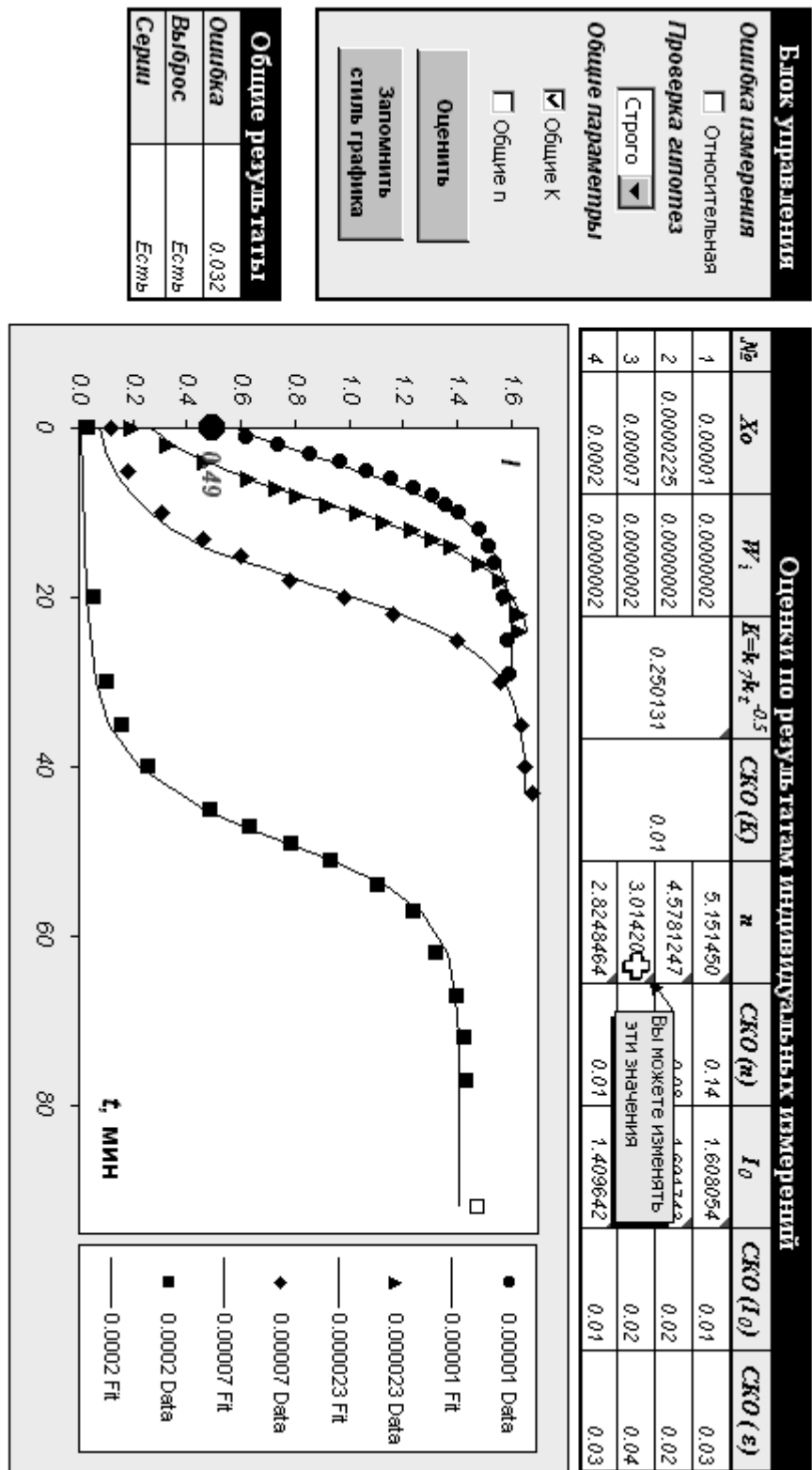


Рис. 11.3 Внешний вид страницы «Результаты» в шаблоне ChemLum.

Результаты обработки выводятся также в таблицу, озаглавленную **Оценки по результатам индивидуальных измерений**. Три первые столбца в этой таблице представляют номер серии и соответствующие значения предикторов X_0 и W_i . Остальные колонки содержат оценки параметров K , n и I_l и их стандартные отклонения – СКО(K), СКО(n) и СКО(I_l). Если какой-то параметр является общим для всех серий (на Рис. 11.3 это – K), то он представлен единой оценкой и отклонением. В последнем столбце, озаглавленном СКО(ϵ), приводятся значения среднеквадратичной ошибки измерения отдельно для каждой серии. Все величины неизвестных параметров, указанные в этой таблице, используются как начальные значения при поиске оценок, поэтому они могут быть изменены, с целью улучшить сходимость поиска. Все другие клетки на этом листе защищены от изменения пользователем.

На **Графике** точки представляют экспериментальные, а кривые – модельные значения. Если имеется выброс, то он отмечается увеличенным значком с указанием значения. Все точки, исключенные из поиска, показаны незакрашенными значками. График доступен для редактирования и пользователь может установить вид графика, который ему больше подходит. Запомнить эти изменения можно с помощью кнопки **Запомнить стиль графика**.

11.3. Программирование шаблона

При нажатии кнопки **Оценить** на листе «*Результаты*», происходит вызов процедуры (макроса) `TestInhibitor`, написанной на языке *Visual Basic for Application* (VBA) [7]. В этой процедуре используются *функции Fitter*, описанные в разделе 5.12. Применение этих функций позволяет осуществить все операции по обработке данных в скрытом от пользователя режиме. Приведем тексты некоторых модулей, которые помогут понять, как программируется работа `Fitter` в скрытом режиме.

Начнем с объявления глобальных переменных.

```
Option Explicit
Option Base 1
Dim dI0Values() As Double           'Массив значений I0
Dim dX0Values() As Double           'Массив значений X0
Dim dWiValues() As Double           'Массив значений W
Dim nCLSeries As Integer             'Число серий
Dim nCLSeriesSize() As Integer       'Массив размерностей серий
Dim nCLErrorType As Integer          'Тип ошибки
Dim dSign As Double                  'Величина значимости
Dim rFitterData As Range              'Область данных на скрытом листе
Dim sCLData As String                'Адрес области данных на скрытом листе
Dim sCLParams As String              'Адрес параметров на скрытом листе
Dim sCLModel As String               'Имя текстового поля модели
```


Хемилюминесцентный метод оценки эффективности ингибиторов

```
Dim bOneN As Boolean           'Общие ли n?
Dim bOneK As Boolean           'Общие ли K?
Dim bOneNOld As Boolean        'Были ли n общими раньше?
Dim bOneKOld As Boolean        'Были ли K общими раньше?
Dim wsUserData As Worksheet   'Лист «Данные»
Dim wsResults As Worksheet    'Лист «Результаты»
Dim wsFitterData As Worksheet 'Скрытый лист
Dim vCalcStatus As Variant    'Статус вычислений
Dim chGraph As ChartObject    'График на листе «Результаты»
```

Эти переменные используются во всех процедурах шаблона.

Приведем текст главной процедуры.

```
Sub TestInhibitor()

    OpenSheets           'Открытие защищенных листов
    SetOptions           'Установка типа ошибки и значимости (см. ниже)
    CollectData          'Перенос данных с листа «Данные» на скрытый лист
    CreatModelBox        'Создание текстового поля модели (см. ниже)
    InitParameters       'Установка нач. значений параметров

    'Инициализация
    If (FitterInit() = 0) Then GoTo ErrorEnd

    'Установка настроек
    If (FitterOptions( _
        dCalcPrecision:=0.00000000000001, kErrorType:=nCLErrorType, _
        kDeviations:=xtYes, kFitResults:=xtYes, _
        dSignificance:=dSign, kTestSeries:=xtYes, kTestOutliers:=xtYes _
    ) = 0) Then GoTo ErrorEnd

    'Регистрация данных
    If (FitterData(sDataRange:=sCLData, sCodeString:="PPPRWF", _
        kTitle:=xtNo) = 0) Then GoTo ErrorEnd

    'Регистрация модели
    If (FitterModel(sModelTextBox:=sCLModel) = 0) Then GoTo ErrorEnd

    'Регистрация параметров
    If (FitterParameters(sParametersLeftTop:=sCLParams, kTitle:=xtNo) = 0) _
        Then GoTo ErrorEnd

    'Построение графика
    If (MakePlot() = 0) Then GoTo ErrorEnd 'Make plot

    Application.ScreenUpdating = True
    Application.Calculation = xlAutomatic
    Application.DisplayAlerts = False

    'Обработка данных
    If (FitterFit(dFitConvergence:=0.00001, nMaxIterations:=100) = 0) _
        Then GoTo ErrorEnd

    'Вывод результатов (см. ниже)
    Application.DisplayAlerts = True
    If (OutputResults() = 0) Then GoTo ErrorEnd 'Output results
    CloseSheets           'Закрытие защищенных листов

Exit Sub
'Обработка ошибки
ErrorEnd:
    FitterError
```

```

CloseSheets           'Заккрытие защищенных листов

End Sub

В ней участвуют как функции Fitter (подчеркнуты), так и специально написанные проце-
дуры. Приведем для примера текст процедуры SetOptions, которая считывает установки
с листа «Результаты» и устанавливает соответствующие значения переменных.

Sub SetOptions()

'Установка значимости
  Select Case wsResults.DropDowns("ddSign").ListIndex
    Case 1
      dSign = 0.05
    Case 2
      dSign = 0.0225
    Case 3
      dSign = 0.001
  End Select

'Установка типа ошибки
  nCLErrorType = xtAbsolute           'Ошибка - абсолютная

  If (wsResults.CheckBoxes("cbError") = xlOn) Then _
    nCLErrorType = xtRelative         'Ошибка - относительная

'Установка признаков общности параметров K и n
  bOneNOld = bOneN
  bOneKOld = bOneK

  bOneN = False                       'Все n - частные
  If (wsResults.CheckBoxes("cbCommonN").Value = xlOn) Then _
    bOneN = True                       'Все n - общие

  bOneK = False                       'Все K - частные
  If (wsResults.CheckBoxes("cbCommonK").Value = xlOn) Then_
    bOneK = True                       'Все K - общие

End Sub

```

Интересно проследить, как в процедуре CreatModelBox автоматически создается тексто-
вое поле модели. Далее мы приведем ее текст, но сначала дадим некоторые пояснения.
Модель (11.6) является неявной функцией относительно отклика I , поэтому, в соответ-
ствие с синтаксисом (см. раздел 5.3), необходимо задать граничные условия для I (сравни
Рис. 11.1). Для этого на скрытом листе (см. Рис. 11.4) отводятся две клетки (L1 и L2) с
именами rI_{max} и rI_{min} , в которые помещаются формулы «=MAX(C2:C16384)*1.02»
и «=rI_{max}*0.1» соответственно, вычисляющие необходимые границы по имеющимся
экспериментальным значениям отклика. Эти клетки используются в модели для определе-
ния динамических констант

`Imin="FitterHiddenSheet!rImin"`

`Imax="FitterHiddenSheet!rImax"`

, являющихся нужными граничными условиями.

Как правило, одновременно обрабатываются несколько серий, имеющие как общие, так и частные параметры. Для ввода частных параметров в модель используется конструкции с импульсной функцией (5.3) $\text{imp}(x_0-x_1)$, где x_0 – это имя предиктора (начальная концентрация ингибитора), а x_1 – это константа, значение которой равно начальной концентрации в первой серии. В этом случае импульсная функция будет равна единице только в первой серии и нулю во всех остальных сериях. Константа x_1 и все аналогичные ей константы x_2, x_3, \dots (в количестве числа серий – переменная `nCLSeries`), определяются в модели динамически через ссылки на клетки `J1, J2, \dots` на скрытой странице, в которые эти значения заносятся при создании таблицы данных (процедура `CollectData`). Поэтому в модели должны присутствовать строчки типа

`x1="FitterHiddenSheet!J1"`

в количестве равном числу серий. Аналогично обстоит дело с неизвестными параметрами I_1, I_2, \dots , отвечающими за предельное значение интенсивности. Все они являются частными, потому в модели должно быть такое же число строчек типа

`I1=?`

Вид модели зависит от того, какие параметры являются общими, а какие частными (см. Табл. 10.3). Соответствующие признаки находятся в переменных `bOneN` и `bOneK`. В зависимости от их значений (`True` или `False`) вид модели изменяется. Так, если параметры K являются общими (`bOneK=True`) то, в модели должна быть одна строка

`K=?`

а если они частные (`bOneK=False`) то несколько строк вида

`K=K1*imp(x0-x1)+K2*imp(x0-x2)+...`

`K1=?`

`K2=?`

`...`

Аналогично меняется модель в зависимости от вида параметров n . Если они общие (`bOneN=True`), тогда в модели должна быть одна строка

`n1=?`

и еще строка для промежуточной переменной a

`a=M*K*n1*(x1/Wi*imp(x0-x1)+x2/Wi*imp(x0-x2)+...)`

Если же они не общие (`bOneN=False`), то должно быть несколько строчек типа

n1=?
n2=?
...

а строка для переменной a должна иметь вид

$$a = M * K * (X1 / Wi * 1 * \text{imp}(X0 - X1) + X2 / Wi * n2 * \text{imp}(X0 - X2) + \dots)$$

Теперь, после этих пояснений мы можем привести текст процедуры CreatModelBox, создающей текстовое поле модели на скрытой странице.

```
Sub CreatModelBox()

    Dim sCLHiddenSheet As String          'Имя скрытого листа
    Dim sTemp As String, sBox As String, i As Integer

    sCLHiddenSheet = wsFitterData.Name

    'Удалить старое текстовое поле и создать новое пустое
    On Error Resume Next
    wsFitterData.TextBoxes("CL_Model").Delete
    On Error GoTo 0
    wsFitterData.TextBoxes.Add Top:=10, Left:=300, Width:=200, Height:=100
    wsFitterData.TextBoxes(1).Name = "CL_Model"

    'Создать пять первых строк модели, которые неизменны
    sBox = "M=1/120" & Chr(10) & _
    "Z=(a*a+1)^0.5" & Chr(10) & _
    "Imin="" & sCLHiddenSheet & "!rImin"" & Chr(10) & _
    "Imax="" & sCLHiddenSheet & "!rImax"" & Chr(10) & _
    "0=(1-(I/Io)^0.5)/(1+(I/Io)^0.5)-(Z-1)/a*exp[-(Io/I)^0.5-K*t+1/(Z-a)]" & _
    & ";Imin<I<Imax" & Chr(10)

    'Создать строчки типа X1=...
    For i = 1 To nCLSeries
        sTemp = "X" & i & "=" & "" & sCLHiddenSheet & _
        "!J" & i & "" & Chr(10)
        sBox = sBox & sTemp
    Next i

    'Создать строчки типа I1=?
    For i = 1 To nCLSeries
        sTemp = "I" & i & "=?" & Chr(10)
        sBox = sBox & sTemp
    Next i

    'Создать строку вида Io=I1*imp(X0-X1)+I2*imp(X0-X2)+...
    sTemp = "Io=I1*imp(X0-X1)"
    For i = 2 To nCLSeries
        sTemp = sTemp & "+I" & i & "*imp(X0-X" & i & ")"
    Next i
    sTemp = sTemp & Chr(10)
    sBox = sBox & sTemp

    'Создать строку (строки) вида K=...
    If (bOneK) Then
        sTemp = "K=?" & Chr(10)          'Один общий K
    Else
        sTemp = "K=K1*imp(X0-X1)"      'Несколько частных K
    End If
End Sub
```

```

For i = 2 To nCLSeries
    sTemp = sTemp & "+K" & i & "*imp(X0-X" & i & ")"
Next i
sTemp = sTemp & Chr(10)

For i = 1 To nCLSeries
    sTemp = sTemp & "K" & i & "=?" & Chr(10)
Next i
End If
sBox = sBox & sTemp

' Создать строку (строки) вида n=...
If (bOneN) Then
    sTemp = "a=M*K*n1*(X1/Wi*imp(X0-X1)"           ' Один общий n
Else
    sTemp = "a=M*K*(X1/Wi*n1*imp(X0-X1)"           ' Несколько частных n
End If
For i = 2 To nCLSeries
    If (bOneN) Then
        sTemp = sTemp & "+X" & i & "/Wi*imp(X0-X" & i & ")"
    Else
        sTemp = sTemp & "+X" & i & "/Wi*n" & i & "*imp(X0-X" & i & ")"
    End If
Next i
sTemp = sTemp & ")" & Chr(10)
sBox = sBox & sTemp
If (bOneN) Then
    sBox = sBox & "n1=?" & Chr(10)
Else
    For i = 1 To nCLSeries
        sTemp = "n" & i & "=?" & Chr(10)
        sBox = sBox & sTemp
    Next i
End If

' Занести строку sBox в текстовое поле
FillTextBox wsFitterData.TextBoxes("CL_Model"), sBox
' Присвоить имена полю модели и таблице параметров
sCLModel = sCLHiddenSheet & "!CL_Model"
sCLParams = sCLHiddenSheet & "!$G$1"
End Sub

```

Результат работы этой процедуры (текстовое поле) для случая двух серий данных при условии, что параметр K – общий, а параметры $n_1, n_2 \dots$ – частные, представлен на Рис. 11.4.

При регистрации этой модели, таблица параметров устанавливается ссылкой на клетку G1 скрытого листа, где собираются все необходимые имена и начальные значения параметров модели.

| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|----|--|-------|----|-------|---|-------|----|----------|----|----------|--------------------|-------|---|
| 1 | X0 | Wi | t | I | w | f | I1 | 1.568115 | X1 | 0.00001 | I _{max} = | 29.58 | |
| 2 | 0.00001 | 2E-07 | 0 | 0.49 | 0 | 0.564 | I2 | 1.685793 | X2 | 2.25E-05 | I _{min} = | 2.958 | |
| 3 | 0.00001 | 2E-07 | 1 | 0.61 | 1 | 0.659 | K | 0.30577 | | | | | |
| 4 | 0.00001 | 2E-07 | 2 | 0.735 | 1 | 0.762 | n1 | 4.591721 | | | | | |
| 5 | 0.00001 | 2E-07 | 3 | 0.85 | 1 | 0.867 | n2 | 4.666331 | | | | | |
| 6 | 0.00001 | 2E-07 | 4 | 0.96 | 1 | 0.971 | | | | | | | |
| 7 | 0.00001 | 2E-07 | 5 | 1.06 | 1 | 1.069 | | | | | | | |
| 8 | 0.00001 | 2E-07 | 6 | 1.15 | 1 | 1.159 | | | | | | | |
| 9 | <pre> M=1/120 Z=(a*a+1)^0.5 Imin="FitterHiddenSheet!rImin" Imax="FitterHiddenSheet!rImax" 14 O=(1-(I/Io)^0.5)/(1+(I/Io)^0.5)-(Z-1)/a*exp[-(I/Io)^0.5-K*t+1/(Z-a)];Imin<I<Imax 15 X1="FitterHiddenSheet!J1" 16 X2="FitterHiddenSheet!J2" 17 I1=? 18 I2=? 19 Io=I1*imp(X0-X1)+I2*imp(X0-X2) 20 K=? 21 a=M*K*(X1/Wi*n1*imp(X0-X1)+X2/Wi*n2*imp(X0-X2)) 22 n1=? 23 n2=? 24 25 2.3E-05 2E-07 9 0.92 1 0.919 26 2.3E-05 2E-07 10 1.03 1 1.039 </pre> | | | | | | | | | | | | |
| 10 | | | | | | | | | | | | | |
| 11 | | | | | | | | | | | | | |
| 12 | | | | | | | | | | | | | |
| 13 | | | | | | | | | | | | | |
| 14 | | | | | | | | | | | | | |
| 15 | | | | | | | | | | | | | |
| 16 | | | | | | | | | | | | | |
| 17 | | | | | | | | | | | | | |
| 18 | | | | | | | | | | | | | |
| 19 | | | | | | | | | | | | | |
| 20 | | | | | | | | | | | | | |
| 21 | | | | | | | | | | | | | |
| 22 | | | | | | | | | | | | | |
| 23 | | | | | | | | | | | | | |
| 24 | | | | | | | | | | | | | |
| 25 | 2.3E-05 | 2E-07 | 9 | 0.92 | 1 | 0.919 | | | | | | | |
| 26 | 2.3E-05 | 2E-07 | 10 | 1.03 | 1 | 1.039 | | | | | | | |

Рис. 11.4 Вид скрытого листа *FitterHiddenSheet* шаблона *ChemLum*. Показано текстовое поле с моделью.

Рассмотрим теперь вопрос о том, как получить результаты обработки. Как видно, из текста процедуры `TestInhibitor`, приведенного выше, эти результаты не выводятся на лист. Аргумент `sResultsLeftTop` функции `FitterFit` – не задан, а по умолчанию он устанавливается пустой строкой. Тем не менее, они находятся в памяти и их можно извлечь оттуда с помощью функций `Fitter`. То, как это делается, ясно из приведенной функции `GetResults`.

```

Function GetResults() As Integer
    Dim nPar As Integer          ' Число параметров
    Dim nObs As Integer         ' Число измерений
    Dim dSS As Double           ' Сумма квадратов
    Dim dVft As Double          ' Дисперсия ошибки
    Dim dEst() As Double        ' Массив оценок параметров
    Dim dDev() As Double        ' Массив СКО оценок параметров
    Dim dRes() As Double        ' Массив откликов
    Dim dFit() As Double        ' Массив модельных значений
    Dim dWgh() As Double        ' Массив весов
    Dim nOut As Integer         ' Результат теста выбросов
    Dim nSer As Integer         ' Результат теста серий

    GetResults = 0

    ' Получить размерности массивов
    If (FitterDimensionsBack(Observations:=nObs, Parameters:=nPar) = 0) _
        Then Exit Function

    ' Установить размеры всех массивов
    ReDim dEst(1 To nPar)

```

```
ReDim dDev(1 To nPar)
ReDim dRes(1 To nObs)
ReDim dFit(1 To nObs)
ReDim dWgh(1 To nObs)

' Получить значения требуемых статистик
If (FitterStatisticsBack(SumSquares:=dSS, VarianceFit:=dVft, _
    Estimates:=dEst, Deviations:=dDev) = 0) Then Exit Function

' Получить значения данных и подгонки
If (FitterValuesBack(Responses:=dRes, Fittings:=dFit, _
    Weights:=dWgh) = 0) Then Exit Function

' Получить результаты проверки тестов
If (FitterTestsBack(Outliers:=nOut, Series:=nSer) = 0) _
    Then Exit Function

    GetResults = 1
End Function
```

В этой функции используются несколько функций Fitter (Подчеркнуты). После выполнения этой процедуры остается только рассчитать нужные показатели и вывести их на лист «Результаты».

11.4. Результаты главы 11

Пример, разобранный в этой главе, показывает, что система Fitter может применяться не только в качестве удобного интерфейса для исследовательского анализа данных, но и как скрытый «мотор» при рутинной обработке стандартных массивов по фиксированной модели. С ее помощью можно создавать стандартные шаблоны, которые может использовать неискушенный в проблематике оператор. Существует также принципиальная возможность (пока, к сожалению, не реализованная) организовать прямую передачу данных от прибора в компьютер, с последующей их автоматической обработкой. Такая полностью автоматизированная процедура могла бы быть полезна при решении задач мониторинга и контроля различных непрерывных процессов.

Заключение

В работе были рассмотрены теоретические, алгоритмические и методологические аспекты математического моделирования сложных физико-химических процессов. Изложенный материал позволяет сделать следующие **выводы**.

Регрессионный анализ по-прежнему является основным инструментом для нелинейного моделирования и прогнозирования поведения сложных систем. Однако широкое использование этого метода затруднено наличием нескольких плохо исследованных вопросов. В частности имеется три главные проблемы. Во-первых, это теоретические проблемы – одновременная обработка разнородных экспериментальных данных больших размеров с использованием различных моделей, имеющих общие и частные параметры и построение доверительных интервалов при прогнозировании. Помимо теоретических задач, имеется и практическая проблема – создание современной компьютерной системы для интерактивной обработки данных и подбора описаний. Последний вопрос, рассматриваемый в работе, это разработка методологии применения методов НЛРА.

В теоретической части работы были получены следующие результаты.

1. Разработан новый метод оценивания параметров регрессионных моделей – *последовательное байесовское оценивание* (ПБО). Доказано, что оценки, построенные этим способом – ПБО–оценки – обладают следующими свойствами. Для линейных моделей они совпадают с оценками метода максимума правдоподобия (ММП), а для нелинейных моделей асимптотически сходятся к ММП–оценкам.

2. Предложен новый способ построения доверительных интервалов для прогноза по нелинейной регрессии, названный “*связанное моделирование*”. Показано, что при той же точности этот метод работает примерно в 1000 раз быстрее и дает хорошие результаты даже для сильно нелинейных моделей.

3. Дано новое определение *коэффициента нелинейности* γ , который характеризует внутреннюю нелинейность регрессионной задачи. Показано, что коэффициент нелинейности γ зависит как от вида модели, так и от плана эксперимента.

Получены следующие практические результаты.

4. Создан *алгоритм*, позволяющий интерпретировать и компилировать модели, задаваемые пользователем в текстовом виде в естественной математической форме записи – как систему явных, неявных и дифференциальных уравнений.

5. Для точного вычисления производных модели по параметрам был разработан другой алгоритм, который автоматически находит эти *производные* в аналитической форме.

6. Создана программа для регрессионного анализа сложных данных – *компьютерная системы Fitter* – надстройка для программы Excel. Эффективность работы системы Fitter была проверена с помощью стандартного, общепризнанного набора тестовых данных, разработанного Национальным институтом стандартов США (NIST).

Создана методология применения методов НЛРА для математического моделирования сложных физико-химических процессов, которую можно представить следующим набором практических рекомендаций.

7. Сложность модели определяется не ее внешним видом, а комплексом свойств, зависящих от многих, подчас сложных обстоятельств. Трудная для вычисления модель может оказаться простой в оценивании, и наоборот. (см. раздел 6.1)

8. Использование упрощенных методов для обработки данных: линеаризация, анаморфозы и т.п., не позволяет оценить все требуемые параметры, не дает нужной точности и может приводить к грубым ошибкам. (см. главу 10)

9. Метод ПБО является корректной процедурой, упрощающей моделирование и позволяющей достичь тех же результатов, которые дали бы традиционные подходы – МНК и ММП. В некоторых случаях ПБО позволяет получить результат, который невозможно достичь другими методами. (см. главы 6, 7 и 8)

10. Для интерполяции допустимы формальные, линейные модели, но для экстраполяции – только содержательные, нелинейные модели. (см. раздел 6.2)

11. При экстраполяции всегда необходимо оценивать точность прогноза – строить доверительный интервал. Метод связанного моделирования позволяет сделать это также точно, как и традиционные методы (например, bootstrap), но во много раз быстрее. Однако он медленнее, чем грубые, приближенные методы (например, стохастическая аппроксимация).

12. Построение доверительных интервалов для прогноза требует предварительной оценки сложности, нелинейности модели. Коэффициент нелинейности характеризует эту

сложность и определяет метод, которым надо пользоваться. Если этот коэффициент близок к единице, то можно использовать быстрые и грубые методы, а если нет – то медленные и точные. (см. раздел 7.4)

13. Коэффициент нелинейности зависит как от формы модели, так и от плана эксперимента. Изменив план эксперимента, можно приблизить нелинейную модель к линейной и наоборот. (см. раздел 3.4)

14. Результаты оценивания часто становятся исходными данными на следующем этапе моделирования. Для их правильной интерпретации нужно использовать метод ПБО, который учитывает погрешности в этих оценках. (см. раздел 6.3)

15. Большие, но однородные экспериментальные данные (например, спектры), описываемые одной моделью с большим числом неизвестных параметров, можно разбивать на части и обрабатывать последовательно, используя метод ПБО. (см. главу 7)

16. Многооткликовые и разнородные данные, описываемые разными моделями, зависящими от общих параметров, можно обрабатывать методом ПБО. При этом каждый отклик оценивается отдельно, но с учетом уже имеющейся априорной информации об общих параметрах. (см. главу 8)

17. Отдельные части экспериментальных данных можно обрабатывать по несколько раз с помощью ПБО и ОПБО. При этом необходимо следить за тем, чтобы одни и те же параметры не оценивались по тем же данным еще раз. (см. разделы 7.3 и 8.3)

18. Мультиколлинеарность – это главная проблема в нелинейном моделировании. Для ее решения необходимо так преобразовать исходные данные и модель, чтобы уменьшить разброс собственных значений информационной матрицы. Центрирование, шкалирование и перепараметризация – это основные методы таких преобразований. (см. раздел 4.4)

19. Уравнение Аррениуса всегда нужно использовать в модифицированном виде. Одно это может уменьшить разброс собственных значений на 5-10 порядков. (см. разделы 6.1 и 8.2)

20. Часто для достижения цели нужно изменить исходную, простую модель, заменив ее более сложной, состоящей из нескольких, «склеенных» формул. В такой форме она может лучше и, главное, стабильнее описывать данные. Это также является и методом борьбы с мультиколлинеарностью. (см. главу 9)

21. Подбор содержательной модели для описания экспериментальных данных – это основная проблема нелинейного моделирования. Конечный выбор зависит, конечно, от опыта и знаний исследователя. Однако применение простого и эффективного инструмента упрощает эту задачу, позволяет быстро проверить различные варианты. Компьютерная система Fitter является именно таким инструментом.

22. Набор моделей, приведенных в работе, можно использовать как базовый для построения других аналогичных моделей. Основные приемы, используемые в системе Fitter для модификации моделей это – «склейка» формул, устранение неопределенностей, дифференцирование по предиктору. (см. главу 9)

23. Используя систему Fitter, можно не только проводить интерактивную интерпретацию данных, но и строить «шаблоны» – программы для автоматической, пакетной обработки стандартных наборов данных. (см. главу 11).

Приложение 1. Список основных обозначений

| Символ | Формула | Смысл |
|--|------------------------|--|
| $\mathbf{a} = (a_1, \dots, a_p)^t$ | (1.6) | Вектор неизвестных параметров |
| $\hat{\mathbf{a}} = (\hat{a}_1, \dots, \hat{a}_p)^t$ | (1.23) | Вектор оценок параметров |
| $\mathbf{a}^* = (a_1^*, \dots, a_p^*)^t$ | (3.24) | Вектор генерируемых оценок параметров |
| \mathbf{A} | (1.38), (2.20), (2.30) | Матрица Гессе |
| $\mathbf{b} = (b_1, \dots, b_p)^t$ | (1.18) | Вектор априорных значений параметров |
| $B(\mathbf{a})$ | (2.12), (2.13) | Байесовский член |
| \mathbf{C} | (1.36) | Ковариационная матрица |
| d | (1.59) | Оценка взвешенной дисперсии по выборке |
| D | (9.1) | Коэффициент диффузии |
| $D()$ | | Дисперсия |
| $E()$ | | Математическое ожидание |
| $E_n()$ | (6.7) | Интегральная экспонента |
| erf | (9.14) | Интеграл ошибок |
| \mathbf{F} | (1.42) | F-матрица |
| $F_{m,n}(P)$ | (1.52) | Квантиль F -распределения |
| \mathbf{H} | (1.19) | Априорная информационная матрица |
| \mathbf{I} | | Единичная матрица |
| $L(\mathbf{a}, \sigma^2)$ | (1.29), (2.8), (2.9) | Функция правдоподобия |
| m | (1.4) | Число предикторов в модели |
| N_0 | (1.21) | Априорное число степеней свободы |
| N | (1.9) | Число наблюдений |
| $N()$ | (4.19) | Число обусловленности матрицы |
| N_f | (1.44), (2.14) | Число степеней свободы по регрессии |
| N_s | (1.60) | Число степеней свободы по выборке |
| N_w | (1.45) | Число наблюдений с ненулевым весом |
| $N(m, S)$ | (3.19) | Плотность нормального распределения |
| p | (1.7) | Число параметров |
| P | | Вероятность, достоверность |

Приложение 1. Список основных обозначений

| | | |
|--|------------------------|--|
| $Q(\mathbf{a})$ | (1.35), (2.10), (2.11) | Целевая функция |
| s^2 | (1.33), (2.16) | Оценка взвешенной дисперсии по регрессии |
| s_0^2 | (1.20) | Априорное значение взвешенной дисперсии |
| $S(\mathbf{a})$ | (1.22) | Сумма квадратов |
| $t_m(P)$ | (1.51) | Квантиль распределения Стьюдента |
| $\text{tr}()$ | | След матрицы |
| V | (1.39) | Матрица взвешенных производных |
| $\mathbf{w} = (w_1, \dots, w_N)^t$ | (1.17) | Вектор весовых коэффициентов |
| $\mathbf{x} = (x_1, \dots, x_m)^t$ | (1.2) | Независимые переменные |
| X | (1.10) | Матрица плана |
| $\mathbf{y} = (y_1, \dots, y_N)^t$ | (1.8) | Вектор данных наблюдения (отклик) |
| $z(P)$ | (1.49) | Квантиль нормального распределения |
| α | | Уровень значимости |
| γ | Табл. 3.3 | Грубый коэффициент нелинейности |
| Γ | (3.44) | Точный коэффициент нелинейности |
| $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_N)^t$ | (1.13) | Вектор ошибок |
| σ^2 | (1.16) | Взвешенная дисперсия |
| $\chi_m^2(P)$ | (1.50) | Квантиль распределения хи-квадрат |

Приложение 2. Список использованных сокращений

| <i>Обозначение</i> | <i>Смысл</i> |
|--------------------|--|
| АО | Антиоксидант |
| ДСК | Дифференциальная сканирующая калориметрия |
| ИК (спектр) | Инфракрасный |
| ММП | Метод максимума правдоподобия |
| МНК | Метод наименьших квадратов |
| ОМП | Оценки максимума правдоподобия |
| ОПБО | Обратное последовательное байесовское оценивание |
| ПБО | Последовательное байесовское оценивание |
| ПВХ | Поливинилхлорид |
| ПЭ | Полиэтилен |
| СКО | Среднеквадратичное отклонение |
| ТГА | Термогравиметрия |
| ТМА | Термомеханический анализ |
| ТНО | Температура начала окисления |
| УИ | Ускоренные испытания |
| УФ (спектр) | Ультрафиолетовый |
| ЭПУИ | Эволюционное планирование ускоренных испытаний |
| ELB | Разрывное удлинение |
| ELN | Удлинение |
| MOD | Условный модуль |
| STR | Напряжение |
| TEN | Разрывное напряжение |
| VBA | Visual Basic for Applications |

Литература

1. Дрейпер Н., Смит Г. *Прикладной регрессионный анализ*, . – М.: Статистика, 1973.
2. Горский В.Г., Зейналов М.З. *Физико-химические и математические основы феноменологической кинетики сложных реакций*, Махачкала, ИПЦ ДГУ, 1997.
3. Бард И. *Нелинейное оценивание параметров*. – М.: Статистика, 1979.
4. Brown S.D., Sum S.T., Despagne F., Lavine B.K., *Chemometrics, Anal.Chem.*, **68**, 21R-61R (1996)
5. Померанцев А.Л., Родионова О.Е. Надстройка FITTER (FITTER). *Свидетельство об официальной регистрации № 2002611562 от 11.09.02*
6. Fitter Add-Inn. [On line], <http://polycert.chph.ras.ru/fitter.htm> [1 января 2002]
7. Walkenbach J. *Excel 2000 Power Programming with VBA*, 2001
8. Померанцев А.Л. Интегрированная компьютерная система Kinetic Trunk. *Свидетельство об официальной регистрации № 940061 от 28.02.94*
9. Быстрицкая Е.В., Максимова Г.А., Померанцев А.Л. Интегрированная компьютерная система Kinetic Trunk. *EMRS 1993 Fall Meeting 4th European East-West Conference & Exhibition on Materials and Process: Тезисы докладов*. С.-Петербург, **1**, 117, (1993)
10. Померанцев А.Л. Интегрированная система Kinetic Trunk, *Материалы I рабочей школы-совещания КЦ "Эксплуатационная устойчивость"*, Можанка, 196-207, (1991).
11. Райфа Г., Шлейфер Р. *Прикладная теория статистических решений*, М. Статистика, 1977
12. Крамер Г. *Математические методы статистики*, М., Мир, 1975
13. Джонс Н., Лион Ф., *Статистика и планирование эксперимента в технике и науке. Методы обработки данных*. М., Мир, 1980
14. Боровиков А.А. *Математическая статистика*, М., Наука, 1984
15. Rodionova O.Ye., Pomerantsev A.L. SIC (Simple Interval Calculation) - a new approach for linear modeling. *J. Chemometrics* (submitted), 2001
16. Carney T.M., Goldwyn R.M. Numerical experiments with various optimal estimators. *J. Optimization Theory Appl.* **1**, 113-130, (1967)
17. Chow G.C. A comparison of alternative estimators for simultaneous equations. *Econometrics*, **32**, 532-553 (1964)

18. Cragg J.G. On the relative small sample properties of several structural-relative estimators. *Econometrics*, **35**, 89-110, 1967
19. Демиденко Е. З. *Линейная и нелинейная регрессии*, М, Финансы и статистика, 1981
20. Chambers J.R. Fitting nonlinear models: numerical techniques, *Biometrika*, **60**, 1-13, (1973)
21. Efron B. Bootstrap methods: another look at the jackknife, *Ann. Statist.*, **7**, 1-26, (1979)
22. Efron B. Better bootstrap confidence intervals, *J.Amer.Statist.Ass.*, **81**, (1987)
23. Эфрон Б. *Нетрадиционные методы многомерного статистического анализа*, Москва, Финансы и Статистика (1988)
24. Breiman, L. Bagging predictors, *Machine Learning* **24**,123-140, (1996).
25. Kong E. B., Dietterich T. G.. Error correcting output coding corrects bias and variance. In Proc. *Twelfth International Conference on Machine Learning*, Tahoe City, CA: Morgan Kaufmann. 313-321, (1995)
26. Madigan D., Raftery A.E., Volinsky C.T., Hoeting J.A. Bayesian model averaging. In Proc. *AAAI-96 Workshop on Integrating Multiple Learned Models*, Portland, OR: AAAI Press, 77-83 (1996).
27. Клейнен Дж. *Статистические методы в имитационном моделировании*, М, Статистика, 1978
28. Quenouille M. Approximate tests of correlation in time series, *J.Roy.Statist.Soc.*, Ser B, **11**, 18-84
29. Cornfield J. Bayes Theorem, *Rev. Inst. Internat. Statist.* **35**, 34-49 (1967)
30. Максимова Г.А., Померанцев А.Л., Последовательное байесовское оценивание параметров регрессии, *Зав. Лаб.*, **61** ,432-435 , (1995)
31. Померанцев А.Л. Байесовское оценивание параметров химических моделей на ПЭВМ. *VI Всесоюзная конференция "Математические методы в химии"*, Тезисы докладов. Новочеркасск, 1989, 68-69
32. Bystritskaya E.V., Pomerantsev A.L., Rodionova O.Ye., Non-linear regression analysis: new approach to traditional implementations *J. Chemometrics*, **14** (2000), 667-692.
33. Karpukhin O.N., Pomerantsev A.L. The Harmonization of Analytical Study, *Proc. of III Inter. Symp. on the Harmonization of Quality Assurance Systems in Chemical Analysis*, Washington, D.C. ,1989.
34. Быстрицкая Е.В., Карпухин О.Н., Померанцев А.Л., Репина Т.С. Планирование и интерпретация ускоренных испытаний полимерных материалов на термовлажност-

- ное старение, *VIII конференция по старению и стабилизации полимеров*. Тезисы докладов. Душанбе, 1989.,87-88.
35. Карпухин О.Н. Использование априорной информации о кинетике и механизме старения полимерного материала при оценке и регулировании его эксплуатационной устойчивости, *Там же* с. 94-95.
 36. Karpukhin O.N. Physico-Chemical Problems in Estimation of Service Steadiness (SS) of Polymer Materials. *Intern. J. Polymeric Mater.*, **13**, 21-33., (1990)
 37. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. *Прикладная статистика основы моделирования и первичная обработка данных.*, М., Финансы и статистика,1983.
 38. Померанцев А.Л., Карпухин О.Н., Брин Э.Ф. Кинетический подход к прогнозированию надежности материалов в изделии. *Кинетика и катализ.* **24** ,1233-1238 (1983)
 39. Померанцев А.Л., Брин Э.Ф., Карпухин О.Н. Математические проблемы прогнозирования надежности материалов в изделии по результатам кинетических исследований. *Зав. лаб.*, **5**,48-52. (1986)
 40. Быстрицкая Е.В., Максимова Г.А., Померанцев А.Л. Интегрированная компьютерная система Kinetic Trunk. *EMRS 1993 Fall Meeting 4th European East-West Conference*. Тезисы докладов, С.-Петербург, **1** ,117, 1993,
 41. Карпухин О.Н., Померанцев А.Л. Банк первичных экспериментальных данных. *Всесоюзное совещание "Создание сети автоматизированных банков о свойствах и применении полимерных материалов*. Тезисы докладов. Ленинград, **7**, 19-20, (1987).
 42. Bystritskaya E.V., Pomerantsev A.L., Rodionova O.Ye. Prediction the behavior of polymer materials at different conditions, *Chemometrics and Intell. Lab. Syst.*, **47**, 175-179, (1999)
 43. Pomerantsev A.L., Successive estimating of reaction rate constants from spectral data: a case study of two-step kinetics. Представлено в *J.Chemometrics*, (2000)
 44. Быстрицкая Е.В., Карпухин О.Н., Максимова Г.А., Прогнозирование срока службы резин методом последовательного байесовского оценивания, *Межд. Конференция "Rubber-94"* , тезисы докладов Москва., **2**, 605-612, (1994).
 45. Померанцев А.Л. , Применение интегрированной компьютерной системы Kinetic Trunk для прогнозирования срока службы РТИ, *Там же*,**4**, 291-297
 46. Быстрицкая Е.В., Померанцев А.Л., Прогнозирование сроков хранения шинных резин методом последовательного байесовского оценивания, *Труды 2-ой межд. конфер. ЭМФ-1*, **1**, 52-61, Барнаул, (2001)

47. Быстрицкая Е.В., Быков Е.В., Кононенко А.И., Померанцев А.Л., Разработка методов определения эксплуатационной устойчивости кабелей для атомных станций. В сб. "Вопросы атомной науки и техники. Серия: Физика радиационного воздействия на радиоэлектронную аппаратуру." Москва, (1995).
48. Быстрицкая Е.В., Быков Е.В., Крючков А.А., Померанцев А.Л. и др. Прогнозирование периода индукции для полиэтиленовой кабельной изоляции. *Там же*
49. Bystritskaya E.V., Pomerantsev A.L., Rodionova O.Ye., Prediction the behavior of polymer materials at different conditions, *Межд. конф. CONFERENTIA CHEMOMETRICA (CC'97)*. Тезисы доклада, Budapest, Lec.16, (1997).
50. Bystritskaya E.V., Pomerantsev A.L., Rodionova O.Ye., Kinetic Trunk - computerized system for nonlinear regression analysis and statistical simulation, *Там же*. Pos.54.
51. Bystritskaya E.V., Pomerantsev A.L., Rodionova O.Ye , Application of nonlinear regression and sequential bayesian estimation technique for testing antioxidants in polypropylene by DSC method., *Там же*, Pos.69.
52. Bystritskaya E.V., Pomerantsev A.L., Rodionova O.Ye., Testing antioxidants in polypropylene with processing of DSC data by nonlinear regression and sequential bayesian estimation technique, *Межд. конф. 5th Scandinavian Symposium on Chemometrics (SSC5)*. Тезисы доклада . Lahti., 46, (1997).
53. Bystritskaya E.V., Pomerantsev A.L., Rodionova O.Ye , Life-time prediction of polymer materials at different conditions, *Там же*, 47
54. Горский В.Г. *Планирование кинетических экспериментов*, М., Наука, 1984
55. Чуев Ю.В., Михайлов Ю.Б., Кузьмин В.И. *Прогнозирование количественных характеристик процессов*, М., «Сов. радио», 1975
56. Efron B., Tibshirani R., Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, *Statistical Science*, **1**(1), 54-75 (1986)
57. Pomerantsev A.L., Confidence intervals for nonlinear regression extrapolation, *Chemometrics and Intell. Lab. Syst.*, **49**, 41-48, (1999)
58. Fox L. *An Introduction to Numerical Linear Algebra*, Oxford Univ. Press (Clarendon), London and New York, 1964
59. Barra J.R., *Notions Fondamentales de Statistique Mathematique*, Dunod, Paris, 1971
60. Ермаков С.М. *Метод Монте-Карло и смежные вопросы*, М., Наука, 1975
61. Beale E.M.L. Confidence regions in nonlinear estimation, *J. Roy. Statist. Soc. Ser B*, **22**, (1), 41-75, (1960)

62. Guttman I., Meeter D.A.. On Beale's measures of nonlinearity, *Technometrics*, **7**, (4), 623-637, 1965
63. Bates D.M., Watts D.G. Relative curvature measures of nonlinearity. *J. Roy. Statist. Soc. Ser B*, **42**, 1-25, 1980
64. Bates D.M., Watts D.G., Parameters transformations for improved approximate confidence regions in nonlinear least squares, *Ann. Statist.*, **9**, (6), 1152-1167, 1981
65. Гантмахер Ф.Р. *Теория матриц*, М, Наука, 1966
66. Кюнцци Г.П., Крелле В., Нелинейное программирование, М., Сов. Радио, 1965.
67. Marquardt D.W. An algorithm for least squares estimation of nonlinear parameters, *SIAM J.*, **11**, 431-441, (1963).
68. Levenberg K. A method for solution of certain nonlinear problems in least squares, *Quart. Appl. Math.*, **2**, 164-168, (1944)
69. Goldfeld S.M., Quandt R.E., Trotter H.F. Maximization by quadratic hill climbing, *Econometrics*, **34**, 541-551, (1966)
70. Jennrich R.I., Sampson P.F. Application of stepwise regression to nonlinear estimation, *Techometrics*, **10**, 63-72, (1961).
71. Golub G. Numerical methods for solving liner least square problems, *Numer. Math.*, **7**, 206-216, (1965)
72. Healy M., Multiple regression with singular matrix, *J. Roy. Statist. Soc. Ser. C Appl. Statist.* **17**, 110-117, (1968).
73. Longley J.W. An appraisal of lest squares program for the electronic computer from the point of view of the user, *J. Amer. Statist. Assoc.*, **62**, 819-841, (1967).
74. Павлов Б.В., Повзнер А.Я., Об одном методе численного интегрирования систем обыкновенных дифференциальных уравнений, *Жур. Вычисл. Мат. Мат. Физ.*, **13**, 1056-1059, (1973).
75. Демидович Б.П., Марон А.И. *Основы вычислительной математики*, М., Наука, 1966
76. Михлин С.Г., Смолицкий Х.Л., *Приближенные методы решения дифференциальных и интегральных уравнений*, СМБ, Наука, 1965.
77. Powell J.D. An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Comput. J.*, **7**, 303-307, (1965).
78. Powell J.D. A method for minimizing a sum of squares of nonlinear functions without calculating derivatives, *Comput. J.*, **7**, 155-162, (1964).

79. Fletcher R., Function minimization without evaluating derivatives – a review *Comput.J.*, **8**, 33-41, (1965), *Comput.J.*, **8**, 33-41, (1965)
80. Stewart III G.W. A modification of Davidon's minimization method to accept difference approximations of derivatives, *J. Assoc. Comput. Math.* **14**, 72-83, (1967).
81. StRD Nonlinear Least Squares Regression Datasets, [On line], http://www.nist.gov/itl/div898/strd/nls/nls_main.shtml [1 Января 2002]
82. More J.J., Garbow B. S., Hillstom K.E., Testing unconstrained optimization software, *ACM Trans. Math. Softw.*, **7**, 17-41, (1981)
83. Meyer, R. R. Theoretical and computational aspects of nonlinear regression. In *Nonlinear Programming*, Rosen, Mangasarian and Ritter (Eds). NY: Academic Press, 465-486. (1970).
84. DataFit, [On line], <http://www.oakdaleengr.com/datafit.htm> [1 Января 2002]
85. GraphPad, [On line], <http://www.graphpad.com> [1 Января 2002]
86. NLREG, [On line], <http://www.nlreg.com/> [1 Января 2002]
87. Microcalc Origin, [On line], <http://www.microcalc.com> [1 Января 2002]
88. SAAM [On line], <http://www.saam.com/software/saam2/saam2software.htm> [1 Января 2002]
89. SigmaPlot [On line], <http://www.spssscience.com/SigmaPlot/index.cfm> [1 Января 2002]
90. XLStat, [On line], <http://www.xlstat.com> [1 Января 1999]
91. Dennis J.E., Gay D.M., Welsch R.E., An adaptive nonlinear least-squares algorithm. *ACM Trans. Math. Softw.* **7**, 3 (1981)
92. Спивак С.И., Горский В.Г. *Докл. АН СССР*, **257**, 2, 412-415 (1981)
93. Клибанов М.В., Спивак С.И., Тимошенко В.И., Слинко М.Г. *Докл. АН СССР*, **208**, 5, 1387-1390 (1973)
94. Горский В.Г., Спивак С.И. *Завод. Лаб.*, **47**, 10, 30-47, (1981)
95. Эмануэль Н.М., Кнорре Д.Г., *Курс химической кинетики*, М., Высшая школа, 1984.
96. Bystritskaya E.V., Pomerantsev A.L., Rodionova O.Ye.. Prediction the behavior of polymer materials at different conditions. *Chemometrics and Intell. Lab. Syst.*, **47** (1999), 175-179
97. Минскер К.С., Козлов Б.Ф., Заиков Г.Е. *Старение и стабилизация полимеров на основе винилхлорида*. М.:Наука, 1982, 272
98. Борисов Б.И., Исследование старения ПВХ пленок в почве, *Жур. прикл. химии*, **43**, 5, 116-120, (1978).
99. Брагинский Р.П., *Док. АН СССР*, **303**, 3, 535-537, (1988).

100. Бочкарев Р.Н., Филатов И.С. *Старение материалов на основе поливинилхлорида в условиях холодного климата*. Новосибирск: Наука. Сиб.отд-ние, 1990.
101. Штаркман Б.П. *Пластификация поливинилхлорида*. М., Химия, 1975.
102. Bystritskaya E.V., Rodionova O.Ye., Overbergh N., NLR fitting of thermomechanic curves of electron-beamed polyethylene and curing kinetics, *Межд. конф. 6th Scandinavian Symposium on Chemometrics (SSC6)*. Тезисы доклада, Porsgrunn, Norway, .86, (1999).
103. Dawkins I.V. (Ed.). *Development in Polymer Characterization*, 1., L., 1978.
104. Charlesby A.. *Atomic Radiation and Polymers*. Pergamon Press, 1960
105. Flory P.J. *Principles of Polymer Chemistry*. NY Cornell. Univ. Press, 1953.
106. Ferry J. D., Kan H.C. Interpretation of deviations from neo-Hookean elasticity by a two-network model with crosslinks and trapped entanglements *Rubb. Chem. Technol.*, **51**, 731 - 737, (1978)
107. Shlyapnikov Yu.A., Polyolefine chain oxidation *In: Development in Polymer Stabilization*, London - New Jersey, Applied Science Publishers, **5**, 1-22, (1981)
108. Pomerantsev A.L Successive estimating of reaction rate constants from spectral data: A case study of two-step kinetics, *Межд. конф. 7th Scandinavian Symposium on Chemometrics (SSC-7)* Тезисы доклада, Copenhagen, Denmark, (2001).
109. Pomerantsev A.L. Evaluation of kinetics rate constants from spectra data by successive Bayesian estimation, *In Aging of polymers, polymer blends and polymer composites*, **2**, 1-18, Novascience Publishers, NY, (2002)
110. Pomerantsev A.L. Estimation of reaction rate constants from spectra data by successive Bayes method, *Межд. конф. Кетометрия '02*, Tata, Hungary, Lec. 4, (2002)
111. Haario H, Taavitsainen V-M, Combining soft and hard modelling in chemical kinetics models. *Chemometrics Intell. Lab. Syst.*; **44**:77-98 (1998)
112. Bijlsma S, Louwerse DJ, Windig W, Smilde AK. Rapid estimation of rate constants using on-line SW-NIR and trilinear models. *Anal.Chim.Acta*; **376**:339-335 (1998)
113. Bijlsma S Louwerse DJ, Smilde AK. Estimating rate constants and pure UV-VIS spectra of a two-step reaction using trilinear models. *J. Chemometrics*; **13**:311-329 (1999)
114. Павлов Б.В., Брин Э.Ф. Обратные задачи химической кинетики, *Хим. физика*, **3**, 3, 393-404, (1984)
115. Брин Э.Ф., Померанцев А.Л. Классификация обратных задач кинетики гомогенных химических реакций. *Хим. физика*, **5**, 12, 1674-1682 (1986)

116. Bijlsma S, Smilde AK. Estimating reaction rate constants from a two-step reaction: a comparison between two-way and three-way methods. *J. Chemometrics*; **14**:541-560, (2000)
117. Datasets [Online] <http://www.its.chem.uva.nl/research/pac>, [1 December 2000]
118. Savitzky A. , Golay M.J.E. , Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**: 1627-1639, (1964)
119. Lawton W.H., Sylvestre E.A. Self modeling curve resolution. *Technometrics*, **13**: 617-633, (1971)
120. Wilson B., Sanchez E., Kowalski B.R. An improved algorithm for the generalized rank annihilation method. *J. Chemometrics*, **3**: 493-498, (1998)
121. Harshman R.A., Lundy M.E. PARAFAC: parallel factor analysis *Comp. Statist. Data Anal.* **18**:39-72, (1994)
122. Esbensen K.H. *Multivariate Data Analysis - In Practice 4-th Ed.*, CAMO, 2000.
123. Martens H., Naes T. *Multivariate Calibration*, Wiley, New York, 1998.
124. Карпухин О.Н. Определение срока службы полимерного материала как физико-химическая задача. *Успехи химии*, **XLIX**, 8, 1523-1553, (1980).
125. Bystritskaya E.V., Rodionova, O.Ye., Pomerantsev A.L., Evolutionary design of experiment for accelerated aging tests, *Pol. Testing*, **19**, 221-229, (1999).
126. Fitter Solutions [On line], <http://polycert.chph.ras.ru/solution.htm> , [1 января 2002].
127. Hösskuldsson A. *Prediction methods in Science and Technology*, Thor Publ., Denmark, 1996.
128. Померанцев А.Л., Кротов А.С., Родионова О.Е. *Компьютерная система FITTER для регрессионного анализа экспериментальных данных*, Учебное пособие, Барнаул, Из-во АГУ, 2001
129. Crank J. *The mathematics of diffusion*. (second edition) – Oxford, UK: Clarendon press, 1975.
130. Заиков Г.Е., Иорданский А.Л., Маркин В.С. , *Диффузия электролитов в полимерах*. М. "Химия" 1984.
131. Репина Т.С., Дуфлот В.Р., Быстрицкая Е.В. *V Всесоюзное совещание по полимерным оптическим материалам*. Тезисы докладов, Ленинград, 84, 1991.
132. Тихонов А.Н. , Самарский А.А., *Уравнения математической физики*, М, Наука, 1972

133. Auriault J.-L., Lewandowska J. Modelling of pollutant migration in porous media with interfacial transfer: local equilibrium/non-equilibrium, *Mechanics of cohesive-frictional materials*, **2**, 205-221, (1997)
134. Startsev O., Krotov A., Mashinskaya G., Climatic ageing of organic fiber reinforced plastics: water effect., *Intern. J. Polymeric Mater.*, **37**, 161-171, (1997).
135. Аскадский А.А., Матвеев Ю.И. *Химическое строение и физические свойства полимеров.* – М.: Химия, 1983.
136. Startsev O.V., Krotov A.S., Golub P.D. Effect of climatic and radiation ageing on properties of glass fibre reinforced epoxy laminates, *Polymers and Polymer Composites*, **6**, 7, 481-488 (1998.).
137. Startsev O.V., Krotov A.S., Startseva L.T. Interlayer shear strength of polymeric composite materials during long term climatic ageing, *Polymer Degrad. & Stab.*, **63**, 183-186, (1999)
138. Startsev O.V. Peculiarities of ageing of aircraft materials in a warm damp climate. *Polymer Yearbook 11* , Ed. By R.A. Pethrick, Glasgow, UK: Harwood Academic Publishers, 91-109.(1993).
139. Xiao G.Z., Shanahan M.E.R. Water absorption and desorption in an epoxy resin with degradation, *J. Polym. Scien. B: Pol. Phys.* **35**, 2659-2670, (1997).
140. Schult K.A., Paul D.R. Water sorption and transport in blends of poly(vinyl pyrrolidone) and polysulfone, . *J. Polym. Scien. B: Pol.Phys* , **35**,.655-674, (1997)
141. Lee S., Knaebel K.S. Effects of mechanical and chemical properties on transport in fluoropolymers. I. Transient sorption, *J.Appl. Pol. Scien*, **64**,.455-476, (1997)
142. Быстрицкая Е.В., Померанцев А.Л. *V Всесоюзное совещание по полимерным оптическим материалам.*, Ленинград, Тезисы докладов,17-19, (1991).
143. Абрамовиц М., Стиган И. *Справочник по специальным функциям* , М, Наука, 1979
144. Фихтенгольц Г.М. *Курс дифференциального и интегрального исчисления* , М, Наука, 1966
145. Быстрицкая Е.В. Действие влаги на гидрофобные полимерные стекла. *I-я рабочая школа-совещание КЦ "Эксплуатационная устойчивость"*, Моженка, Материалы, 207-229 (1991).
146. Марьянов Б.М. Хемометрика в титровании, *Химики ТГУ на пороге третьего тысячелетия.* Томск, Изд-во ТГУ, 48-58, (1998)
147. Waldmeier K, Rellstab W. Auswertung potentiometrischer Titration mittels nichtlinearer Regression, *Z. anal. Chem.* , **264**, 5, 337-341. (1973)

148. Sato H., Momoki K. Least squares curve-fitting method for end-point detection of chelometric titration with metal indicator, *Anal. Chem.*, **42**, 13, 1477-1482, (1970).
149. Джонсон К. *Численные методы в химии*. - М.: Мир, 1983, 280-299
150. Barry D.M., Meites L. Titrimetric applications of multiparametric curve-fitting. Part 1. Potentiometric titrations of weak bases with strong acids at extreme dilutions. *Anal. Chim. Acta.*, **68**, 2, 435-445, (1974)
151. Barry D.M. et al. Titrimetric applications of multiparametric curve-fitting. Part II. Potentiometric titration with an unstandardised reagent *Anal. Chim. Acta.*, **69**, 1, 143-151, (1974).
152. Марьянов Б.М. Точка эквивалентности и константа реакции при линейных титрованиях, *Ж Аналит. Химии*, **27**, 11, 2099-2106, (1972).
153. Сидоренко В.И., Гордиенко В.И., *Ж Аналит. Химии*, **24**, 315, (1969).
154. Марьянов Б.М., Богданов В.С. Точка эквивалентности и константа реакции при линейных титрованиях, *Ж Аналит. Химии*, **30**, 3, 459-464, (1975).
155. Марьянов Б.М. Строгий метод спрямления кривых потенциометрического ионно-ассоциативного титрования с участием электроактивного реагента, *Ж Аналит. Химии*, **30**, 10, 1892-1897, (1975).
156. Марьянов Б.М., Пронин В.А. Применение ЭВМ для обработки данных потенциометрического титрования при определении серебра в медеэлектролитных шламах, *Ж Аналит. Химии*, **39**, 9, 1625-1629, (1984).
157. Марьянов Б.М. Оценка константы ионизации и концентрации слабой одноосновной кислоты по данным потенциометрического титрования с помощью ЭВМ, *Ж Аналит. Химии*, **34**, 4, 655-660, (1979).
158. Марьянов Б.М., Никифорова И.А.. Применение ЭВМ для обработки логарифмических кривых кулонометрического осадительного титрования, *Ж Аналит. Химии*, **42**, 3, 435-440, (1987).