

ВМК



А.Н.ТИХОНОВ, М.В.УФИМЦЕВ

**Статистическая
обработка
результатов
экспериментов**

**ИЗДАТЕЛЬСТВО
МОСКОВСКОГО
УНИВЕРСИТЕТА**

А.Н.ТИХОНОВ, М.В.УФИМЦЕВ

Статистическая обработка результатов экспериментов

Допущено Министерством высшего
и среднего специального образования СССР
в качестве учебного пособия
для студентов вузов,
обучающихся по специальности
«Прикладная математика»

ИЗДАТЕЛЬСТВО
МОСКОВСКОГО
УНИВЕРСИТЕТА
1988

Тихонов А. Н., Уфимцев М. В. Статистическая обработка результатов экспериментов: Учеб. пособие. — М.: Изд-во Моск. ун-та, 1988. — 174 с. — ISBN 5—211—00105—2.

В учебном пособии рассматриваются вопросы статистической обработки экспериментальных данных, являющейся важным этапом при проведении полной математической обработки результатов экспериментов. Выделены основные проблемы, связанные с применением вероятностно-статистических методов, начиная от создания математической модели и вплоть до интерпретации результатов. Наряду с традиционными в статистике вопросами изложены методы робастного оценивания, планирование эксперимента, факторный анализ; большое внимание уделено применению устойчивых регуляризирующих алгоритмов.

Рецензенты:

кафедра прикладной математической физики МИФИ,
член-корреспондент АН СССР *Н. Н. Говорун*

Т 1704010000(4309000000)—048 105—88
077(02)—88

ISBN 5—211—00105—2

© Издательство
Московского университета, 1988 г.

ОГЛАВЛЕНИЕ

<i>Список принятых обозначений и сокращений</i>	4
<i>Предисловие</i>	6
Глава 1. Основные этапы и задачи статистической обработки экспериментальных данных	8
Глава 2. Проверка закона распределения наблюдений	17
Глава 3. Выявление аномальных наблюдений в одномерных выборках	29
Глава 4. Выявление аномальных наблюдений в совокупности связанных величин	40
Глава 5. Робастные методы и оценки	46
Глава 6. Линейный регрессионный анализ	58
Глава 7. Линейный регрессионный анализ в случае нормальности. Анализ остатков	70
Глава 8. Вычислительные аспекты метода наименьших квадратов. Робастный МНК. Нелинейный метод наименьших квадратов	78
Глава 9. Регуляризирующие алгоритмы. Регуляризованный метод наименьших квадратов	94
Глава 10. Введение в планирование регрессионных экспериментов	111
Глава 11. Применение метода Монте-Карло в анализе и планировании экспериментов	124
Глава 12. Применение факторного анализа для оценки точности данных	133
Глава 13. Учет систематических искажений	142
Глава 14. Применение рекуррентного оценивания и фильтрации при обработке данных	149
<i>Литература</i>	164

Список принятых обозначений и сокращений

$\overset{\cdot}{=} \overset{\cdot}{=}$ — по определению есть (равно);

\sim — распределено как (по закону);

X^T, \vec{y}^T — транспозиция матрицы X , вектора \vec{y} ;

$\det \Phi, |\Phi|$ — определитель матрицы Φ ;

$\text{diag} \{\lambda_1, \dots, \lambda_n\}$ — диагональная матрица с элементами $\lambda_1, \dots, \lambda_n$ на главной диагонали (возможно также обозначение $\text{diag} \{H\}$ — диагональная матрица, имеющая на главной диагонали диагональные элементы $H_{11}, H_{22}, \dots, H_{nn}$ матрицы H);

I_k — единичная матрица размера k ;

$x_0 = \text{Arg min } \varphi(x)$ — значение x , доставляющее минимум $\varphi(x)$ (аналогично $\text{Arg max } \varphi(x)$, $\text{Arg inf } \varphi(x)$);

$\text{rang } \Phi$ — ранг матрицы Φ ;

СЛАУ — система линейных алгебраических уравнений;

с. з. — собственные значения;

с. в. — случайная величина;

ф. р. — функция распределения;

ф. п. в. — функция плотности вероятностей;

МП — максимум правдоподобия;

ФП — функция правдоподобия;

ОП — отношение правдоподобия;

$\overset{P}{x_n} \rightarrow x_0$ — сходимость по вероятности;

$x_0 = p - \lim x_n = x_0$ — есть предел $\{x_n\}$ в смысле сходимости по вероятности;

$E\xi$ — математическое ожидание (м. о.) случайной величины ξ ;

$D\xi$ — дисперсия с. в. ξ ;

$\text{cov}(\xi, \eta)$ — ковариация случайных величин ξ и η ;

\bar{y} — выборочное среднее для выборки y_1, \dots, y_n ;

$\varphi_i(t)$ — характеристическая функция (х. ф.) с. в. ξ_i ;
 $N(a, \sigma)$ — одномерное нормальное распределение с м. о. a и стандартным отклонением σ ;
 $N(\vec{a}, Q)$ — многомерное нормальное распределение с вектором м. о. \vec{a} и ковариационной матрицей Q ;
 $Po(\mu)$ — пуассоновское распределение с параметром μ ;
 χ_k^2 — распределение χ^2 с k степенями свободы;
 $St_l(x)$ — распределение Стьюдента с l степенями свободы;
 $Sn(k, m)$ — распределение Снедекора—Фишера с параметрами k, m ;
 $R(a, b)$ — равномерное (непрерывное) распределение на отрезке (a, b) ;
 НК — наименьшие квадраты;
 МНК — метод наименьших квадратов;
 НЛНО — наилучшая линейная несмещенная оценка;
 РА — регуляризирующий алгоритм;
 РМНК — регуляризованный метод наименьших квадратов;
 RSS — остаточная сумма квадратов (Residual sum of squares);
 SVD — декомпозиция по сингулярным значениям (Singular value decomposition);
 МГШ — модифицированный алгоритм Грама—Шмидта;
 $\text{tr } \Phi$ — след матрицы Φ .

Остальные обозначения либо стандартны, либо носят локальный характер и разъяснены в соответствующем месте текста.

Предисловие

Важное место в математической обработке экспериментальных данных занимают статистические методы. При проведении полной обработки результатов экспериментов различных классов, создании систем полной автоматизированной обработки на ЭВМ экспериментальных данных статистическая обработка является необходимым этапом решения задачи. Результаты статистического анализа позволяют оценить достоверность выводов, полученных при интерпретации экспериментального материала. При этом информация, извлеченная на статистическом этапе из наблюдений, является входной по отношению к этапу интерпретации с использованием регуляризирующих алгоритмов [79] и позволяет дать оценку неопределенности входных данных, необходимую для корректного применения этих алгоритмов и оценки неопределенности регуляризованного решения.

Двадцатилетний опыт применения метода регуляризации при обработке экспериментальных данных в различных областях науки позволил выделить ряд характерных задач и методов прикладной статистики, используемых на этапе статистической обработки. Эти задачи рассмотрены в публикациях авторов, упомянутых в последующих ссылках; создано соответствующее программное обеспечение [9]. Наряду с опытом конкретных приложений статистики большое влияние на понимание проблемы обработки экспериментальных данных оказало развитие теории регуляризирующих алгоритмов [79], в частности регуляризованного метода наименьших квадратов [81; 82].

В основу книги положен материал спецкурса М. В. Уфимцева по статистической обработке результатов экспериментов, на протяжении ряда лет читаемого на физическом факультете и факультете вычислительной математики и кибернетики МГУ. При обучении студентов выявилась трудность, связанная с разбросанностью материала в книгах, статьях и отчетах, порой труднодоступных. Попыткой дать кон-

центрированную основу для усвоения материала является настоящая книга. Наряду с рассмотрением приложений статистики к обработке наблюдений в конкретных классах экспериментов затрагиваются также вопросы, недостаточно подробно освещаемые на указанных факультетах в общем курсе теории вероятностей и математической статистики: проверка статистических гипотез и критерии согласия, вычислительные аспекты метода наименьших квадратов, теория планирования эксперимента, статистическое моделирование и т. д. Сведения из общей статистики, необходимые для понимания материала, приводятся в ходе изложения. Более глубокое их усвоение читателем и практическое применение возможно с помощью упомянутых в списке литературы книг и статей.

Гл. 1, 8, 12 написаны авторами совместно, гл. 9 принадлежит А. Н. Тихонову, остальные главы книги написаны М. В. Уфимцевым.

Выражаем глубокую благодарность А. В. Гончарскому и П. Н. Заикину за многочисленные ценные замечания.

А. Н. Тихонов, М. В. Уфимцев

ОСНОВНЫЕ ЭТАПЫ И ЗАДАЧИ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Широкое применение методов статистической обработки экспериментальной информации связано с «математизацией» биологических наук, с совершенствованием техники экспериментальных исследований, с широким применением ЭВМ (в том числе встроенных в измерительную аппаратуру и снабженных математическим обеспечением для обработки данного класса экспериментов), с повышением требований к качеству данных, диктуемым развитием теории.

Так, изучение процессов размножения и гибели растений, микроорганизмов, насекомых в биологии, сельском хозяйстве, экологии привело к рассмотрению статистических свойств обобщенных распределений: пуассон-биномиального, Неймана, Эрмита и т. д. [8]. Определение характеристик, описывающих процессы в соответствующей биологической популяции, возможно на основе наблюдений за численностью объектов на выбранных контрольных участках и сводится к задаче статистического оценивания параметров распределения. Математизация биологии и медицины приводит к появлению количественных моделей (например, в виде системы интегродифференциальных уравнений), описывающих биохимические процессы в организме, в клетке, механизмы действия лекарств, процессы развития опухолей и т. п. Обычно параметры этих моделей неизвестны и подлежат определению по результатам измерений, полученных для подопытных индивидов. Возникающие математические задачи обычно могут быть решены методами линейного или нелинейного регрессионного анализа с интерпретацией результатов на основе проверки статистических гипотез. Следует отметить, что обработка и интерпретация экспериментальных данных указанного класса не тривиальна, поскольку их специфика (малые объемы выборок, вариации индивидуальных характеристик) нередко приводит к неприменимости

«классических», основанных на выборках из нормально распределенных популяций, подходов, что приводит к необходимости использовать робастные методы.

Современный уровень естественнонаучного эксперимента характеризуется большими потоками информации. Так, в ядерной физике регистрируется до 10^6 — 10^7 событий за один эксперимент, огромен объем телеметрической информации в космических исследованиях и т. д. При этом даже визуальный просмотр данных, не говоря об анализе, невозможен без применения ЭВМ. Машинная автоматизация обработки приводит к потере значимости таких факторов, как опыт, интуиция исследователя: если при создании комплекса программ обработки не предусмотреть возможности выявления абсурдных, «нефизических» данных во входном потоке информации, ЭВМ обработает все подряд и выдаст результат, далекий от истинного. Ясно, что в силу наличия ошибок измерений, случайного характера экспериментальных данных способы обнаружения и парирования абсурдных наблюдений должны быть статистическими.

Многие эксперименты (например, при исследованиях космоса, при поиске новых элементарных частиц) имеют уникальный характер и требуют дорогостоящего оборудования. При этом необходимы максимальное использование полученной информации, извлечение из нее всех возможных выводов с корректным указанием степени доверия к ним, обусловленной методикой проведения эксперимента и используемыми методами обработки.

Расширение области применимости статических методов связано и с внедрением достижений атомной и ядерной физики, таких как использование радиоактивных изотопов для исследования характеристик изучаемых процессов, применение эффекта Мессбауэра и ядерного магнитного резонанса и т. д. в смежных с физикой науках (химии, биологии, геологии, медицине). Объекты, рассматриваемые в ядерной физике, описываются уравнениями квантовой механики, имеющими стохастический характер. Естественно, что при обработке экспериментов, основанных на поведении микрочастиц, использование статистических методов продиктовано самой вероятностной природой причинно-следственных связей. Сами математические

модели для рассматриваемых процессов являются вероятностными. Отметим также широкое распространение радиоэлектронной экспериментальной и регистрирующей аппаратуры; учет приборных искажений от нее основан на теории случайных процессов, анализе временных рядов, теории фильтрации.

Статистическая обработка результатов экспериментов предполагает знание основных понятий и методов теории вероятностей и математической статистики [44; 45; 49; 87]. Выявление характерных классов задач в обработке экспериментальных данных и стандартных методов их решения позволяет выделить обработку результатов экспериментов из многообразия задач и методов прикладной статистики. Так, для рассматриваемого класса спектрометрических экспериментов наиболее распространенным статистическим методом обработки являются метод наименьших квадратов или его заменители (метод наименьших модулей и другие робастные методы). В то же время такие широко применяемые в социальных науках и биологии методы, как дисперсионный или корреляционный анализ, сравнительно мало используются в обработке спектрометрических экспериментов. Кроме того, при статистической обработке результатов экспериментов большее внимание, чем обычно в статистике, уделяется особенностям организации измерений и процесса обработки информации на ЭВМ. Эти особенности приводят к видоизменению или полной замене оптимальных статистических методов при обработке реальных данных.

Обсудим вкратце типы измерений и характер ошибок в них. По отношению к интересующим нас величинам x_1, \dots, x_k измерения могут иметь прямой (измеряются непосредственно x_1, \dots, x_k) и косвенный (измеряются функции $\xi_i(x_1, \dots, x_k)$, $i = \overline{1, l}$, от совокупности интересующих нас величин x_1, \dots, x_k) характер. Ясно, что прямое измерение — частный случай косвенного при

$$\xi_m(x_1, \dots, x_m, \dots, x_k) = \xi_m(x_m) = x_m.$$

Неопределенность экспериментальных данных может быть вызвана различными причинами. В простейшем случае, когда изучаемый объект описывается в рамках лапласовского детерминизма, а неопреде-

ленность в наблюдениях обусловлена инструментальными ошибками и носит аддитивный характер, для наблюдаемых величин y справедливо соотношение

$$y = A(x) + \varepsilon, \quad (1)$$

где A — оператор, описывающий причинно-следственные связи, x — набор факторов (причин), ε — инструментальные искажения. По y определяется либо механизм процесса (оператор A), либо значения ненаблюдаемых факторов (x). В более сложных случаях, например в ядерной физике, сама причинно-следственная модель стохастична; неопределенность, вызванная ошибками измерений, при этом сохраняется. Возможна гораздо более детальная классификация источников ошибок измерений [62].

Ошибки могут иметь как систематический, так и случайный характер. В первом случае искажения возникают всегда, когда имеют место обуславливающие их причины, и обычно приводят к характерным изменениям результатов. Так, систематические ошибки, связанные с наличием «мертвого времени» при регистрации событий в ядерной физике [27], приводят к заниженным результатам подсчета. Наоборот, фоновые эффекты «повышают» значения регистрируемых величин. Что касается случайных ошибок, их действие имеет (при одинаковых условиях эксперимента и процесса измерений) нерегулярный характер, приводит как к завышению, так и к занижению измеряемых величин по сравнению с истинными значениями. Конечно, при обработке результатов экспериментов необходим корректный учет и систематических и случайных ошибок.

Охарактеризуем вкратце важнейшие задачи, возникающие при статистической обработке результатов для типичных классов экспериментов. Прежде всего для применения статистического анализа необходима вероятностная модель регистрируемых величин. Получаемые выводы о природе измеряемых величин могут быть как весьма общими (независимость измерений, аддитивный характер искажений, одинаковость ошибок), так и содержать детальную информацию о статистических свойствах данных (вплоть до закона распределения, полностью определенного или известного с точностью до параметров). Эта проблема не

является чисто математической, требует детального знания изучаемого в эксперименте процесса, методики проведения эксперимента и измерений и обычно решается в тесном сотрудничестве с экспериментаторами.

Даже самые убедительные доводы в пользу конкретной модели измеряемых величин являются неполными, могут приводить к ошибочным заключениям из-за наличия неучитываемых факторов, влияния величин друг на друга через посредство нерассмотренных третьих величин и т. д. Поэтому предлагаемая модель должна быть проверена на экспериментальном материале как на справедливость исходных предпосылок, так и (что особенно важно) на выполнение свойств наблюдений, которые вытекают из модели. Например, в случае задания функции распределения измеряемых величин речь идет о проверке соответствия этой функции экспериментальным данным. Такая проверка выполнима, если имеется достаточно большое число наблюдений (сотни), и позволяет выделить целесообразные методы дальнейшей статистической обработки.

Полученная экспериментальная информация обычно неоднородна по качеству. Наряду с основной массой данных, измеренных с возможно малыми ошибками, имеются грубые промахи в измерениях, вызванные экспериментальными ошибками или сбоями при фиксации результатов, вводе их в ЭВМ. Без анализа данных на качественность, устранения грубых промахов или уменьшения их влияния на результаты последующей обработки можно прийти к ложным выводам об изучаемом объекте. Поскольку визуальный контроль большого объема информации во многих экспериментах невозможен, приведение данных к однородности осуществляется с помощью ЭВМ по статистическим критериям.

Для разумной интерпретации полученной экспериментальной информации необходимо «свернуть» ее так, чтобы потери информации были минимальными, и осуществить редукцию большого количества данных к компактной форме. Это возможно при учете закона распределения данных, выборе соответствующих оптимальных методов. Поэтому процесс свертки экспериментальной информации имеет в основе стати-

стическую теорию оценок параметров. Здесь же возникает вопрос о корректной оценке точности редуцированных данных, об учете их коррелированности. Нередко выбранный метод редукции не является оптимальным с точки зрения статистики, а представляет собой компромисс между использованием эффективных, но вычислительно трудоемких оценок и вычисляемых неитерационными методами, но не сто процентно эффективных оценок.

До сих пор речь шла в основном об учете искажений, имеющих случайный характер. Однако полученная свертка информации еще не пригодна для интерпретации изучаемых явлений из-за наличия систематических аппаратных искажений, фоновых, «загрязняющих» эффектов. Полный учет этих искажений и фонов потребовал бы усилий по созданию их математических моделей и проведению измерений для определения параметров этих моделей, сравнимых по трудоемкости с основным экспериментом и его анализом. Ввиду малости вклада искажений в наблюдаемые величины используют огрубленные модели для систематических искажений, а параметры этих моделей определяют из небольшого числа сопутствующих основному эксперименту измерений [20].

Не касаясь вопросов интерпретации, кроме случая определения параметров модели при заданном функциональном виде, рассмотрим задачи, возникающие после интерпретации. К ним относится прежде всего проблема соответствия предлагаемой модели экспериментальным данным. Нередко она является огрубленной, не вполне адекватной изучаемому процессу. Эта приближенность обнаруживается как сопоставлением результатов интерпретации с фактами соответствующей конкретной науки (например физики), так и статистическими методами. Вычитая из наблюдаемых величин результаты предсказания их в рамках рассматриваемой модели, получаем совокупность остатков, которые могут статистически исследоваться на случайность (либо, напротив, на определенного вида тренды). Если результаты такого сравнения указывают на систематические расхождения между моделью и экспериментом, то модель нуждается в доработке, введении в нее дополнительных факторов, а иногда и в полной замене.

С другой стороны, после проведения обработки и интерпретации экспериментальной информации нередко целесообразно вернуться к постановке эксперимента. Обычно некоторые условия эксперимента (например, число точек измерений и их расположение, энергия, температура) поддаются контролю и управлению. В зависимости от их изменения меняются и величины случайных и систематических искажений, и неточность редуцированных данных. Можно ставить вопрос об условиях эксперимента, обеспечивающих наилучшие по точности редуцированные данные или гарантирующих достижение требуемой точности. Эти проблемы решаются в математической теории планирования эксперимента, а также (особенно в случаях, когда изучаемые зависимости слишком сложны для применения теории планирования эксперимента) с помощью имитационного моделирования на основе метода Монте-Карло.

Обратимся теперь к соотношению эксперимента и обработки его результатов во времени. Часто запланированная серия опытов выполняется до конца, получают весь набор экспериментальных данных, который подвергают обработке после проведения эксперимента (off-line-обработка). Если результаты требуют коррекции условий проведения эксперимента, то меняют значения регулируемых параметров и проводят новую серию наблюдений, которые затем обрабатывают. При этом используют методы, излагаемые в общих курсах статистики [44; 45; 49; 87]. Нередко, однако, обработка наблюдений проводится параллельно с их получением (on-line-обработка), особенно если результаты обработки используются для непрерывного управления экспериментом (в космических исследованиях, в сложных технологических процессах). При этом алгоритмы off-line-обработки нередко неудобны, требуют слишком много времени, что делает невозможным эффективное управление. Например, необходимо вычислить выборочное среднее по выборке объема n наблюдений x_1, \dots, x_n

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (2)$$

В случае off-line-обработки можно непосредственно

использовать (2), что требует $n+1$ операции. Если же проводится on-line-обработка, данные x_i поступают непрерывно; при общем числе наблюдений N вычисление \bar{x} на N шагах требует всего $N^2/2 + 3/2N$ операций. Простое видоизменение (2)

$$\bar{x}^{(n)} = \bar{x}^{(n-1)} + \frac{1}{n} (x_n - \bar{x}^{(n-1)}), \quad (3)$$

где $\bar{x}^{(n)}$ — выборочное среднее по n наблюдениям, снижает (при $N > 3$) число операций до $3 \cdot N$, что уже существенно при обработке телеметрической информации. В более сложных ситуациях выигрыш во времени при использовании рекуррентных методов и алгоритмов еще более значим [62].

Рассматриваемые вопросы статистической обработки тесно связаны с теорией решения некорректно поставленных задач [78; 79]. Во-первых, результаты предварительной обработки экспериментов далее нередко подвергаются интерпретации на основе применения регуляризирующих алгоритмов. Важно согласование статистических свойств, получаемых при обработке величин, с требованиями, накладываемыми этими алгоритмами (малость стандартных отклонений по сравнению с самими величинами, учет коррелированности данных и т. д.). Во-вторых, сами статистические процедуры, например метод наименьших квадратов, приводят к некорректно поставленным задачам [81; 82].

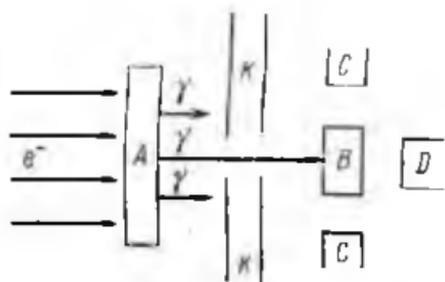
Остановимся в заключение на примере типичного спектрометрического эксперимента и возникающих при этом задач обработки, для которых были применены многие из рассматриваемых далее методов. При нахождении сечений фотонейтронных реакций искомые сечения σ (как функции энергии) связаны косвенным образом с наблюдаемыми величинами [78; 76; 136]:

$$Y(E) = \text{const} \int_0^E W(E, E') \sigma(E') dE', \quad (4)$$

где E — максимальная энергия частиц первичного пучка; ядро интегрального уравнения $W(E, E')$, показывающее распределение налетающих на мишень частиц по энергиям E' , предполагается известным.

Нахождение $\sigma(E')$ по $Y(E)$ из (4) требует решения некорректно поставленной задачи. Разумные результаты получаются лишь при сведении к минимуму систематических и случайных искажений в Y . В противном случае локальные особенности типа перегибов в Y вызовут ложную структурность $\sigma(E')$.

Схема эксперимента вкратце такова (рис. 1): из импульсного ускорителя периодически (50—100 раз в секунду) выводится пучок электронов высокой энергии и направляется на мишень A . Тормозясь в веществе мишени, электроны вызывают излучение γ -квантов, распределение которых по энергиям E' описывается функцией $W(E, E')$. С помощью системы коллиматоров K пучок γ -квантов



сосредоточивается на изучаемом образце B . При этом большинство γ -квантов пронизывает образец без взаимодействия и регистрируется дозиметром D , измеряющим интенсивность облучения образца. Взаимодействие γ -квантов с ядрами мишени приводит к вылету вторичных нейтронов, регистрируемых детекторами C , расположенными в стороне от оси пучка γ -квантов, чтобы исключить регистрацию первичного излучения и его взаимодействие с детектором.

В качестве первого шага необходимо построение стохастической модели эксперимента, так как идеальную картину, описываемую (4), искажает ряд факторов. К ним относится вероятностный (обычно пуассоновский) характер изучаемых в ядерных экспериментах событий. Однако в данном случае при достаточно больших энергиях наряду с изучаемой реакцией (γ, n) происходит и побочная $(\gamma, 2n)$, причем продукты реакций неразличимы при регистрации. Встает задача вероятностного описания множественности и статистического разделения выходов (среднего числа актов реакций на единицу облучения) Y_1 и Y_2 по наблюдениям выхода частиц $Y = Y_1 + 2Y_2$ [8; 18]. Наблюдаемое количество частиц может отличаться от Y также вследствие влияния случайных колебаний ус-

ловий эксперимента, приводящих к плавным искажениям формы кривой $Y(E)$ (аппаратурному дрейфу), наличия систематических искажений: просчетов детекторов, вклада фоновых излучений, вызванных взаимодействием γ -квантов с коллиматором, стенками камеры и т. д. Искажения возможны и при нормировке выхода на дозу облучения, поскольку зависимость чувствительности дозиметра от энергии нелинейна.

При этом возникает весь перечисленный выше комплекс проблем: определение действительного вида распределения числа регистрируемых частиц; выявление «сбоев», вызванных грубыми ошибками при проведении измерений или дешифровке данных, либо их парирование; учет сопутствующих измерений, фоновых эффектов, систематических и случайных аппаратурных искажений; корректная редукция экспериментальных данных и оценка точности и коррелированности результатов свертки; планирование условий эксперимента с целью получения максимальной информации. Интересны также видоизменения методов анализа при on-line-обработке и применение теории фильтрации для решения задачи (4) [40].

Программное обеспечение для многих рассматриваемых ниже алгоритмов статистической обработки имеется в библиотеках программ [9; 57]. Тексты ряда фортранных и алгольных программ опубликованы в [31; 67; 86; 96; 123].

Глава 2

ПРОВЕРКА ЗАКОНА РАСПРЕДЕЛЕНИЯ НАБЛЮДЕНИЙ

Выбор критериев или оценок, имеющих оптимальные свойства, определяется законом распределения наблюдаемых величин. Обычно он не известен точно, но нередко априорные соображения позволяют высказать те или иные гипотезы о его виде. Такого рода гипотезы могут возникать также при рассмотрении эмпирической функции распределения и сравнении ее с распределениями известного вида. Однако подобные соображения требуют статистической про-

верки на достаточно больших совокупностях наблюдений.

Пусть имеется выборка x_1, \dots, x_n наблюдений, взятых из генеральной совокупности с ф.р. $F(x)$. В дальнейшем обычно под $F(x)$ подразумевается непрерывная ф.р. Пусть \mathcal{E} — множество всех непрерывных функций распределения.

Напомним некоторые понятия из теории статистических гипотез [45; 87]. Выборка наблюдений $(x_1, \dots, x_n) = \vec{x}$ трактуется как точка n -мерного пространства R_n . Пусть ω — некоторая область в R_n . Рассматривая ф.р., отвечающую (x_1, \dots, x_n) ,

$$dF(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i) dx_i,$$

можно вычислить вероятность

$$P\{(x_1, \dots, x_n) \in \omega\} = \int_{\omega} dF(x_1, \dots, x_n).$$

Любая гипотеза, связанная с такой вероятностью, называется *статистической*. Приведем примеры статистических гипотез: а) $f(x)$ — ф.п.в. нормального распределения, имеющая заданные математическое ожидание и дисперсию; б) $f(x)$ имеет заданную (например, единичную) дисперсию, а математическое ожидание произвольно; в) $f(x)$ — ф.п.в. нормального распределения. Первые две гипотезы касаются значений параметров распределения известной формы. Такие гипотезы называются *параметрическими*. Гипотеза в) является примером непараметрической гипотезы. Рассмотрим основные понятия теории статистических гипотез на примере параметрических.

Итак, пусть (x_1, \dots, x_n) отвечает ф.р. $F(x_1, \dots, x_n, \vec{\theta})$, где $\vec{\theta} \in \Omega \subset R_k$ — вектор параметров. Пусть ω — подмножество Ω . Обычно $\vec{\theta}_0$, истинное значение вектора $\vec{\theta}$, неизвестно. Назовем *статистической гипотезой* $\mathcal{H} = \mathcal{H}(\omega; \Omega)$ утверждение $\vec{\theta}_0 \in \omega$; его альтернативой служит $\vec{\theta}_0 \in \Omega \setminus \omega$. \mathcal{H} истинна, если $\vec{\theta}_0 \in \omega$, в противном случае \mathcal{H} ложна (иногда говорят, что $\vec{\theta}_0 \in \omega$ есть *нулевая гипотеза* H_0). В случае, когда ω состоит из

одной точки, *гипотеза простая*. Если же ω содержит более одной точки, то *гипотеза сложная*. Так, а) — простая гипотеза, б) — сложная (первая компонента вектора параметров (μ, σ^2) — произвольное число: $-\infty < \mu < \infty$). Пусть, далее, ω_n — множество в R_n , не зависящее от θ и такое, что \mathcal{H} отвергается при $(x_1, \dots, x_n) \in \omega_n$ и принимается в противном случае. Можно вычислить вероятности $P\{(x_1, \dots, x_n) \in \omega_n | \omega\}$ и $P\{(x_1, \dots, x_n) \in \omega_n | \Omega \setminus \omega\}$. Множество ω_n называется *критическим множеством* (критерием) для проверки \mathcal{H} . Выбор ω_n и отклонение \mathcal{H} при $(x_1, \dots, x_n) \in \omega_n$ называется *статистическим критерием* (тестом) для проверки \mathcal{H} . При использовании статистического критерия возможны два вида ошибок: 1) отклонение гипотезы \mathcal{H} , если она истинна (ошибка 1-го рода); 2) принятие гипотезы \mathcal{H} , если она ложна (ошибка 2-го рода). Поскольку принятие или отклонение \mathcal{H} зависит от попадания (x_1, \dots, x_n) в ω_n , величина

$$P\{\omega_n | \vec{\theta}\} = \int_{\omega_n} dF(x_1, \dots, x_n; \vec{\theta})$$

характеризует статистические свойства критерия и является функцией θ с областью определения Ω , принимающей значения из $[0, 1]$. Эта функция называется *мощностью* критерия ω_n , а дополнительная ей $P\{R_n \setminus \omega_n | \vec{\theta}\} = 1 - P\{\omega_n | \vec{\theta}\}$ — *оперативной характеристикой* критерия. Тогда $\alpha = P\{\omega_n | \vec{\theta} \in \omega\}$ и $\beta = 1 - P\{\omega_n | \vec{\theta} \in \Omega \setminus \omega\}$ — вероятности ошибок 1-го и 2-го рода соответственно. Обычная практика проверки гипотез состоит в том, что вероятность ошибок 1-го рода не должна превышать заданного значения α_0 , называемого *размером* (объемом) критерия. Тогда критическая область ω_n зависит от α_0 : $\omega_n = \omega_{n, \alpha_0}$. Желательными свойствами критерия являются его *несмещенность*: $\alpha \leq \alpha_0$, $P\{\omega_n | \vec{\theta} \in \Omega \setminus \omega\} > \alpha_0$, и *состоятельность*: если $\lim_{n \rightarrow \infty} P\{\omega_{n, \alpha_0} | \vec{\theta}\} = 1$, $\vec{\theta} \in \Omega \setminus \omega$, то ω_{n, α_0} — *состоятельный критерий* размера α_0 для гипотезы \mathcal{H} . Множество всех возможных критериев может быть частично упорядочено по следующему пра-

вилу: пусть w_α и w_α^* — два критерия размера α для гипотезы \mathcal{H} и

$$P\{w_\alpha^* | \vec{\theta}\} \leq P\{w_\alpha | \vec{\theta}\}, \vec{\theta} \in \omega; P\{w_\alpha^* | \vec{\theta}\} > P\{w_\alpha | \vec{\theta}\}, \vec{\theta} \in \Omega \setminus \omega.$$

Тогда говорят, что w_α^* равномерно мощнее w_α . Если это правило выполняется для любого (измеримого) подмножества $\omega \subset R_n$, то w_α^* называется *равномерно наиболее мощным* критерием для проверки \mathcal{H} .

К сожалению, равномерно наиболее мощные и несмещенные критерии существуют лишь в относительно несложных случаях; для простой гипотезы против простой альтернативы (лемма Неймана—Пирсона) и в ситуациях, сводимых к указанной. Для сложных гипотез во многих практически важных задачах удается построить асимптотически (при $n \rightarrow \infty$) наиболее мощные критерии исходя из принципа МП, предложенного Фишером [109]. Применение этого принципа при проверке статистических гипотез приводит к критерию ОП [127]. Пусть ФП записывается в виде

$$L(x | \vec{\theta}) = \prod_{i=1}^n f(x_i; \vec{\theta}), \quad (1)$$

где $\vec{\theta} = (\theta_r, \theta_s)$ — вектор $r+s=k$ параметров ($r \geq 1$, $s \geq 0$). Проверяется гипотеза

$$H_0: \vec{\theta}_r = \vec{\theta}_{r0}, \quad (2)$$

являющаяся сложной при $s > 0$, против альтернативы $H_1: \vec{\theta}_r \neq \vec{\theta}_{r0}$. Ищем оценки максимального правдоподобия (МП-оценки) вектора $(\vec{\theta}_r, \vec{\theta}_s)$, дающие безусловный максимум (1), а также МП-оценки для $\vec{\theta}_s$, если справедлива H_0 , т. е. оценки, реализующие условный экстремум функции (1) при условии (2). Пусть $\hat{\theta}_r$, $\hat{\theta}_s$ — оценки в первом случае, а $\hat{\theta}_s'$ — во втором. Рассмотрим ОП

$$0 \leq l = L(x | \vec{\theta}_{r0}, \hat{\theta}_s') / L(x | \hat{\theta}_r, \hat{\theta}_s) \leq 1. \quad (3)$$

Интуитивно ясно (обоснования имеются, например, в [45; 87]), что, если H_0 истинна, значения l должны

быть с большой вероятностью близки к 1. Поэтому критическая область для H_0 имеет вид $l \leq C_\alpha$, где C_α определяется исходя из ф. п. в. $g(l)$ случайной величины так, чтобы

$$\int_0^{C_\alpha} g(l) dl = \alpha.$$

Понятия, введенные для параметрических гипотез, переносятся и на случай непараметрических гипотез. Различия имеют место лишь в формулировке проверяемых гипотез. Так, для задачи проверки вида распределения проверяется следующая гипотеза: (x_1, \dots, \dots, x_n) имеют ф. р. $F_0(x)$. Эта функция может быть известна полностью (простая гипотеза), а может быть задана лишь ее форма при неизвестных параметрах (сложная гипотеза, см. пример в)). Альтернативой служит утверждение $F(x) \in \mathcal{E} \setminus F_0(x)$.

Ясно, что такая гипотеза, вообще говоря, не может быть проверена, поскольку число наблюдений x_1, \dots, x_n , по которому делается заключение о значениях непрерывной функции, определяемой счетным множеством значений, конечно. В действительности сформулированная выше гипотеза заменяется некоторой другой, близкой ей и допускающей проверку по выборке. Однако при такой подмене всегда остается указанная неопределенность, в чем мы убедимся в дальнейшем.

Простейший критерий для проверки нашей гипотезы, тесно связанный с ОП, — критерий χ^2 Пирсона [33; 45]. Пусть область значений величины x произвольным образом разбита на k непересекающихся классов G_1, \dots, G_k . Проверяется простая гипотеза $H_0: x \sim F_0(x)$. Тогда можно вычислить вероятности

$$p_{0i} = P\{x \in G_i\}, \quad i = \overline{1, k}, \quad \sum_{i=1}^k p_{0i} = 1.$$

Рассмотрим число попаданий r_1, \dots, r_k наблюдений x_1, \dots, x_n в классы G_1, \dots, G_k ($\sum_{i=1}^k r_i = n$). Очевидно,

что если справедлива H_0 , то r_1, \dots, r_k имеют мультиномиальное распределение $M(n; p_{01}, \dots, p_{0k-1})$, для

которого [44; 87] $Er_i = np_{0i}$, $Dr_i = np_{0i}(1 - p_{0i})$, $i = 1, k-1$. Тогда

$$L(r_1, \dots, r_k | H_0) = n! \sum_{i=1}^k \frac{p_{0i}^{r_i}}{r_i!}$$

Альтернативой H_1 служит любая $F_1(x)$, для которой найдется хотя бы один номер i такой, что

$$p_{1i} = \int_{G_i} dF_1(x) \neq p_{0i}.$$

Таким образом, в действительности речь идет о возможности различить распределения, имеющие заданные вероятности p_{0i} на G_1, \dots, G_k , и распределения, у которых вероятности $P\{x \in G_i\}$ другие. $F_1(x)$ соответствует

$$L(r_1, \dots, r_k | H_1) = n! \prod_{i=1}^k \frac{p_{1i}^{r_i}}{r_i!} \quad (4)$$

Найдем безусловный максимум для ФП (4); для упрощения выкладок перейдем к логарифмам:

$$\ln L(r_1, \dots, r_k | H_1) = \ln \left(n! / \prod_{i=1}^k r_i! \right) + \sum_{i=1}^k r_i \ln p_{1i}.$$

Подставляя $p_{1k} = 1 - p_{11} - \dots - p_{1k-1}$ и дифференцируя по p_{1i} , $i = 1, k-1$, получим

$$\frac{\partial \ln L}{\partial p_{1i}} = \frac{r_i}{p_{1i}} - \frac{r_k}{p_{1k}} = 0,$$

откуда p_{1i} пропорционально r_i в точке экстремума.

Вспомнив, что $\sum_{i=1}^k p_{1i} = 1$, получим $p_{1i} = \frac{r_i}{n}$. Не-

трудно проверить, что это точка максимума (устремить p_{1i} к 0 или к 1). Значение ФП в точке максимума

$$L_{\max} = \frac{n!}{n^n} \prod_{i=1}^k \frac{r_i^{r_i}}{r_i!}.$$

Составим ОП, используя обозначение $np_{0i} \doteq a_i$:

$$l = \frac{L(r_1, \dots, r_k | H_0)}{L_{\max}} = \prod_{i=1}^k \left(\frac{np_{0i}}{r_i} \right)^{r_i} = \prod_{i=1}^k \left(\frac{a_i}{r_i} \right)^{r_i}. \quad (5)$$

Логарифмируя (5), получим

$$\begin{aligned} \ln l &= \sum_{i=1}^k r_i \ln \left(\frac{a_i}{r_i} \right) = \\ &= - \sum_{i=1}^k [(r_i - a_i) + a_i] \ln \left[1 + \frac{r_i - a_i}{a_i} \right]. \end{aligned} \quad (6)$$

Если H_0 истинна, то $Er_i = a_i$, $\sigma r_i = \sqrt{a_i(1-p_{0i})} \approx \sqrt{a_i}$, т. е. $r_i - a_i \sim \sqrt{a_i} \sim \sqrt{n}$. Поэтому можно разложить логарифм в правой части (6) по степеням $n^{-1/2}$ и, учитывая, что $\sum_{i=1}^k (r_i - a_i) = 0$, получить

$$\begin{aligned} \sum_{i=1}^k [(r_i - a_i) + a_i] \left[\frac{r_i - a_i}{a_i} - \frac{(r_i - a_i)^2}{2a_i^2} + \right. \\ \left. + \frac{(r_i - a_i)^3}{3a_i^3} - \dots \right] = - \frac{1}{2} \sum_{i=1}^k \frac{(r_i - a_i)^2}{a_i} + O(n^{-1/2}). \end{aligned}$$

Главный член здесь записывается в виде

$$- \frac{1}{2} \sum_{i=1}^k \frac{(r_i - np_{0i})^2}{np_{0i}}.$$

Поэтому статистика ОП l эквивалентна использованной Пирсоном [128] статистике

$$X^2 = \sum_{i=1}^k \frac{(r_i - np_{0i})^2}{np_{0i}}. \quad (7)$$

Поскольку малым критическим значениям ОП l соответствуют большие значения X^2 , то гипотеза H_0 отвергается при $X^2 > X^2_{n, \alpha}$. Статистика (7) проще для

вычислений и анализа, чем (5), и используется чаще.

При малых n распределение статистики (7) зависит от вида гипотетической ф.р. $F_0(x)$, однако при большом объеме выборки распределение (7) в случае справедливости H_0 свободно от проверяемого распределения $F_0(x)$. Рассмотрим вопрос о выборе критического значения $X^2_{n,\alpha}$ при большом n . Для этого нужно вспомнить ряд фактов из общего курса теории вероятностей [44; 87].

Утверждение 1. Если $r = (r_1, \dots, r_k)$ имеет мультиномиальное распределение $M(n; p_1, \dots, p_k)$, то $E r_i = n p_i$, $\text{cov}(r_i, r_j) = n p_i \delta_{ij} - n p_i p_j$.

Утверждение 2. Пусть x_1, \dots, x_n — выборка k -мерных случайных величин, $z = (z^{(1)}, \dots, z^{(k)})$ — вектор выборочных сумм x_1, \dots, x_n по компонентам $x^{(1)}, \dots, x^{(k)}$, $\vec{\mu} = E x$ конечно и ковариационная матрица $\text{cov}(x, x) = \Sigma = \|\sigma_{ij}\| > 0$. Тогда z имеет асимптотически нормальное распределение $N(n\vec{\mu}, n\Sigma)$, а $\vec{\eta} = (z - n\vec{\mu})/\sqrt{n}$ асимптотически $N(0, \Sigma)$.

Утверждение 3. Если выполнены условия из утверждения 2, то квадратичная форма $Q_n = \vec{\eta}^T \Sigma^{-1} \vec{\eta}$ при $n \rightarrow \infty$ сходится по распределению к χ^2_k .

Если теперь рассматривать величины r_1, \dots, r_k как суммы n случайных величин τ_1, \dots, τ_n , где τ_i относится к i -му испытанию, то согласно утверждению 1 $E \tau_i = p_i$, $\text{cov}(\tau_i, \tau_j) = p_i \delta_{ij} - p_i p_j$ и, как легко показать перемножением матриц,

$$\|\text{cov}(\tau_i, \tau_j)\|^{-1} = \|p_i \delta_{ij} - p_i p_j\|^{-1} = \left\| \frac{\delta_{ij}}{p_i} + \frac{1}{p_{k+1}} \right\|,$$

где $p_{k+1} = 1 - \sum_{i=1}^k p_i$. Применив к r_1, \dots, r_k утверждения 2 и 3, нетрудно найти, что

$$\begin{aligned} Q_n &= \sum_{i=1}^{k+1} \frac{(r_i - n p_i)^2}{n p_i} = \\ &= \frac{(\vec{r} - E \vec{r})^T}{\sqrt{n}} \|\text{cov}(\tau_i, \tau_j)\|^{-1} \frac{(\vec{r} - E \vec{r})}{\sqrt{n}}, \end{aligned}$$

где $r_{k+1} = n - \sum_{i=1}^k r_i$, при $n \rightarrow \infty$ имеет асимптотически распределение χ^2_k .

Возвращаясь к распределению статистики (7), видим, что в случае справедливости H_0 при $n \rightarrow \infty$ величина X^2 асимптотически имеет распределение X^2_{k-1} . Поэтому критическое значение выбирается как α -процентная точка распределения χ^2_{k-1} , т. е. как решение уравнения

$$\int_{\chi^2_{n,\alpha}}^{\infty} d\chi^2_{k-1} = \alpha.$$

Иначе говоря, при больших n величина $X^2_{n,\alpha}$ в действительности не зависит от n , а вполне определяется значением k и α : $X^2_{n,\alpha} = X^2_{\alpha}$.

Рассмотрим теперь вопрос о состоятельности критерия χ^2 . Вообще говоря, она вытекает из состоятельности критерия ОП [45; 87] и доказанной эквивалентности критериев (5) и (7). Однако ее можно доказать и непосредственно. Пусть ω_{α} — множество точек выборочного пространства, для которых при истинности H_0 $X^2 > X^2_{\alpha}$. Пусть в действительности справедлива H_1 , т. е. существует i такое, что $p_{1i} = \int_{G_i} dF_1(x) \neq p_{0i}$. Из центральной предельной теоремы

(утверждение 2) следует, что при $n \rightarrow \infty (r_1/n, \dots, r_k/n) \xrightarrow{P} (p_{11}, \dots, p_{1k})$. Тогда можно рассмотреть при заданном n множество E_n точек (x_1, \dots, x_n) , для которых

$$\sum_{i=1}^k \left(\frac{r_i}{n} - p_{1i} \right)^2 < D^2, \quad (8)$$

где $0 < D^2 < \rho^2(\vec{p}_0, \vec{p}_1)$. Условие $X^2 > X^2_{\alpha}$ для статистики (7) можно переписать в виде

$$\sum_{i=1}^k \frac{1}{p_{0i}} \left(\frac{r_i}{n} - p_{0i} \right)^2 > \frac{X^2_{\alpha}}{n}.$$

Ясно, что найдется такое n_1 , что при $n > n_1$ множество $E_n \subset \omega_\alpha$. Поэтому при $n > n_1$ $P\{E_n | F_1\} \leq P\{\omega_\alpha | F_1\}$. Но по определению сходимости по вероятности для любого $\delta > 0$ найдется n_2 такое, что (8) выполняется с вероятностью $P\{E_n | F_1\} > 1 - \delta$. Теперь, взяв $n > \max(n_1, n_2)$, получим для него $P\{\omega_\alpha | F_1\} > 1 - \delta$, т. е. при $n \rightarrow \infty$ $\lim P\{\omega_\alpha | F_1\} = 1$, что и доказывает состоятельность критерия ω_α для проверки гипотезы H_0 . Однако ясно, что на самом деле статистика (7) позволяет проверить гипотезу о принадлежности $F_0(x)$ множеству \mathcal{E}_0 , где \mathcal{E}_0 состоит из всех ф. р. (непрерывных), имеющих для данного разбиения G_1, \dots, G_k одинаковые вероятности $P\{x \in G_i\} = p_{0i}$, $i = 1, k$. Для того чтобы различить распределения внутри \mathcal{E}_0 , критерий χ^2 , конечно, несостоятелен. Поэтому приходится изменять границы классов G_i , либо число k таких классов.

Пусть теперь гипотетическое распределение $F_0(x)$ зависит от параметров $\bar{\theta} = (\theta_1, \dots, \theta_s)$, значения которых неизвестны и подлежат оценке по выборке, т. е. мы рассматриваем случай сложной гипотезы. Тогда величины $p_{0i}(\bar{\theta})$ (здесь $\hat{\theta}$ — оценка параметров $\bar{\theta}$) сами являются случайными и приведенное выше обоснование асимптотического распределения χ^2 не годится. В частности, $E(r_i - np_{0i}(\hat{\theta})) \neq 0$.

Однако можно доказать [45], что если оценки $\hat{\theta}$ ищутся из (7) методом минимума χ^2 [49] или как МП-оценки для мультиномиального распределения $M(n; p_{01}(\hat{\theta}), \dots, p_{0k}(\hat{\theta}))$, то это эквивалентно наложению s линейных связей на (асимптотически) квадратичную форму от стандартных нормальных величин так, что (7) асимптотически распределена как χ^2_{k-1-s} и условие отклонения H_0 есть $\chi^2 > \bar{\chi}^2_\alpha$, где $\bar{\chi}^2_\alpha$ — решение уравнения

$$\int_{\bar{\chi}^2_\alpha}^{\infty} d\chi^2_{k-1-s} = \alpha.$$

Если же в качестве $\hat{\theta}$ используются обычные МП-оценки, то при $n \rightarrow \infty$ асимптотическое распределение (7) точно неизвестно, но показано, что его процент-

ные точки заключены между X^2_{α} и X^2_{α} . Поэтому для надежности рекомендуется использовать X^2_{α} [45; 66].

Ниже приведены другие рекомендации по использованию критерия χ^2 .

1. Поскольку критическое значение X^2_{α} выбирается исходя из асимптотического распределения (7), то критерий следует использовать при достаточно большом n (обычно порядка сотен).

2. Число классов разбиения k и границы классов выбираются так, чтобы np_{0i} не были слишком малы. Обычно считается [45], что $np_{0i} \geq 5$, однако Кокрэн [106; 107] показал, что если $k \geq 7$ и $\alpha = 0,05$ или $0,01$, то допустимо уменьшение одной или двух np_{0i} до 1.

3. В случае простой гипотезы доказана локальная несмещенность критерия χ^2 , если выбрано равновероятное разбиение области значений x , т. е. если $p_{01} = p_{02} = \dots = p_{0i} = \dots = p_{0k} = k^{-1}$. Поэтому рекомендуется использовать такое равновероятное разбиение. Кроме того, для этого случая, как показано Слэктором [134], аппроксимация χ^2 -распределением хороша даже при $np_{0i} = \frac{n}{k} \approx 1$.

Интуитивно ясным недостатком критерия χ^2 является потеря информации от группировки значений непрерывной с. в. в классы G_i . Был разработан ряд критериев, свободных от этого недостатка. Наиболее важен и часто применяем критерий Колмогорова — Смирнова [122]. Он основан на сравнении гипотетической ф. р. $F_0(x)$ с эмпирической

$$F_n(x) = \begin{cases} 0, & x < x_{(1)}, \\ r/n, & x_{(r)} \leq x < x_{(r+1)}, \\ 1, & x_{(n)} \leq x, \end{cases}$$

где $x_{(1)} \leq \dots \leq x_{(r)} \leq \dots \leq x_{(n)}$ — члены вариационного ряда, построенного по x_1, \dots, x_n . Используется мера расхождения ф. р.

$$D_n = \sup_x |F_n(x) - F(x)|. \quad (9)$$

Применимость этого критерия, как и некоторых других, основанных на сравнении эмпирической ф. р. с гипотетической $F_0(x)$ [45; 48], основывается на уси-

ленном законе больших чисел [55]: при истинности H_0 $P\{\lim_{n \rightarrow \infty} F_n(x) = F_0(x)\} = 1$. Статистика D_n не зависит от вида распределения $F_0(x)$, если проверяемая гипотеза простая [45]. Это следует из того, что любое непрерывное распределение $F_0(x)$ можно преобразовать в любое другое, используя подходящее преобразование аргумента (оно интерпретируется как деформация оси абсцисс), и при таком преобразовании максимальное расстояние между двумя ф. р. по вертикали D_n остается неизменным. Асимптотическое распределение $\sqrt{n}D_n$ известно [122] и заетабулировано [12] для случая простой гипотезы. Гипотеза H_0 отвергается, если $D_n > d_{n, \alpha}$; критическое значение выбирается из таблиц или вычисляется стандартными программами на ЭВМ.

Если же распределение $F_0(x)$ зависит от неизвестных параметров, подлежащих оценке по выборке, т. е. гипотеза H_0 сложная, то в общем случае ничего не известно о распределении (9), даже асимптотическом. Однако в некоторых частных случаях применение D_n при сложных гипотезах изучалось. Так, показано, что при проверке нормальности распределения достаточно использовать при вычислении $d_{n, \alpha}$ поправочные множители, зависящие от n и α , но не зависящие от математического ожидания и дисперсии.

В заключение рассмотрим простые критерии для проверки нормальности, поскольку нормальность чаще всего предполагается для распределения наблюдаемых величин $\text{arg} \text{ог} \text{г}$. Известно [44; 87], что характеристическим свойством нормального распределения является равенство пулю семиинвариантов κ_r , $r > 2$. Поэтому возникает идея использовать выборочные аналоги младших семиинвариантов, чтобы влияние статистического разброса было минимальным. Обычно используются κ_3 и κ_4 , связанные с центральными моментами следующими соотношениями:

$$\kappa_3 = \mu_3, \quad \kappa_4 = \mu_4 - 3\mu_2^2.$$

Таким образом, для с. в. $\xi \sim N(0, \sigma^2)$ рассматриваются асимметрия $\beta_1 = \mu_3/\sigma^3$ и эксцесс $\beta_2 = (\mu_4 - 3\sigma^4)/\sigma^4 = \mu_4/\sigma^4 - 3$, не зависящие от выбранного масштаба измерений с. в. ξ . Их выборочные аналоги

$$b_1 = m_3/s^3, \quad b_2 = m_4/s^4 - 3 \quad (10)$$

$$(s^2 \text{ — выборочная дисперсия: } s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2)$$

не зависят от σ и распределение их затабулировано [12] для стандартной нормальной величины. Поскольку β_1 и β_2 равны 0 при нормальном распределении, использование (10) приводит к критериям

$$w_{\alpha}^{(1)} : |b_1| > d_1, \quad w_{\alpha}^{(2)} : b_2 > d_2, \quad (11)$$

значения d_1 и d_2 для заданного размера α находят из таблиц. Гипотеза H_0 отвергается, если выполнено хотя бы одно из неравенств (11). Эти критерии нередко называют критериями нормальности, однако фактически проверяется равенство нулю μ_3 и μ_4 , в то время как вид ф. р. определяется всей совокупностью семинвариантов или моментов [44].

Применение критериев проверки нормальности разумно при наличии достаточно представительных выборок (несколько сотен или тысяч наблюдений), поскольку статистики b_1 и особенно b_2 очень чувствительны к отклонениям выборочных значений от среднего \bar{x} и при небольших и средних объемах выборок результат применения (11) в значительной степени случаен. Вообще, если обработка наблюдений ведется на мощных ЭВМ (БЭСМ-6 или ЕС ЭВМ), снабженных соответствующим программным обеспечением, лучше использовать не (11), а критерии согласия χ^2 или Колмогорова—Смирнова. Необходимые для этого программы имеются на БЭСМ-6 в составе библиотек SSP [67] и БОИРЭ [9], а на ЕС — в библиотеке БИМ [57].

Глава 3

ВЫЯВЛЕНИЕ АНОМАЛЬНЫХ НАБЛЮДЕНИЙ В ОДНОМЕРНЫХ ВЫБОРКАХ

При обработке результатов экспериментов нередко возникает задача проверки качества информации, анализа принадлежности всех наблюдений к одной генеральной совокупности или, напротив, вывода о том, что имело место «засорение» выборки. Обычно неизвестно, какие наблюдения не относятся к генеральной совокупности, характеризующей основную

часть выборки, и являются аномальными. Однако некоторые из наблюдений имеют слишком экстремальные (высокие или низкие, или и то и другое) значения, чтобы быть в согласии с предположением, что они порождены тем же распределением. Требуются объективные критерии, чтобы проверить основательность этого подозрения. Особенно важно иметь такие критерии, если обработка ведется на ЭВМ автоматизированно, ибо при этом отпадают такие факторы, как опыт, интуиция исследователя.

Обнаружение аномальных наблюдений может производиться для того, чтобы: а) провести сортировку данных перед их анализом и либо полностью исключить аномальные наблюдения из рассмотрения, либо придавать им меньшее доверие (вес) при дальнейшем анализе; б) выявить наличие аномальных наблюдений и тем самым указать на необходимость более тщательного проведения эксперимента и измерений; в) иногда именно аномальные наблюдения представляют особый интерес и даже являются целью эксперимента (например, след новой элементарной частицы на фотопленке).

Причины появления резко выделяющихся наблюдений могут быть различными: наличие сопутствующих по отношению к изучаемому «фоновых» эффектов; нечистота изучаемого объекта, наличие примесей; резкие изменения условий эксперимента (скажем, напряжения питания приборов); «сбои» при регистрации (например, флуктуации шума в электронной аппаратуре). Возможно также появление грубых промахов при записи показаний приборов, вводе в ЭВМ и т. д.

Чтобы выявлять аномальные наблюдения, нужно знать, что же является нормой. Иными словами, должен быть известен закон распределения основной массы наблюдений. Далее обычно вычисляют некоторые характеристики с. в. с заданным законом распределения и исследуют значимость отличий выборочных величин от теоретических.

Рассмотрим пример многократного бросания монеты. Предполагается биномиальный закон распределения с $p = 1/2$. Пусть осуществлено по 900 испытаний с 10 монетами и для одной из них зафиксировано 510 появлений «орла». Насколько это согласуется с

гипотезой о равновероятности выпадения лицевой и реверсной сторон монеты? Для 900 испытаний при $p=1/2$ м. о. числа появлений «орла» есть $E\xi=np=450$; дисперсия $D\xi=np(1-p)=225$, поэтому стандартное отклонение $\sigma=\sqrt{D\xi}=15$. Поскольку n достаточно велико, биномиальное распределение удовлетворительно описывается нормальным законом $N(450, 15)$. В случае «подозрительной» монеты отклонение от $E\xi$ составляет $60=4\sigma$. Вероятность такого события для нормально распределенной с. в. равна 0,00013. Всего было проведено 10 независимых серий испытаний, поэтому вероятность появления такого отклонения хотя бы в одной серии есть $1 - (1 - 0,00013)^{10} \approx 1 - 1 + 0,0013 = 0,0013$. Следовательно, с вероятностью, большей 99,8%, можно утверждать, что подозреваемая серия является аномальной.

Большинство критериев обнаружения аномальных наблюдений предполагает нормальность генеральной совокупности. Это связано как с тем, что статистическая теория наиболее разработана именно для нормального закона, так и с некоторыми априорными соображениями о статистических свойствах наблюдаемых величин. Измерительные погрешности трактуются как сумма большого числа «элементарных», примерно одинаково распределенных ошибок, что дает согласно центральной предельной теореме асимптотическую нормальность суммарной погрешности. Нередко условия проведения эксперимента таковы, что закон распределения, отличный от нормального, становится близким к нормальному при выбранных параметрах эксперимента.

Так, в ядерной физике число частиц, образующихся в результате реакции, подчиняется закону Пуассона (это связано с малой вероятностью реакции для каждого ядра, независимостью процессов для различных ядер и большим их числом в исследуемом образце) $P_0(\mu)$, где μ — среднее число образующихся частиц. Часто, чтобы добиться высокой величины отношения эффект — фон, а также ускорить набор необходимой статистики, измерения ведут при высокой интенсивности μ . Нетрудно показать, что при $\mu \rightarrow \infty$ $P_0(\mu)$ сходится к $N(\mu, \sqrt{\mu})$. Действительно х. ф. с. в. $\xi \sim P_0(\mu)$ равна $\varphi_\xi(t) = E(e^{it\xi}) = \exp\{\mu(e^{it} - 1)\}$.

Рассмотрим нормированную и центрированную с. в. $\tau = (\xi - \mu) / \sqrt{\mu}$, ее х. ф. $\varphi_{\tau}(t) = \exp\{-it\sqrt{\mu}\} \cdot \varphi_{\xi}(t/\sqrt{\mu}) = \exp\{-it\sqrt{\mu} + \mu[1 + it/\sqrt{\mu} - t^2/2\mu + O(\mu^{-3/2})] - \mu\} = \exp\{-it\sqrt{\mu} + \mu[1 + it/\sqrt{\mu} - t^2/2 + O(\mu^{-1/2})]\}$. Устремив μ к бесконечности, получим $\varphi_{\tau}(t) \rightarrow \exp\{-t^2/2\}$, т. е. τ сходится по распределению [44] к нормальной с. в. $N(0,1)$. Поэтому исходная величина $\xi = \tau\sqrt{\mu} + \mu$ сходится по распределению к нормальной $N(\mu, \sqrt{\mu})$, что и утверждалось.

Подчеркнем, однако, что эти априорные соображения являются лишь наводящими. Например, пуассоновский закон регистрируемых событий может искажаться просчетами регистрирующей аппаратуры. Гипотеза об «элементарных» ошибках трудно проверяема. В любом случае необходимо проверять гипотезу о нормальности по статистическим критериям, используя экспериментальные данные.

Поставим задачу выявления аномальных наблюдений более строго: пусть x_1, \dots, x_n — наблюдения, которые независимы и распределены по нормальному закону. Согласно Фергюсону [108], возможны две общие модели аномальных наблюдений. В модели А (гипотеза отличия в сдвиге) предполагается, что $n - k$ величин имеют одинаковое распределение $N(\mu, \sigma)$, а оставшиеся k (заранее неизвестно, какие именно) имеют математическое ожидание $\mu + \lambda_i \sigma$ ($i = \overline{1, k}$) и общую дисперсию σ^2 . Модель Б (гипотеза отличия в масштабе) предполагает, что k с. в. имеют закон распределения $N(\mu, \lambda'_i \sigma)$, $\lambda'_i > 1$. Нулевой гипотезой в каждой модели является $k = 0$.

Поскольку аномальное наблюдение трактуется как экстремальное по величине, то многие критерии их выявления используют порядковые статистики [33; 87] и их распределения. Другие критерии основаны на вычислении величин, чувствительных к отклонениям от нормальности. Так, рассмотренные выше выборочная асимметрия и выборочный эксцесс, зависящие соответственно от третьей и четвертой степеней величин вида $(x_i - \bar{x})$, подобным свойством обладают и используются при анализе аномалий.

Даже ограничивая рассмотрение критериев выявления аномальных наблюдений случаем выборок из нормальной совокупности, можно видеть, что возмож-

ны различные постановки задачи в зависимости от информации о параметрах распределения. Пусть $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ — вариационный ряд для выборки x_1, \dots, x_n , $\omega_n = x_{(n)} - x_{(1)}$ — выборочный размах. Следуя [35], укажем часто применяемые статистики для вычисления аномальных наблюдений и ситуации, в которых они применяются:

а) известны μ и σ ;

$$A_1 = (x_{(n)} - \mu)/\sigma, \quad A_2 = \max_i |x_i - \mu|/\sigma,$$

$$A_3 = \chi_n^2 = \sum (x_i - \mu)^2/\sigma^2;$$

б) известно только σ ;

$$B_1 = (x_{(n)} - \bar{x})/\sigma, \quad B_2 = \max_i |x_i - \bar{x}|/\sigma,$$

$$B_3 = (x_{(n)} - x_{(n-1)})/\sigma, \quad B_4 = \omega_n/\sigma,$$

$$B_5 = \chi_{n-1}^2 = \sum (x_i - \bar{x})^2/\sigma^2;$$

в) значения μ и σ неизвестны, но имеются независимые оценки s_v для σ^2 ;

$$C_1 = (x_{(n)} - \bar{x})/s_v, \quad C_2 = \max_i |x_i - \bar{x}|/s_v,$$

$$C_3 = \omega_n/s_v,$$

$$C_4 = \frac{x_{(n)} - \bar{x}}{(\sum (x_i - \bar{x})^2 + v s_v^2)^{1/2}}, \quad C_5 = \frac{\max_i |x_i - \bar{x}|}{(\sum (x_i - \bar{x})^2 + v s_v^2)^{1/2}};$$

г) μ и σ неизвестны;

$$D_1 = (x_{(n)} - \bar{x})/s, \quad D_2 = \max_i |x_i - \bar{x}|/s,$$

$$D_3 = \omega_n/s,$$

$$D_4 = \sqrt{n} \frac{\sum (x_i - \bar{x})^3}{s^3}, \quad |D_4|, \quad D_5 = \frac{n \sum (x_i - \bar{x})^4}{s^4},$$

$$D_6 = \frac{1}{s^2} \sum_{i=1}^{n-2} (x_i - \bar{x}_{n,n-1})^2,$$

$$\bar{x}_{n,n-1} = \frac{1}{n-2} \sum_{i=1}^{n-2} x_{(i)}.$$

Процентные точки многих из этих и других статистик приводятся в [12; 33]. Обращаясь к применению этих статистик, отметим, что стоящие слева статистики главным образом односторонние, позволяющие обнаружить аномально большие наблюдения, а справа указаны статистики двухсторонних критериев, обнаруживающих и слишком большие, и слишком малые значения наблюдаемых величин. Статистика A_2 фактически приводит к проверке неравенства

$$\max_i |x_i - \mu| < \gamma_\alpha \cdot \sigma, \quad (1)$$

где γ_α выбирается для данного уровня значимости из таблиц нормального распределения аналогично тому, как это сделано выше для примера с бросанием монет, т. е. как решение уравнения

$$2 \frac{1}{\sqrt{2\pi}} \int_{\gamma_\alpha}^{\infty} e^{-t^2/2} dt = \frac{\alpha}{n}.$$

Эта статистика издавна использовалась экспериментаторами в форме «правила трех сигм»: поскольку для нормальной с. в. вероятность отклонения на 3σ от μ около 0,001. Отклонения, превышающие эту границу (при числе наблюдений порядка нескольких десятков), характеризуются как аномальные, и соответствующие наблюдения исключаются из рассмотрения с основной массой данных и анализируются особо. Однако поскольку такой подход негибок (при сотнях и тысячах наблюдений большие отклонения не маловероятны) и обычно имеет место ситуация г) неизвестного σ , более интересны другие критерии.

Показано [108], что при неизвестных μ и σ критерий, основанный на D_4 , является в модели А локально наиболее мощным против односторонней альтернативы $H_1: \lambda\sigma = \Delta > 0$. Для двухсторонней альтернативы $H'_1: \Delta \neq 0$ локально наиболее мощным является критерий D_5 , основанный на выборочном эксцессе. Однако эти локально оптимальные свойства не очень важны для практики, поскольку выявляются наблюдения, резко выделяющиеся, т. е. Δ сильно отличается от 0. Показано, что в модели А с не более чем одним аномальным наблюдением вероятность правильно его выявить максимизируется, если использовать стьюден-

тизированное экстремальное отклонение D_2 (D_1 для односторонней альтернативы $\Delta > 0$, $D'_1 = (\bar{x} - x_{(1)})/s$ для $\Delta < 0$) [45]. Оптимальность этого критерия доказана и в модели Б при наличии не более одного аномального наблюдения [108]. Однако в случае двух выбросов, имеющих распределение $N(\mu + \lambda\sigma, \sigma)$, при $\lambda \rightarrow \infty$ вероятность обнаружения их стремится к нулю («маскирующий эффект») [35].

Рассмотрим подробнее статистику D_2 . Оценим вероятность события $\mathcal{A} = \{D_2 \geq z\}$, если справедлива нулевая гипотеза H_0 в модели А. Пусть событие

$$\mathcal{A}_\alpha = \left\{ \frac{|x_\alpha - \bar{x}|}{s^*} \geq z \right\}, \quad \alpha = \overline{1, n}, \quad (s^*)^2 = \frac{n-1}{n} s^2. \quad (2)$$

Поскольку \mathcal{A} — объединение событий \mathcal{A}_α (возможно и пересекающихся), то в силу формулы Бонферрони

$$\begin{aligned} P\{\mathcal{A}\} &= P\left\{ \bigcup_{\alpha=1}^n \mathcal{A}_\alpha \right\} = \sum_{\alpha=1}^n P\{\mathcal{A}_\alpha\} - \sum_{\alpha=1}^{n-1} \sum_{\beta=\alpha+1}^n P\{\mathcal{A}_\alpha \cap \\ &\cap \mathcal{A}_\beta\} + \sum_{\alpha=1}^{n-2} \sum_{\beta=\alpha+1}^{n-1} \sum_{\delta=\beta+1}^n P\{\mathcal{A}_\alpha \cap \mathcal{A}_\beta \cap \mathcal{A}_\delta\} - \dots \\ &\dots + (-1)^{n-1} P\left\{ \bigcap_{\nu=1}^n \mathcal{A}_\nu \right\} \end{aligned} \quad (3)$$

из (3) следует неравенство

$$\begin{aligned} \sum_{\alpha=1}^n P\{\mathcal{A}_\alpha\} - \sum_{\alpha=1}^{n-1} \sum_{\beta=\alpha+1}^n P\{\mathcal{A}_\alpha \cap \mathcal{A}_\beta\} &\leq P\{\mathcal{A}\} \leq \\ &\leq \sum_{\alpha=1}^n P\{\mathcal{A}_\alpha\}. \end{aligned} \quad (4)$$

Поскольку все стоящие в (2) величины \mathcal{A}_α распределены одинаково, получим

$$P\{D_2 \geq z\} \leq nP\{\xi_1 \geq z\}, \quad (5)$$

где $\xi_1 = |x_1 - \bar{x}|/s^*$. При этом левая часть неравенства (4) позволяет оценить погрешность приближе-

ния (5): можно показать [11; 13], что она имеет порядок ε^2 . Преобразуем

$$\xi_1 = \frac{\left| x_1 - \frac{1}{n} \sum x_i \right|}{\sqrt{\frac{1}{n} \sum \left(x_i - \frac{1}{n} \sum x_i \right)^2}},$$

вводя нормированные величины $z_\alpha = (x_\alpha - \mu)/\sigma$. Тогда

$$\xi_1 = \frac{\left| z_1 - \frac{1}{n} \sum z_i \right|}{n^{-1/2} \sqrt{\sum z_i^2 - \frac{1}{n} (\sum z_i)^2}}. \quad (6)$$

В знаменателе (6) стоит квадратичная форма с матрицей

$$I - \frac{1}{n} E = Q = \|q_{\alpha\beta}\| = \begin{cases} 1 - \frac{1}{n} & \text{при } \alpha = \beta, \\ -\frac{1}{n} & \text{при } \alpha \neq \beta, \end{cases}$$

где $E_{\alpha\beta} = 1$, $\alpha, \beta = \overline{1, n}$, I — единичная матрица.

Подвергнем z_α , $\alpha = \overline{1, n}$, ортогональному преобразованию. Пусть $z_\alpha = \sum_{\beta=1}^n r_{\alpha\beta} y_\beta$, где матрица $R = \|r_{\alpha\beta}\|$ имеет следующий вид: ее первый столбец $r_{\alpha 1} = (n^{-1/2}, \dots, n^{-1/2})^T$, $\alpha = \overline{1, n}$, а остальные

$$r_{\alpha\gamma} = \left(\underbrace{0, \dots, 0}_{\gamma-2}, \sqrt{\frac{n-\gamma+1}{n-\gamma+2}}, \right. \\ \left. \frac{1}{\sqrt{(n-\gamma+1)(n-\gamma+2)}}, \dots, \dots, \frac{1}{\sqrt{(n-\gamma+1)(n-\gamma+2)}} \right)^T, \quad \gamma = \overline{2, n}, \quad \alpha = \overline{1, n}.$$

Нетрудно убедиться, что

$$\sum_{\alpha=1}^n r_{\alpha 1} = \sqrt{n}, \quad \sum_{\alpha=1}^n r_{\alpha \gamma} = \sqrt{\frac{n-\gamma+1}{n-\gamma+2}} -$$

$$-(n-\gamma+1) \frac{1}{\sqrt{(n-\gamma+1)(n-\gamma+2)}} = 0, \quad \gamma = \overline{2, n}. \quad (7)$$

Также легко проверить, что

$$\sum_{\alpha=1}^n r_{\alpha i}^2 = 1, \quad i = \overline{1, n}, \quad \sum_{\alpha=1}^n r_{\alpha 1} r_{\alpha \gamma} = 0, \quad \gamma = \overline{2, n}, \quad (8)$$

$$\sum_{\alpha=1}^n r_{\alpha \gamma} r_{\alpha \delta} = 0, \quad \gamma > \delta \geq 2,$$

т. е. матрица R ортогональная.

Поскольку $z_{\alpha} \sim N(0, 1)$ и независимы, а R — ортогональная матрица, величины y_{β} также независимы и нормальны $N(0, 1)$ [68]. При этом квадратичная форма в (6), которая может быть записана как $\vec{z}^T Q \vec{z}$, при переходе к $\vec{y} = R^T \vec{z}$ преобразуется в $\vec{y}^T R^T Q R \vec{y}$ [68; 87]. Матрица преобразованной формы равна

$$B = R^T Q R = R^T \left(I - \frac{1}{n} E \right) R = R^T R - \frac{1}{n} R^T E R = I -$$

$$- \frac{1}{n} R^T E R = \text{diag} \{ 0, 1, \dots, 1 \}$$

в силу (7) и ортогональности (8). Кроме того,

$$z_1 - \bar{z} = \sum_{\alpha=1}^n q_{1\alpha} z_{\alpha} = \sum_{\alpha=1}^n q_{1\alpha} \sum_{\beta=1}^n r_{\alpha\beta} y_{\beta} =$$

$$= \sum_{\beta=1}^n y_{\beta} \sum_{\alpha=1}^n q_{1\alpha} r_{\alpha\beta} = y_1 \left[\left(1 - \frac{1}{n} \right) \frac{1}{\sqrt{n}} -$$

$$- \frac{1}{\sqrt{n}} \frac{n-1}{n} \right] + y_2 \left[\left(1 - \frac{1}{n} \right) \sqrt{\frac{n-1}{n}} +$$

$$+ \frac{1}{n} \sqrt{\frac{n-1}{n}} \Big] - \frac{1}{n} \sum_{\beta=3}^n y_{\beta} \sum_{\alpha=1}^n r_{\alpha\beta} = y_2 \sqrt{\frac{n-1}{n}}.$$

Таким образом, величина ξ_1 может быть записана через с. в. y_{β} , $\beta = \overline{2, n}$, в виде

$$\xi_1 = |y_2| \sqrt{\frac{1}{n-1} \sum_{\beta=2}^n y_{\beta}^2}, \quad y_{\beta} \sim N(0, 1), \quad \beta = \overline{2, n}.$$

Рассмотрим теперь $\tau = y_2 \sqrt{\frac{1}{n-1} \sum_{\beta=2}^n y_{\beta}^2}$, где y_{β} — независимые и нормальные $N(0, 1)$ с. в. Введя взаимно-однозначное преобразование [49]

$$\begin{aligned} \tau' &= \sqrt{\frac{n-2}{n-1}} \frac{\tau}{\sqrt{1 - \frac{\tau^2}{n-1}}} = \\ &= y_2 \sqrt{\frac{1}{n-2} \sum_{\nu=3}^n y_{\nu}^2}, \end{aligned}$$

убедимся, что τ' имеет распределение Стьюдента с $n-2$ степенями свободы: $\tau' \sim St_{n-2}(x)$. В силу монотонности преобразования $\tau' \rightarrow \tau$ условие

$$\tau' < \sqrt{\frac{n-2}{n-1}} \frac{x}{\sqrt{1 - \frac{x^2}{n-1}}}$$

эквивалентно условию $\tau < x$, поэтому

$$P\{\tau < x\} = P\left\{\tau' < \sqrt{\frac{n-2}{n-1}} \frac{x}{\sqrt{1 - \frac{x^2}{n-1}}}\right\}.$$

Поскольку распределение Стьюдента симметрично и $\xi_1 = |\tau|$, получим

$$P\{\xi_1 \geq x\} = 2P\{\tau \geq x\} = \\ = 2 \left(1 - St_{n-2} \left(\frac{x \sqrt{n-2}}{\sqrt{n-1-x^2}} \right) \right).$$

Поэтому для определения ε -процентной критической точки статистики D_2 с учетом (5) получим

$$P\{D_2 \geq x\} \approx nP\{\xi_1 \geq x\} = \\ = 2n \left(1 - St_{n-2} \left(\frac{x \sqrt{n-2}}{\sqrt{n-1-x^2}} \right) \right).$$

т. е. критическая точка x_ε определяется из уравнения

$$\frac{\varepsilon}{2n} = 1 - St_{n-2} \left(\frac{x_\varepsilon \sqrt{n-2}}{\sqrt{n-1-x_\varepsilon^2}} \right). \quad (9)$$

Для получения функции мощности критерия, основанного на D_2 , требуются более громоздкие обоснования и выкладки [13].

Кратко остановимся теперь на выявлении аномальных наблюдений в модели Б. Наряду с ситуацией, описанной выше, когда имеется единственная выборка, возможно содержащая аномальные наблюдения, для которой применимы вышеуказанные критерии, можно рассмотреть также задачу выявления аномальной группы наблюдений среди k групп [11; 66]. Пусть имеются k групп независимых наблюдений по m наблюдений в каждой и пусть s_1^2, \dots, s_k^2 — соответствующие выборочные дисперсии. Если все k групп однородны (т. е. дисперсии $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$), то s_1^2, \dots, s_k^2 — взаимно независимые и одинаково распределенные с. в. Есть подозрение, что одна из групп имеет отличающуюся дисперсию. В случае справедливости нулевой гипотезы H_0 об однородности групп все s_i^2 ($i = \overline{1, k}$) распределены как $\sigma^2 \chi^2_{m-1} / (m-1)$. Составим отношения $G_i = s_i^2 / \sum_{j=1}^k s_j^2$ ($i = \overline{1, k}$)

и положим $G = \max_i G_i$. Гипотеза H_0 отвергается, если $G > \omega_\varepsilon$, где ω_ε — решение уравнения, получаемого аналогично (5): $kP\{G_1 > \omega_\varepsilon\} = \varepsilon$. Можно пока-

зять, что G_1 подчиняется бета-распределению: $G_1 \sim \sim B(p, q)$ с параметрами $p = (m - 1)/2$, $q = (m - 1) \times \times (k - 1)/2$ [11; 12].

Глава 4

ВЫЯВЛЕНИЕ АНОМАЛЬНЫХ НАБЛЮДЕНИЙ В СОВОКУПНОСТИ СВЯЗАННЫХ ВЕЛИЧИН

До сих пор речь шла о выявлении резко выделяющихся наблюдений в выборке для одномерной с. в. В экспериментах нередко изучается зависимость некоторой величины от контролируемого параметра t (например, число вторичных частиц как функция от энергии первичных частиц). Если при этом условия эксперимента и режим измерений для соседних точек t_i, t_j ($i \neq j$) каждый раз регулируют заново, то результаты измерений в соседних точках статистически независимы и измерения кривой представляют собой набор n (n — размер сетки аргумента) одномерных выборок. К каждой из них можно применить критерии выявления аномальных наблюдений, обсуждавшиеся выше. Однако, поскольку экспериментальные кривые несут некую информацию об изучаемом явлении, воспроизводят какую-то структуру, можно привлечь эту дополнительную информацию для получения более обоснованных заключений.

Например, есть теоретическая модель, описывающая форму кривой и содержащая неизвестные, подлежащие подгонке с помощью эксперимента, параметры. Возможен случай, что такой модели нет, но есть априорные соображения о свойствах кривой (ее гладкости, интервалах роста и убывания, периодичности и т. д.). Тогда обычно экспериментальную кривую приближают по методу наименьших квадратов функциями с соответствующими свойствами. Приближение это может быть глобальным (на всей сетке $\langle t_i \rangle$, $i = 1, n$) либо локальным (сетка $\langle t_i \rangle$ разбивается на участки с небольшим числом точек и для каждого такого участка выбирается своя аппроксимирующая функция). Далее применяют критерии выявления сбоев, используя вместо среднего значения в D_2 построенную приближающую функцию [13].

При получении кривых в автоматическом режиме регистрации и накопления информации аппаратурный дрейф приводит к согласованным изменениям в соседних точках, к появлению корреляционных связей (подробнее см. гл. 12). При этом техника выявления одномерных резко выделяющихся наблюдений не годится. Однако при некоторых предположениях о характере измеряемой информации и способе измерений удается свести проблему к известным стандартным критериям в рамках модели А [19].

Пусть в результате регистрации $Y(t)$ получается набор $\Omega^{(m)} = \{Y_1(t_i), \dots, Y_m(t_i)\}$, $i = \overline{1, n}$, составленный из n -мерных с. в., принадлежащих к различным генеральным совокупностям. Предполагается, что:

1) $\vec{Y}_\alpha = \{Y_\alpha(t_i)\}$ — n -мерная нормальная с. в. с независимыми компонентами;

2) математическое ожидание отношения $\vec{Y}_\alpha / \vec{Y}_\beta$ является гладкой функцией t и может быть аппроксимировано в l -окрестности произвольной точки t_{i_0} полиномом степени k ;

3) нелинейные эффекты, вследствие которых $E(Y_\alpha(t)/Y_\beta(t)) \neq \text{const}$, считаются малыми:

$$\frac{\max_i E\left(\frac{Y_{\alpha i}}{Y_{\beta i}}\right) - \min_i E\left(\frac{Y_{\alpha i}}{Y_{\beta i}}\right)}{E\left(\frac{Y_{\alpha i_0}}{Y_{\beta i_0}}\right)} \ll 1,$$

$$i = \overline{i_0 - l, i_0 + l}, \alpha, \beta = \overline{1, m};$$

$$4) \sigma Y_{\alpha i} / E Y_{\alpha i} \ll 1;$$

5) считается, что отношение $\frac{\sigma^2 Y_{\alpha i}}{E Y_{\alpha i}} = \eta = \text{const}$, $\alpha = \overline{1, m}$, $i = \overline{1, n}$ (условие равноточности процесса измерений);

6) вероятность появления резко выделяющегося наблюдения мала. Конкретно предполагается, что вероятности сбоя более одного значения $Y_{\alpha i_0}$ в точке i_0 ($\alpha = \overline{1, m}$) и более одного значения $Y_{\alpha i}$ на небольшом участке (в l -окрестности точки i_0) малы.

В качестве первого шага найдем нормирующие множители $N_\alpha(t)$ такие, что $E[N_\alpha(t)Y_\alpha(t)] = EY_0(t)$ (здесь \bar{Y}_0 — эталонная кривая либо $\bar{Y}(t)$). Составим отношение $K_\alpha(t) = \frac{Y_\alpha(t)}{Y_0(t)}$, $\alpha = \overline{1, m}$, и, используя условия $\frac{\sigma Y_\alpha}{EY_\alpha} \ll 1$ и $\frac{\sigma Y_0}{EY_0} \ll 1$, проведем разложение K_α по разностям $(EY_\alpha)^{-1}(Y_\alpha - EY_\alpha)$ и $(EY_0)^{-1}(Y_0 - EY_0)$. Тогда можно показать, что $K_\alpha(t_i) \sim N\left(\frac{EY_\alpha}{EY_0}, \sigma K_{\alpha i}\right)$, причем для вычисления дисперсии $\sigma^2 K_{\alpha i}$ получены формулы [19].

Используя условие 2, представляем $EK_{\alpha i}$ в l -окрестности точки i_0 в виде полинома

$$EK_{\alpha i} = \mathcal{P}_{\alpha i} \equiv u_1^{(\alpha)}(i - i_0)^x + u_2^{(\alpha)}(i - i_0)^{x-1} + \dots \\ \dots + u_{x+1}^{(\alpha)}.$$

Пусть $\tilde{\mathcal{P}}_{\alpha i}$ — оценка $\mathcal{P}_{\alpha i}$ по методу наименьших квадратов. Определим нормирующие множители $N_{\alpha i}$ в точке i_0 равенством

$$N_{\alpha i_0} = (\tilde{\mathcal{P}}_{\alpha i_0})^{-1} = (\tilde{u}_{x+1}^{(\alpha)})^{-1}.$$

Тогда $Ey_{\alpha i_0} = E[Y_{\alpha i_0}/\tilde{u}_{x+1}^{(\alpha)}] = EY_{0i_0}$, с точностью до членов порядка $\left(\frac{\sigma Y}{EY}\right)^2$. Действительно, с точностью до $O\left(\frac{\sigma Y}{EY}\right)^2$ имеем

$$y_{\alpha i_0} = \frac{EY_{\alpha i_0}}{E\tilde{u}_{x+1}^{(\alpha)}} \left(1 + \frac{Y_{\alpha i_0} - EY_{\alpha i_0}}{EY_{\alpha i_0}} - \frac{\tilde{u}_{x+1}^{(\alpha)} - E\tilde{u}_{x+1}^{(\alpha)}}{E\tilde{u}_{x+1}^{(\alpha)}} \right).$$

Используя соотношение

$$EK_{\alpha i_0} = \mathcal{P}_{\alpha i_0} = u_{x+1}^{(\alpha)} = E\tilde{u}_{x+1}^{(\alpha)} = \frac{EY_{\alpha i_0}}{EY_{0i_0}},$$

найдем

$$Ey_{\alpha i_0} = \frac{EY_{\alpha i_0}}{E\tilde{u}_{x+1}^{(\alpha)}} = \frac{EY_{\alpha i_0}}{EY_{0i_0}} \cdot \frac{EY_{0i_0}}{E\tilde{u}_{x+1}^{(\alpha)}} = EY_{0i_0}.$$

Заметим, однако, что величины $y_{1i_0}, \dots, y_{mi_0}$ не являются выборкой, ибо они статистически зависимы вследствие проведенной нормировки. Вычисления показывают [19], что ковариационная матрица $\|\text{cov}(y_{\alpha i_0}, y_{\beta i_0})\|$ недиагональна.

Проверяется гипотеза $H_0: Ey_{1i_0} = Ey_{2i_0} = \dots = Ey_{mi_0} = a$ против альтернативы H_1 : существует такое a^* , что $Ey_{\alpha^*} \neq a$. Если количество элементов в $\omega^{(m)} = \{y_{1i_0}, \dots, y_{mi_0}\}$ не очень велико, а смещение м.о. y_{α^*} значительно, то с вероятностью, близкой к 1, y_{α^*} совпадает с y_{α} , наиболее удаленной от среднего \bar{y} , т. е. будем использовать критерий, основанный на статистике

$$\xi = \max_{\alpha} \frac{|y_{\alpha} - \bar{y}|}{s_{\alpha}}$$

Назовем y_{α^*} , соответствующее ξ , *сбитым*, если $\xi \geq z_{\epsilon}$, где z_{ϵ} — $(1 - \epsilon)$ -квантиль функции распределения с.в. ξ , т. е. решение уравнения $1 - F_{\xi}(z) = \epsilon$ (критерий А).

Можно показать, что выбор z производится из такого же уравнения, как рассмотренное ранее для D_2 ; в [19] построена оценка дисперсии для кривой $\bar{y}^*(t)$ (т. е. для среднего по оставшимся после отбраковки наблюдениям). В случае, если число точек для построения $\mathcal{P}_{\alpha i}$ достаточно большое и сбой в первичной информации не очень велики, так что резко выделяющемуся $y_{\alpha^* i^*}$ отвечает резко выделяющееся наблюдение $Y_{\alpha^* i^*}$, удастся оценить также дисперсию для кривой $\mathcal{P}^*(t)$.

При сильных сбоях статистический анализ начинают с анализа хода каждой отдельной кривой $Y_{\alpha}(t)$ и устранения резко выпадающих точек. Целесообразность такого подхода состоит в том, что при малом m критерий А недостаточно эффективен.

Если бы вид математического ожидания $EY_{\alpha}(t)$ был известен, можно было построить критерий, аналогичный критерию А, с заменой выборочного среднего на соответствующую эмпирическую кривую регрессии. Однако часто вид $EY_{\alpha}(t)$ неизвестен, структура этой кривой весьма сложна. Поэтому непосредственно к $Y_{\alpha}(t)$ регрессионный анализ неприменим. Поделим кривые, чтобы уничтожить структуру:

$$K_{\alpha}(t) = \frac{Y_{\alpha}(t)}{Y_0(t)}, \quad \alpha = \overline{1, m}.$$

Смысл $Y_0(t)$ тот же, что и для критерия А. Как и выше, пусть $EK_{\alpha}(t)$ может быть приближена полиномом в l -окрестности точки i_0 . Рассмотрим совокупность величин $K_i = K_{\alpha i}$, $i = i_0 - l, i_0 + l$. Проверке подлежит гипотеза $H_0: EK_i = \mathcal{P}_i$ против альтернативы H_1 : существует такое i^* , что $EK_{i^*} \neq \mathcal{P}_{i^*}$. Рассмотрим

$$\xi = \max_i \xi_i = \max_i \frac{|K_i - \tilde{\mathcal{P}}_i|}{\theta_i s_K}$$

где $\tilde{\mathcal{P}}_i$ — оценка по методу наименьших квадратов полинома \mathcal{P}_i ,

$$s_K^2 = \frac{1}{2l - \kappa} \sum_{i=i_0-l}^{i_0+l} g_i (K_i - \tilde{\mathcal{P}}_i)^2,$$

g_i — весовые множители; θ_i — константы, выбираемые так, чтобы все ξ_i были одинаково распределены [19].

Значение K_{i^*} считается *сбитым*, если $\xi \geq z_{\varepsilon}$, где z_{ε} — ε -процентная точка функции распределения с. в. ξ (критерий Б). Аналогично критерию А удается показать, что z_{ε} является решением уравнения

$$1 - St_N \left(z \sqrt{\frac{N}{N+1-z^2}} \right) = \frac{\varepsilon}{2(2l+1)},$$

где $N = 2l - \kappa - 1$. Анализ мощности критериев А и Б показывает, что во многих практических ситуациях эффективность их применения мало отличается от использования студентизированного абсолютного отклонения D_2 .

В модельном эксперименте по критерию Б были обработаны фотоядерные кривые выхода [19]. Рассматривались 8 кривых, значения которых известны в 62 точках. На экспериментальных кривых было смоделировано 8 сбитых точек: две из них были сбиты на 5 статошибок, остальные — на 3. После обработки по критерию Б исправлению подверглись 4 сбитых точки из 8 и, кроме того, 3 точки из исходной информации. При этом были использованы значения пара-

Видно, что проекции сбитой точки на оси $x^{(1)}$ и $x^{(2)}$ по величине не выделяются из основной массы наблюдений. В то же время сочетание пары координат $(x^{(1)}, x^{(2)})$ является аномальным. Приходится учитывать не только отклонения, но и ковариации компонент наблюдений, что сделать сложнее и теоретически, и с точки зрения реализации критериев.

Так, Уилксом [138] был предложен критерий, основанный на вычислении отношения определителей выборочных ковариационных матриц

$$\omega_v = \frac{\det Q_v}{\det Q},$$

где

$$Q = \frac{1}{N} \sum_{\alpha=1}^N (\vec{x}_\alpha - \vec{x})(\vec{x}_\alpha - \vec{x})^T,$$

$$Q_v = \frac{1}{N-1} \sum_{\substack{\alpha=1, \\ \alpha \neq v}}^N (\vec{x}_\alpha - \vec{x})(\vec{x}_\alpha - \vec{x})^T.$$

Критерий имеет вид

$$\omega = \min_v \omega_v \leq \omega_\epsilon,$$

где ω_ϵ — критическое значение для ϵ -процентного размера критерия. Имеется оценка для приближенного определения (типа (5) из гл. 3) критических точек статистики ω , однако функция мощности этого критерия не исследована. Кроме того, его применение требует большого объема вычислений.

Глава 5

РОБАСТНЫЕ МЕТОДЫ И ОЦЕНКИ

Обсуждая приведенный выше пример выявления аномальных наблюдений в модельной задаче, мы видели, что не все аномальные наблюдения были исправлены и в то же время исправлению подверглись наблюдения из исходной нормальной совокупности.

Таким образом, исправленные данные не образуют популяции с нормальным законом распределения, даже если исходные «неиспорченные» данные были нормально распределены. Такая ситуация характерна не только при обработке наблюдений с выявлением среди них аномальных, но, как мы увидим ниже, типична в задачах анализа и обработки экспериментальной информации.

Отметим прежде всего ограничения, сделанные при рассмотрении критериев аномальности наблюдений.

1. Предполагался нормальный закон распределения совокупности, если верна нулевая гипотеза, причем параметры его одинаковы для всех наблюдений. Если же нулевая гипотеза ошибочна, основная часть наблюдений распределена нормально с одинаковыми параметрами, а аномальные наблюдения имеют другое м. о. (модель А) или дисперсию (модель Б), однако также нормальны.

2. Аномальных наблюдений немного, в идеале — одно на выборку. В противном случае сказывается «маскирующий эффект» и эффективность критериев обнаружения аномалий резко падает.

Однако те соображения, которые обычно приводятся при обосновании нормальности, по меньшей мере спорны. Так, гипотеза о суммировании большого числа «элементарных» погрешностей практически не проверяема. Утверждение центральной предельной теоремы носит асимптотический характер, и никогда нет уверенности, что число «элементарных» ошибок достаточно велико, чтобы гарантировать нормальность. Кроме того, и это главное, однородность «элементарных» ошибок также сомнительна. В частности, возникновение аномальных наблюдений (грубых промахов, нарушающих однородность, распределение которых может резко отличаться от нормального) отлично от ошибок основной массы наблюдений. Второй фактор, приводящий к нормальности, — асимптотическое поведение распределений — также обычно не действует в чистом виде. Так, хотя число частиц, образующихся в ядерных реакциях, обычно считается пуассоновским, процесс регистрации, наличие «мертвого времени» у детекторов частиц искажает пуассоновский закон [27]. Кроме того, отсчеты, связанные с

фоновой подложкой, также нарушают предполагаемую асимптотическую нормальность регистрируемых количеств частиц.

Таким образом, априорные соображения о нормальности регистрируемых величин представляются сомнительными. Проверка закона распределения наблюдаемых величин с помощью опыта показывает, что обычно нормальность не имеет места [99; 120]. Так, Шеффе [99] приводит значения выборочной асимметрии b_1 и выборочного эксцесса b_2 для величин, регистрируемых в различных областях науки. Для технических измерений значение b_2 оказалось в пределах от 0,7 до 0,9, а b_1 — от $-0,4$ до 1,8. При химических анализах для 100 опытов получено $b_2=7$, что указывает на наличие грубых случайных ошибок (например, нечистоту посуды или реактивов для проведения опытов). Другие примеры, полученные при обработке очень большого числа наблюдений (10 000 и более), также подтверждают эту тенденцию. Величины (b_1 ; b_2) принимают следующие значения: (2; 6,3) для распределения возраста невест в Австралии в 1907—1914 гг.; (2; 5,3) для распределения возраста женихов; (0,5; 0,4) для барометрического давления в Гринвиче за 1848—1922 гг.; ($-0,9$; 1,8) для длины бобов; ($-0,4$; 0,6) для их ширины. Таким образом, реальные распределения наблюдаемых величин асимметричны и имеют более тяжелые «хвосты», чем нормальное распределение. Для астрономических наблюдений этот факт был отмечен еще Бесселем [103]. Он объяснил более частое появление больших ошибок, чем это предсказывается нормальным законом, наличием наряду с «обычными» инструментальными факторами источников редких грубых промахов (усталость наблюдателя, неисправность инструмента и т. д.). Таким образом, наблюдаемое распределение ошибок обусловлено действием большого числа малых, примерно одинаково распределенных погрешностей и небольшого числа больших ошибок. Тогда результирующее распределение является смесью распределений: близкого к нормальному для основной массы наблюдений и «загрязняющего», отличного от нормального, для аномальных наблюдений. Фактически модель Б аномальных наблюдений является частным случаем описанной смеси распределений, когда

«загрязняющее» распределение также нормально, но с большей дисперсией.

Ясно, что устранение резко выделяющихся наблюдений, изменяя форму «хвостов» распределения, оставляет нетронутой среднюю часть. Поэтому ненормальность сохраняется даже после выявления и устранения сбоев и приводит к тому, что оптимальные для нормального закона методы, критерии и оценки не являются таковыми при обработке обычных в экспериментальной практике выборок. Рассмотрим это на примере точечных оценок параметров ф. р.

Напомним некоторые факты из теории статистических оценок [45; 87]. Пусть x_1, \dots, x_n — выборка наблюдений из популяции с ф. р. $F(x, \theta)$, где θ — параметр, истинное значение которого θ_0 неизвестно. Под *точечной оценкой* $\hat{\theta}$ параметра θ понимается с. в. $\hat{\theta}(x_1, \dots, x_n)$, которую используют в дальнейшем вместо истинного значения. Желательно, чтобы $\hat{\theta}$ удовлетворяла некоторым (или всем) разумным требованиям:

- 1) несмещенности: $E\hat{\theta} = \theta_0$;
- 2) состоятельности: $P\text{-}\lim_{n \rightarrow \infty} \hat{\theta} = \theta_0$.

Эти требования могут и не выполняться для используемых оценок. Так, оценка x_1 для м. о. несмещенная, но несостоятельная. Оценка дисперсии σ^2

$$(s^*)^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

состоятельная, но $E(s^*)^2 = (n-1)\sigma^2/n$. Поэтому $(s^*)^2$ — смещенная оценка. Однако при $n \rightarrow \infty$ ее смещение стремится к 0, следовательно, она несмещенная асимптотически.

Даже ограничиваясь несмещенными и состоятельными оценками, обычно можно указать бесконечное множество таких оценок. Так, м. о. μ нормального распределения можно оценить через выборочное среднее \bar{x} и выборочную медиану m , причем обе оценки состоятельные и несмещенные. Этими же свойствами обладает оценка $x_\beta = (1-\beta)\bar{x} + \beta m$, $0 \leq \beta \leq 1$. Встает вопрос о выборе оптимальной (в каком-то смысле) оценки.

При прочих равных условиях лучше та оценка, которая «в среднем» ближе к истинной величине оцениваемого параметра, т. е. ф. р. которой «плотнее» локализована около θ_0 . Однако задача вычисления ф. р. решена лишь для некоторых распределений и оценок. В то же время показано, что ряд оценок имеет при $n \rightarrow \infty$ нормальное распределение, для которого плотность локализации обратна корню из дисперсии. Поэтому обычно за меру качества оценки принимают ее дисперсию [45]: чем меньше дисперсия оценки, тем она лучше. Можно показать, что дисперсия несмещенной оценки не может быть произвольно малой [45; 49; 87; 132].

Теорема (Крамера, Рао). Пусть (x_1, \dots, x_n) — n -мерная с. в. из R_n с ф. р. $F_n(x_1, \dots, x_n; \theta)$, причем F_n регулярна относительно первой производной по θ , т. е.

$$\int_{R_n} \frac{\partial}{\partial \theta} dF_n(\vec{x}, \theta) = \frac{\partial}{\partial \theta} \int_{R_n} dF_n(\vec{x}, \theta) = 0$$

в некотором открытом интервале Ω , содержащем θ_0 . Тогда если $\hat{\theta}(x_1, \dots, x_n)$ — несмещенная оценка для θ , то

$$D\hat{\theta} \geq \left[\int_{R_n} \left(\frac{\partial \ln dF_n}{\partial \theta} \right)_{\theta=\theta_0}^2 dF_n \right]^{-1}. \quad (1)$$

Доказательство. В силу несмещенности $\hat{\theta}$

$$\int_{R_n} (\hat{\theta} - \theta) dF_n \equiv 0.$$

Дифференцируя по θ , получим в силу регулярности F_n :

$$\int_{R_n} (\hat{\theta} - \theta) \frac{\partial \ln dF_n}{\partial \theta} dF_n = 1.$$

Используя неравенство Коши-Буняковского, получим

$$1 = \left[\int_{R_n} (\hat{\theta} - \theta) \frac{\partial \ln dF_n}{\partial \theta} dF_n \right]^2 \leq$$

$$\leq \int_{R_n} (\hat{\theta} - \theta)^2 dF_n \int_{R_n} \left(\frac{\partial \ln dF_n}{\partial \theta} \right)^2 dF_n$$

и окончательно

$$D\hat{\theta} = \int_{R_n} (\hat{\theta} - \theta)^2 dF_n \geq \left[\int_{R_n} \left(\frac{\partial \ln dF_n}{\partial \theta} \right)^2 dF_n \right]^{-1},$$

причем равенство имеет место тогда и только тогда, когда с вероятностью 1

$$K(\theta) [\hat{\theta}(x_1, \dots, x_n) - \theta] = \frac{\partial \ln dF_n(x_1, \dots, x_n; \theta)}{\partial \theta}.$$

Положив $\theta = \theta_0$, получим утверждение теоремы.

Следствие. В случае выборки наблюдений

$$dF_n = \prod_{i=1}^n dF(x_i; \theta),$$

$$\frac{\partial \ln dF_n}{\partial \theta} = \sum_{i=1}^n \frac{\partial \ln dF(x_i; \theta)}{\partial \theta},$$

поэтому

$$D\hat{\theta} \geq \left[n \int_{-\infty}^{\infty} \left(\frac{\partial \ln dF}{\partial \theta} \right)_{\theta=\theta_0}^2 dF \right]^{-1},$$

причем необходимое и достаточное условие равенства записывается в виде

$$K(\theta) (\hat{\theta} - \theta_0) = \sum_{i=1}^n \frac{\partial \ln dF(x_i; \theta)}{\partial \theta}$$

с вероятностью 1.

Определение 1. Если $\hat{\theta}$ — несмещенная оценка θ_0 с конечной дисперсией и при этом не существует другой несмещенной оценки с меньшей дисперсией, то $\hat{\theta}$ называется эффективной оценкой для θ_0 .

Рассмотрим пример оценивания м. о. и нормального распределения, если дисперсия σ^2 известна. Имеем

$$\frac{\partial \ln f(x; \mu)}{\partial \mu} = \frac{\partial}{\partial \mu} \left\{ \text{const} - \frac{(x - \mu)^2}{2\sigma^2} \right\} = -\frac{x - \mu}{\sigma^2},$$

так что

$$D\hat{\mu} \geq \sigma^4 \left[n \int_{-\infty}^{\infty} (x - \mu)^2 dF(x; \mu) \right]^{-1} = \frac{\sigma^2}{n}.$$

Поскольку для выборочного среднего имеет место точное равенство $D\bar{x} = \frac{\sigma^2}{n}$, то $\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ является эффективной оценкой.

Для произвольного распределения может не быть эффективной оценки параметра, однако для некоторых оценок неравенство Крамера — Рао (или более точное неравенство Бхаттачария [45]) переходит в равенство при $n \rightarrow \infty$. Такие оценки называются асимптотически эффективными. Обычно они могут быть получены для распределений, описывающих результаты наблюдений, по методу максимального правдоподобия [45; 87]: МП-оценки являются решением уравнения

$$\sum_{i=1}^n \frac{\partial \ln dF(x_i; \theta)}{\partial \theta} = 0.$$

Теорема Крамера — Рао и метод максимального правдоподобия обобщаются на случай векторного k -мерного параметра θ .

Определение 2. Эффективностью оценки $\hat{\theta}$ относительно оценки $\hat{\theta}'$ называется отношение их дисперсий:

$$\text{eff}(\hat{\theta} | \hat{\theta}') = D\hat{\theta}' / D\hat{\theta}. \quad (2)$$

Если в качестве $\hat{\theta}'$ используется (асимптотически) эффективная оценка МП, величина

$$\text{eff}(\hat{\theta}) = [D\hat{\theta}]^{-1} \left[\int_{R_n} \left(\frac{\partial \ln dF_n}{\partial \theta} \right)_{\theta=\theta_0}^2 dF_n \right]^{-1} \quad (2')$$

называется эффективностью оценки $\hat{\theta}$. Очевидно, что $0 < \text{eff } \hat{\theta} < 1$.

Вернемся к рассмотрению задач обработки наблюдений. Поскольку для измеряемых величин нормальность обычно не имеет места, в принципе возможно предложить закон, описывающий их распределение, и оценить параметры по методу МП для данной ф. р. F . Однако нужно иметь в виду, что мы никогда не знаем точно закона распределения наблюдений. Критерии согласия скорее отвечают на вопрос о том, какое распределение не описывает наблюдаемые величины, чем позволяют однозначно выбрать истинное распределение. Особенно трудно сделать выбор в случае смеси распределений.

Но даже если бы удалось точно установить закон распределения наблюдаемых величин и построить МП-оценки параметров, все равно остается проблема аномальных наблюдений, которые вызываются сильнодействующими, но редкими факторами, в результате чего их трудно обнаружить. Они нерегулярны, но способны сильно ухудшить эффективность МП-оценок.

Вместо того чтобы, отказываясь от нормальности, искать оценки, «настроенные» на данное конкретное распределение, можно поступить более радикально. Не будем искать эффективных для данного закона оценок, а рассмотрим оценки, которые имеют высокую эффективность для широкого класса распределений. Такая позиция связана с концепцией робастности [38; 61; 69; 120; 137].

Вообще говоря [33], «слабая чувствительность к отклонениям от стандартных условий называется робастностью. Методы, применимые в широком диапазоне реальных условий, являются робастными. Можно говорить о робастности по отношению к неоднородности дисперсий, к отсутствию независимости, к нарушению нормальности и т. д.». Чаще всего, говоря о робастности, имеют в виду нечувствительность к нарушениям нормальности. Различают робастность к предпосылкам — терпимость к ненормальности распределений на «хвостах» (ее пример — студентово t , используемое при построении доверительных интервалов, проверке гипотез; пример неробастной к предпосылкам статистики — χ^2 [45; 61]) и робаст-

ность к эффективности — высокую эффективность, несмотря на ненормальность «хвостов».

Рассмотрим второй вид робастности на примере оценивания параметра сдвига. Для него предложено наибольшее количество оценок [38; 120], их свойства исследованы в различных возможных ситуациях теоретически и методом Монте-Карло.

Предположим, что производятся многократные измерения интенсивности θ некоторого процесса, причем отсутствуют систематические ошибки. Тогда, если распределение случайных ошибок унимодально [44], θ является центром некоторого распределения, описывающего ошибки, т. е. ф. п. в. $f(x-\theta)$ симметрична относительно вертикальной прямой $x=\theta$. Величина θ называется *параметром сдвига*. В частности, если существует м. о., то

$$E\xi = \alpha_1 = \int_{-\infty}^{\infty} x dF(x) = \int_{-\infty}^{\infty} (x-\theta) dF(x-\theta) + \theta \int_{-\infty}^{\infty} dF(x-\theta) = \theta.$$

Однако, например, для распределения Коши с функцией плотности вероятностей

$$f(x) = \frac{1}{\pi} \frac{1}{1+(x-\theta)^2},$$

θ — параметр сдвига, но конечного первого момента не существует.

Если наблюдаемые величины нормальны, то, как показано выше, выборочное среднее \bar{x} — эффективная оценка м. о. μ . Если же нормальность не имеет места, эффективность выборочного среднего сильно падает. Так, для смеси нормальных распределений с ф. п. в.

$$f(x, \theta) = (1-\varepsilon) \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}} + \frac{\varepsilon}{\sqrt{2\pi} \sigma} e^{-\frac{(x-\theta)^2}{2\sigma^2}} \quad (3)$$

дисперсия \bar{x} возрастает квадратично с ростом σ :

$$D\bar{x} = \frac{1}{n} [(1 - \varepsilon) + \varepsilon\sigma^2],$$

в то время как нижняя граница примерно постоянна и равна n^{-1} . В этих условиях использование даже такой неэффективной для нормального закона оценки, как выборочная медиана m (ее эффективность равна $2/\pi = 0,637$ [45]), выгоднее, чем применение \bar{x} , если только σ достаточно велико, потому что m робастна к отклонениям от нормальности. В последние 30 лет усилия статистиков были направлены на поиск и исследование робастных оценок, почти столь же устойчивых, как выборочная медиана, но высокоэффективных (90% и выше) для выборок из нормального распределения.

Многие робастные оценки вычисляются по части членов вариационного ряда $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$, в частности $m = x_{((n+1)/2)}$ для нечетных n и $m = (x_{(n/2)} + x_{(n/2+1)})/2$ для четного n . Часто используются α -«обрубленное» среднее

$$x_{T(\alpha)} = \frac{1}{n - 2[n\alpha]} \sum_{k=[n\alpha]+1}^{n-[n\alpha]} x_{(k)}, \quad (4)$$

α -«винзоризованное» среднее [137]

$$x_{w(\alpha)} = \frac{1}{n} \left\{ \sum_{k=[n\alpha]+1}^{n-[n\alpha]} x_{(k)} + [n\alpha](x_{([n\alpha])} + x_{(n-[n\alpha]+1)}) \right\}, \quad (5)$$

оценка Ходжеса — Лемана [113]

$$x_{H-L} = \text{median} (x_i + x_j)/2, \quad (6)$$

$1 \leq i \leq j \leq n$

M-оценка Хубера получается при минимаксном подходе [119]: для выборки, являющейся смесью распределений с ф. п. в.

$$f(x) = (1 - \varepsilon) \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}} + \varepsilon\varphi(x-\theta), \quad (7)$$

где $\varphi(x-\theta)$ унимодальна, симметрична относительно

0 и имеет более «тяжелые» хвосты, чем нормальная $N(0, 1)$, строится оценка МП-параметра θ . Рассматривая ее как функционал от φ и найдя «наихудшую» ф. п. в. φ , при которой дисперсия МП-оценки максимальна, Хубер построил минимаксную оценку T параметра θ , которая получается для распределений, имеющих конечный первый момент α_1 при $\varphi(x - \theta) = \frac{1}{2} \exp\{-|x - \theta|\}$. Тогда М-оценка Хубера находится минимизацией функционала

$$\Phi = \sum_{i=1}^n \rho(x_i - T),$$

где

$$\rho(t) = \begin{cases} t^2/2, & |t| < k(\varepsilon), \\ k|t| - k^2/2, & |t| \geq k(\varepsilon), \end{cases}$$

а $k = k(\varepsilon)$ — решение уравнения $(2\pi)^{-1/2} \int_{-k}^k e^{-t^2/2} dt = \varepsilon$.

Доказана состоятельность, асимптотическая несмещенность этих и ряда других оценок, их асимптотическая нормальность, вычислены эффективности робастных оценок относительно выборочного среднего (если $\hat{\theta}$ — робастная оценка, ее эффективность относительно \bar{x} $\text{eff}(\hat{\theta}|\bar{x}) = D\bar{x}/D\hat{\theta}$). Оценки $x_{T(\alpha)}$ и $x_{w(\alpha)}$ имеют при $\alpha = 0,1 \div 0,25$ эффективности $0,87 \div 0,95$ для $\varepsilon = 0$ в (3), а при $\sigma = 3$ и $\varepsilon = 0,05$; $\varepsilon = 0,1$ и $\varepsilon = 0,3$ относительные эффективности равны $1,20$; $1,43$ и $1,64$ для x_{H-L} , $1,12$; $1,36$ и $1,77$ для $x_{T(0,25)}$, $1,21$; $1,46$ и $1,59$ для $x_{T(0,1)}$ и т. д.

Вычисление этих оценок более трудоемко, чем расчет \bar{x} , поскольку требуется упорядочить наблюдения для построения $x_{T(x)}$, $x_{w(x)}$ и x_{H-L} ; построение вариационного ряда требует порядка $n \ln n$ операций [47], а для вычисления x_{H-L} необходимо число действий порядка $n^2 \ln n$. Оценка Хубера требует значительных вычислений при минимизации Φ , однако она имеет важное преимущество перед (4) — (6): оценивание параметра сдвига М-оценкой можно проводить, внося уточнения в оценку по мере поступления информации [38], что существенно при on-line-обработке.

Однако рассмотренные оценки имеют недостаток: для их оптимального использования требуется знать величину «загрязнения» ε (для выбора параметра α в (4), (5), для задания $k(\varepsilon)$ в М-оценке), обычно в практике неизвестную. Поэтому весьма перспективным представляется использование адаптивных робастных оценок. Так, в [121] предложено выбирать α для α -«обрубленных» оценок (4) из условия минимума α -«винзоризованной» оценки дисперсии. Хоггом [116—118] была предложена процедура выбора оценки в зависимости от значения выборочного эксцесса. Теми же причинами вызвано применение итеративных робастных оценок. Л. Д. Мешалкин [59; 126] предложил способ последовательного построения оценок математического ожидания и дисперсии (ЭВ-оценки [2]), используя ф. п. в. нормального закона: если $\hat{\theta}$ и \hat{s}^2 — робастные оценки сдвига и дисперсии на l -м шаге, то на $(l+1)$ -м шаге оценки вычисляются как

$$\hat{\theta}' = \frac{1}{\omega} \sum_{i=1}^n \omega_i x_i, \quad \hat{s}'^2 = \frac{1}{(n-1)\omega} \sum_{i=1}^n \omega_i (x_i - \hat{\theta})^2, \quad (8)$$

где

$$\omega_i = \left[\frac{1}{\sqrt{2\pi} \hat{s}} \exp \left\{ -\frac{(x_i - \hat{\theta})^2}{2\hat{s}^2} \right\} \right]^\lambda, \quad \omega = \sum_{i=1}^n \omega_i,$$

λ характеризует быстроту спадания весовых множителей с удалением от центра. На первом шаге $\hat{\theta} = \bar{x}$, $\hat{s}^2 = s^2$. При $\lambda \approx 1,5$ процедура обычно сходится после 3—5 итераций. Мостеллер и Тьюки [61] рекомендуют использовать бивес-оценку сдвига

$$x^* = \left(\sum_{i=1}^n \omega_i \right)^{-1} \sum_{i=1}^n \omega_i x_i,$$

где

$$\omega_i = \begin{cases} \left[1 - \left(\frac{x_i - x^*}{cs} \right)^2 \right]^2, & \text{если } |x_i - x^*|/cs < 1, \\ 0 & \text{в противном случае,} \end{cases}$$

а $s = \text{median } |x_i - x^*|$ или s есть половина межквартильного размаха [44], константа c часто берется равной 6 и 9.

Другая трудность при применении робастных оценок вызвана тем, что, поскольку механизм появления аномальных наблюдений, «загрязнения» иной, чем у основной доли наблюдений, центр загрязняющего распределения обычно не совпадает с центром основного.

В этих условиях, например, М-оценка Хубера, полученная для модели (7) с совпадающими центрами, не очень эффективна для оценивания положения центра, а адаптивные и итерационные оценки обнаруживают свое преимущество [10; 100].

Свойства робастных оценок многомерного параметра сдвига изучены меньше, чем одномерного. Однако для оценок по компонентам типа (5) показано теоретически [104], а для $x_{j(\alpha)}$ и $x_{w(\alpha)}$ методом Монте-Карло [23; 90], что в случае $k \geq 3$ для k -мерного нормального распределения с корреляционной матрицей ρ эффективность робастных оценок относительно \bar{x} произвольно мала при $\det \rho \rightarrow 0$. Итерационная оценка (8) может быть обобщена на многомерную ситуацию [59], если в w_i использовать ф. п. в. k -мерного нормального распределения. На модельных примерах она показала высокую эффективность. Тем не менее вопрос о парировании многомерных «загрязнений» требует дальнейших исследований.

Более подробно современное состояние в области робастных оценок и методов рассматривается в [88], там же даны необходимые литературные ссылки для углубленного изучения интересующих читателя вопросов. О применении робастных оценок в линейных моделях регрессии вкратце рассказано ниже.

Глава 6

ЛИНЕЙНЫЙ РЕГРЕССИОННЫЙ АНАЛИЗ

Нередко величины Y_1, \dots, Y_n , изучаемые в эксперименте, зависят от значений другой величины x , которая либо известна точно, либо может быть с высокой точностью измерена. Тогда физическая модель,

характеризующая зависимость Y от x , приводит к выражению

$$Y = \varphi(x, \vec{\theta}), \quad (1)$$

где x — одна или несколько величин, которые могут контролированно изменяться в ходе эксперимента, $\vec{\theta}$ — вектор обычно неизвестных параметров $\theta_1, \dots, \theta_k$, вид функции φ предполагается известным. В действительности из-за ошибок измерений регистрируется не Y , а $y = \varphi + \varepsilon$, и если ошибка ε случайная (т. е. ее м.о. равно 0), то (1) выполняется для $Ey = \varphi$. Изменяя контролируемые переменные x и регистрируя соответствующее y , получим в итоге вектор

$$\vec{y} = \vec{\varphi}(x, \vec{\theta}) + \vec{\varepsilon}. \quad (2)$$

Тогда можно попытаться оценить вектор параметров $\vec{\theta}$ по вектору наблюдений \vec{y} (задача оценивания параметров модели).

Например, в задачах активационного анализа смеси радиоактивных изотопов в моменты времени t_1, \dots, t_n регистрируется число распадов ядер y , для математического ожидания которого справедливо экспоненциальное представление

$$Ey(t_i) = \sum_{j=1}^m A_j e^{-\lambda_j t_i},$$

где λ_j характеризует быстроту распада j -й компоненты смеси, A_j — ее количество. Величина y , как обычно в ядерной физике, подчиняется закону Пуассона с параметром $Ey(t_i)$, и если среднее число распадов велико, то пуассоновское распределение переходит в нормальное, так что

$$y(t_i) = \sum_{j=1}^m A_j e^{-\lambda_j t_i} + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sqrt{Ey(t_i)}).$$

Имея достаточное количество измерений, можно попытаться определить $k = 2m$ параметров модели A_j, λ_j .

Близка к этой постановке и задача свертки (редукции) данных и интерполяции: на сетке $\langle x_i \rangle$, $i = \overline{1, n}$, аргумента заданы значения y_1, \dots, y_n вели-

чины y . Желательно представить данные в компактном виде (например, получить полиномиальное представление $y = \hat{P}(x; a_0, a_1, \dots, a_{k-1})$, a_i — коэффициенты полинома, подлежащие определению по y) для удобства хранения и возможности интерполяции при промежуточных значениях аргумента.

Указанные две задачи различаются по существу: в первой модель определяется изучаемым процессом, во второй — произвольна и обычно выбирается по виду графика $y(x)$ из соображений удобства. С другой стороны, в первой задаче обычно имеется информация о характере ошибок (их дисперсии, законе распределения и т. д.), что позволяет, найдя оценки параметров $\vec{\theta}$, ставить вопрос об их точности, об адекватности модели экспериментальным данным и т. д. Во второй постановке никакой статистической информации не заложено, и качество приближения данных оценивается субъективно. Однако с точки зрения техники вычислений при нахождении параметров обе задачи эквивалентны и обычно сводятся при $n > k$ к минимизации взвешенной с весами w_i суммы квадратов отклонений

$$\sum_{i=1}^n w_i (y_i - \varphi(x_i, \vec{\theta}))^2. \quad (3)$$

Простейшим и важнейшим для практики является случай линейной зависимости $\varphi(x, \vec{\theta})$ от параметров

$$\vec{\varphi} = X\vec{\theta}, \quad (4)$$

где $\vec{\theta} = (\theta_1, \dots, \theta_k)^T$ — вектор параметров (коэффициентов регрессии), X — матрица ($n \times k$), j -й столбец которой представляет собой значения j -й контролируемой переменной (регрессора) в 1-м, ..., n -м измерениях, $\vec{\varphi} = (\varphi_1, \dots, \varphi_n)^T$ — вектор математических ожиданий наблюдаемой величины y (отклика) в 1-м, ..., n -м измерениях.

Линейная модель (4) глубоко исследована теоретически, разработаны эффективные алгоритмы и программы для решения связанных с ней задач. Многие зависимости описываются в ее рамках. Например, в

задаче активационного анализа при известном качественном составе смеси (известных λ_j) определение количественного состава (величины A_j) приводит к задаче (4). Другой пример: имеются измерения спектральных кривых $x_i(t_j)$ известных веществ (t_j — частота или энергия, $i = \overline{1, k+1}$); измеряется спектр исследуемого образца, содержащего 1-е, ..., $k+1$ -е вещества, однако процентный состав его неизвестен. Тогда спектр образца $x(t_j)$ допускает представление

$$x(t_j) = \theta_1 x_1(t_j) + \dots + \theta_k x_k(t_j) + \theta_{k+1} x_{k+1}(t_j),$$

причем $\theta_1 + \dots + \theta_{k+1} = 1$, $\theta_i \geq 0$, $i = \overline{1, k+1}$. Таким образом, задача определения процентного состава смеси также приводит к модели (4).

Нелинейная модель (2) иногда сводится к линейной подходящим функциональным преобразованием. Так если

$$\varphi(\vec{x}_i, \vec{\theta}) = \prod_{j=1}^k \exp(\theta_j x_{ij}),$$

то переход к (4) осуществляется логарифмированием. В общем случае определение параметров нелинейной модели сводят к решению последовательности линейных задач, линеаризуя функцию φ в (2) и используя итерационную процедуру.

Итак, пусть

$$\vec{y} = X\vec{\theta} + \varepsilon. \quad (5)$$

Если закон распределения ошибок произволен (но известен), оптимальные по эффективности (при больших n) оценки среди всех несмещенных можно найти по методу МП; вообще говоря, эти оценки $\vec{\theta} = \vec{\theta}(\vec{y})$ не будут линейными по \vec{y} , несмотря на линейность (5). Если же ограничиться классом линейных оценок $\vec{\theta} = A\vec{y}$, то независимо от закона распределения ошибок оценки с оптимальными свойствами получаются при заданном виде матрицы.

Сначала рассмотрим простейший частный случай (5), важный для практики. Пусть независимым обра-

зом n раз измеряется скаляр μ , причем точность измерений различна:

$$\vec{y} = \vec{1} \cdot \mu + \varepsilon, \quad E\varepsilon = \vec{0}, \quad \text{cov}(\varepsilon, \varepsilon) = \text{diag}\{\sigma_1^2, \dots, \sigma_n^2\}.$$

Найдем оптимальную линейную оценку $\hat{\mu} = \vec{a}^T \vec{y}$, являющуюся несмещенной и имеющую минимальную дисперсию среди всех линейных оценок. Требование несмещенности $\mu = E\hat{\mu} = E(\vec{a}^T \vec{y}) = \vec{a}^T E\vec{y} = \mu \vec{a}^T \vec{1}$ приводит к условию $\sum_{i=1}^n a_i = 1$. Дисперсия линейной оценки

$\sigma^2(\hat{\mu}) = \sum_{i=1}^n \sigma_i^2 a_i^2$ должна быть минимизирована по

a_1, \dots, a_n при условии несмещенности. Метод множителей Лагранжа позволяет свести задачу к минимизации

$$F(a_1, \dots, a_n) = \sum_{i=1}^n \sigma_i^2 a_i^2 + \lambda \left(\sum_{i=1}^n a_i - 1 \right).$$

Дифференцируя F по a_1, \dots, a_n , получим

$$\frac{\partial F}{\partial a_i} = 2a_i \sigma_i^2 + \lambda = 0 \quad (i = \overline{1, n})$$

или при $\sigma_i^2 > 0$

$$2a_i + \frac{\lambda}{\sigma_i^2} = 0. \quad (6)$$

Складывая уравнения системы (6) и используя условие несмещенности, найдем

$$\lambda = -2 \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right)^{-1}, \quad a_i = \frac{1}{\sigma_i^2} \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right)^{-1}.$$

Поэтому

$$\hat{\mu} = \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right)^{-1} \sum_{i=1}^n y_i / \sigma_i^2$$

является линейной несмещенной оценкой μ с мини-

мальной дисперсией для неравноточных измерений (средневзвешенная оценка).

Аналогично только что рассмотренной задаче можно построить решение задачи (5) в классе линейных несмещенных оценок с минимальной дисперсией компонент $\hat{\theta} = Ay$, $E\hat{\theta} = \theta$, A :

$$(a_{i1}, \dots, a_{in}) = \underset{\tilde{a}_{i1}, \dots, \tilde{a}_{in}}{\text{Arg min}} D\hat{\theta}_i = \underset{\tilde{a}_{i1}, \dots, \tilde{a}_{in}}{\text{Arg min}} \sum_{l=1}^n \tilde{a}_{il}^2 \sigma_l^2,$$

$$i = \overline{1, k}.$$

Однако проще подойти к задаче (5) с другой стороны: будем искать такую оценку $\hat{\theta}$ вектора параметров, чтобы при подстановке ее в (5) сумма квадратов отклонений экспериментальных \vec{y} от предсказываемых $\vec{y} = X\hat{\theta}$ была минимальной. При этом предполагаем, что:

- 1) $E\varepsilon = \vec{0}$;
- 2) $\text{cov}(\varepsilon, \varepsilon) = \sigma^2 I_n$;
- 3) $\text{rang } X = k$;
- 4) ошибки ε_i независимы.

Минимизация по θ выражения

$$Q = \sum_{i=1}^n \left(y_i - \sum_{j=1}^k X_{ij} \theta_j \right)^2 = (\vec{y} - X\theta)^T (\vec{y} - X\theta) \quad (7)$$

приводит к исследованию решений системы

$$X^T X \theta - X^T y = 0. \quad (8)$$

В силу условия 3 матрица $X^T X$ невырожденная, так что система нормальных уравнений (8) имеет единственное решение

$$\hat{\theta} = (X^T X)^{-1} X^T y. \quad (9)$$

Очевидно, что оно соответствует минимуму (7), поскольку при $\theta_j \rightarrow \pm \infty$ Q неограниченно возрастает. Способ нахождения оценки вектора параметров из условия минимума (7) называется *методом наимень-*

ших квадратов (МНК), а оценка (9) — оценкой метода наименьших квадратов (МНК-оценкой).

Рассмотрим геометрическую интерпретацию метода наименьших квадратов. Пусть E_n — n -мерное евклидово пространство. Наблюдения y_1, \dots, y_n определяют в E_n конец вектора \vec{y} , проведенного из начала координат. Столбцы матрицы X , являющиеся в силу условия 3 независимыми n -мерными векторами $\vec{x}_1, \dots, \vec{x}_k$, порождают в E_n линейное многообразие E_k векторов $\vec{y} = \vec{\theta}_1 \vec{x}_1 + \dots + \vec{\theta}_k \vec{x}_k$. Вообще говоря, \vec{y} не принадлежит к этому многообразию. Вектор $\vec{\epsilon} = \vec{y} - X\vec{\theta}$ представляет собой наклонную из конца \vec{y} в точку $X\vec{\theta}$ многообразия. Минимизируя

$$\vec{\epsilon}^T \vec{\epsilon} = (\vec{y} - X\vec{\theta})^T (\vec{y} - X\vec{\theta})$$

в (7), мы решаем задачу нахождения наклонной минимальной длины, т. е. построения перпендикуляра из \vec{y} на плоскость E_k . При соответствующем выборе $\vec{\theta}$ вектор $\vec{\epsilon}$ перпендикулярен E_k , а значит, он ортогонален любому из $\vec{x}_j, j = \overline{1, k}$: $\vec{x}_j^T (\vec{y} - X\vec{\theta}) = 0$, откуда $X^T (\vec{y} - X\vec{\theta}) = 0$, т. е. получаем нормальные уравнения (8). В силу линейной независимости \vec{x}_j $\text{rang}(X^T X) = k$, откуда следует (9).

Рассмотрим свойства МНК-оценки $\hat{\vec{\theta}}$.

1. $\hat{\vec{\theta}}$ — несмещенная оценка $\vec{\theta}$.

Доказательство. Действительно,

$$\begin{aligned} E\hat{\vec{\theta}} &= E(X^T X)^{-1} X^T \vec{y} = (X^T X)^{-1} X^T E\vec{y} = \\ &= (X^T X)^{-1} X^T X\vec{\theta} = \vec{\theta}. \end{aligned}$$

2. Ковариационная матрица для $\hat{\vec{\theta}}$ записывается в виде

$$\text{cov}(\hat{\vec{\theta}}, \hat{\vec{\theta}}) = \sigma^2 (X^T X)^{-1}. \quad (10)$$

Доказательство. Это следует из свойства ковариационного оператора [68; 87]: если с. в. $\vec{\eta} = L\vec{\xi}$, то $\text{cov}(\vec{\eta}, \vec{\eta}) = L \text{cov}(\vec{\xi}, \vec{\xi}) L^T$. Применяя это к (9), найдем

$$\begin{aligned} \text{cov}(\vec{\hat{\theta}}, \vec{\hat{\theta}}) &= (X^T X)^{-1} X^T \text{cov}(\vec{y}, \vec{y}) X (X^T X)^{-1} = \\ &= \sigma^2 (X^T X)^{-1} X^T I_n X (X^T X)^{-1} = \\ &= \sigma^2 (X^T X)^{-1} X^T X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1}. \end{aligned}$$

3. Пусть \vec{c} — k -мерный вектор. Тогда оценка $\vec{c}^T \vec{\hat{\theta}}$ является единственной оценкой, обладающей минимальной дисперсией, в классе всех линейных несмещенных оценок линейной комбинации $\vec{c}^T \vec{\hat{\theta}}$ (такую оценку называют наилучшей линейной несмещенной оценкой (НЛНО) для $\vec{c}^T \vec{\theta}$) [34; 68; 130].

Доказательство. Несмещенность $\vec{c}^T \vec{\hat{\theta}}$ следует из свойства 1:

$$E(\vec{c}^T \vec{\hat{\theta}}) = \vec{c}^T E \vec{\hat{\theta}} = \vec{c}^T \vec{\theta}.$$

Пусть $\vec{\tilde{\theta}}$ — линейная несмещенная оценка $\vec{\theta}$, не равная $\vec{\hat{\theta}}$. В силу линейности $\vec{\tilde{\theta}} = A\vec{y}$. Условие несмещенности дает $AX = I_k$. Рассмотрим дисперсию с. в. $\vec{c}^T \vec{\tilde{\theta}}$:

$$\begin{aligned} D[\vec{c}^T \vec{\tilde{\theta}}] &= \vec{c}^T \text{cov}(\vec{\tilde{\theta}}, \vec{\tilde{\theta}}) \vec{c} = \vec{c}^T A \text{cov}(\vec{y}, \vec{y}) A^T \vec{c} = \\ &= \sigma^2 \vec{c}^T A A^T \vec{c}. \end{aligned}$$

Аналогично $D[\vec{c}^T \vec{\hat{\theta}}] = \sigma^2 \vec{c}^T (X^T X)^{-1} \vec{c}$.

Рассмотрим подробнее величину $\vec{c}^T A A^T \vec{c}$. Представив A в виде $A = (X^T X)^{-1} X^T + (X^T X)^{-1} X^T$, найдем

$$\begin{aligned} \vec{c}^T A A^T \vec{c} &= \vec{c}^T (A - (X^T X)^{-1} X^T) (A - (X^T X)^{-1} X^T)^T \vec{c} + \\ &+ \vec{c}^T (X^T X)^{-1} X^T X (X^T X)^{-1} \vec{c} + \vec{c}^T (A - \end{aligned}$$

$$\begin{aligned}
 & - (X^T X)^{-1} X^T) X (X^T X)^{-1} \vec{c} + \vec{c}^T (X^T X)^{-1} X^T (A - \\
 & - (X^T X)^{-1} X^T) \vec{c}.
 \end{aligned}$$

Так как $AX = I_k$, $(X^T X)^{-1} X^T X = I_k$, то третий и четвертый члены равны 0. Итак, $\vec{c}^T A A^T \vec{c} = \vec{c}^T (A - (X^T X)^{-1} X^T) \times \times (A - (X^T X)^{-1} X^T)^T \vec{c} + \vec{c}^T (X^T X)^{-1} \vec{c}$. Вторая квадратичная форма правой части не зависит от $\vec{\theta}$ и является положительно определенной. Первая — положительно полуопределенная и обращается в 0 при произвольном \vec{c} , если $A = (X^T X)^{-1} X^T$, т. е. минимальное значение $D[\vec{c}^T \vec{\theta}]$ достигается при $\vec{\theta} = \vec{\hat{\theta}}$.

Следствие. Полагая $c_1 = (1, 0, \dots, 0)^T$, $c_2 = (0, 1, 0, \dots, 0)^T, \dots, c_k = (0, \dots, 0, 1)^T$, получаем, что $\vec{\hat{\theta}}$ реализует НЛНО с минимальной дисперсией компонент. Это вытекает из того, что

$$D[\vec{c}_i^T \vec{\hat{\theta}}] = D\hat{\theta}_i, \quad i = \overline{1, k}.$$

Объединяя результаты, полученные в (7) — (10), с приведенным выше следствием, получим утверждение теоремы Гаусса — Маркова.

Пусть y_1, \dots, y_n — независимые с. в. со средними значениями $\sum_{j=1}^k x_{ij} \theta_j$ и дисперсиями σ^2 ; X_{ij} — элементы известной матрицы X , $\text{rank } X = k$. Тогда несмещенные линейные оценки коэффициентов регрессии с минимальной дисперсией компонент совпадают с МНК-оценками (9). При этом ковариационная матрица оценок задается выражением (10).

Замечание. Откажемся от условия 2. Пусть $\text{cov}(\vec{\varepsilon}, \vec{\varepsilon}) = \Sigma = \text{diag}\{\sigma_1^2, \dots, \sigma_n^2\}$. Тогда утверждения теоремы Гаусса — Маркова сохраняют силу с заменой (9) на

$$\vec{\hat{\theta}} = (X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1} \vec{y} \quad (11)$$

и (10) на

$$\text{cov}(\vec{\hat{\theta}}, \vec{\hat{\theta}}) = (X^T \Sigma^{-1} X)^{-1}. \quad (12)$$

Более того, если $\Sigma > 0$ — произвольная симметричная невырожденная матрица, то обобщенная МНК-оценка, вычисляемая по формуле (11) с ковариационной матрицей (12), также есть НЛНО для $\vec{c}^T \vec{\theta}$.

Вывод формул (11) — (12) опирается на разложение Холецкого матрицы $\Sigma: \Sigma = U U^T$, где U — нижняя треугольная матрица с положительными диагональными элементами [68; 91]. Тогда, введя $\vec{z} = U^{-1} \vec{y}$, $T = U^{-1} X$, $\vec{\eta} = U^{-1} \vec{\epsilon}$, получим вместо (5) представление

$$\vec{z} = T \vec{\theta} + \vec{\eta}, \quad (13)$$

для которого $E \vec{\eta} = \vec{0}$, $\text{cov}(\vec{\eta}, \vec{\eta}) = U^{-1} \Sigma (U^{-1})^T = I_n$, $\text{rang } T = \text{rang } X = k$. Поэтому развитая выше теория применима к (13). Получаем $\hat{\vec{\theta}} = (T^T T)^{-1} T^T \vec{z} =$

$$= (X^T (U^{-1})^T U^{-1} X)^{-1} X^T (U^{-1})^T U^{-1} \vec{y} = (X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1} \vec{y},$$

т. е. соотношение (11). Вычисление $\text{cov}(\hat{\vec{\theta}}, \hat{\vec{\theta}})$ для МНК-оценки в модели (13) и переход от T к X дает (12). Применяя рассуждения, аналогичные рассуждениям, использованным при доказательстве свойства 3 МНК-оценок, можно показать, что обобщенная МНК-оценка (11) является НЛНО для $\vec{c}^T \vec{\theta}$.

Нетрудно проверить, что обобщенная МНК-оценка (11) доставляет минимум квадратичной форме

$$\bar{Q} = (\vec{y} - X \vec{\theta})^T \Sigma^{-1} (\vec{y} - X \vec{\theta}); \quad (7')$$

в частном случае при $\Sigma = \text{diag}\{\sigma_1^2, \dots, \sigma_n^2\}$ минимизируется «взвешенная» квадратичная форма

$$\bar{Q} = \sum_{i=1}^n (y_i - (X \vec{\theta})_i)^2 / \sigma_i^2.$$

Отметим также, что, хотя исходные величины y_1, \dots, y_n в модели (5) предполагались независимыми, для компонент вектора $\vec{\theta}$ некоррелированность уже не имеет места, как это видно из (10). Однако в частном случае, если столбцы матрицы X взаимно ортогональные, матрица $X^T X$ будет диагональной и компоненты $\vec{\theta}$ некоррелированы.

Вернемся к рассмотрению модели (5). Обычно дисперсия заранее неизвестна и подлежит оценке по \vec{y} . Хотя вид МНК-оценки (9) не зависит от σ^2 , ее ковариационная матрица содержит множитель σ^2 . Обозначим $\vec{\hat{y}} = X\vec{\hat{\theta}} = X(X^T X)^{-1} X^T \vec{y} = P\vec{y}$,

$$\vec{\hat{\varepsilon}} = \vec{y} - \vec{\hat{y}} = (I_n - P)\vec{y}, \quad Q = \vec{\varepsilon}^T \vec{\varepsilon} = (\vec{y} - X\vec{\theta})^T (\vec{y} - X\vec{\theta}),$$

$$Q_1 = \vec{\varepsilon}^T \vec{\hat{\varepsilon}} = (\vec{y} - \vec{\hat{y}})^T (\vec{y} - \vec{\hat{y}}), \quad Q_2 = (\vec{\hat{\theta}} - \vec{\theta})^T X^T X (\vec{\hat{\theta}} - \vec{\theta}).$$

Преобразуя Q , найдем $Q = (\vec{y} - X\vec{\theta} + X\vec{\hat{\theta}} - \vec{\hat{y}})^T (\vec{y} - X\vec{\theta} + X\vec{\hat{\theta}} - \vec{\hat{y}}) = (\vec{y} - \vec{\hat{y}})^T (\vec{y} - \vec{\hat{y}}) + (\vec{\hat{\theta}} - \vec{\theta})^T X^T X (\vec{\hat{\theta}} - \vec{\theta}) + 2(\vec{y} - \vec{\hat{y}})^T X (\vec{\hat{\theta}} - \vec{\theta})$.

Последнее слагаемое равно 0 в силу (8), так что

$$Q = Q_1 + Q_2. \quad (14)$$

Для дальнейшего нам понадобятся некоторые понятия и факты из линейной алгебры [68].

Определение 1. Матрица P называется идемпотентной, если $P^2 = P$.

Определение 2. Симметричная идемпотентная матрица называется проекционной.

Определение 3. Следом матрицы A размера $n \times k$ называется сумма ее диагональных элементов:

$$\text{tr } A = \sum_{i=1}^{\min\{n, k\}} a_{ii}.$$

Отсюда легко получить, что: а) $\text{tr}(A+B) = \text{tr } A + \text{tr } B$; б) $\text{tr } AB = \text{tr } BA$ (размеры матриц A и B таковы, что соответствующие операции над матрицами определены); в) если A — симметричная матрица, то $\text{tr } A$ равен сумме ее собственных значений.

Легко проверить также, что у симметричной матрицы A $\text{rank } A$ равен числу ненулевых собственных значений. Действительно, найдется вещественная ортогональная матрица C такая, что $C^T A C = \Lambda$, где Λ — диагональная матрица. Но для невырожденных матриц R и S $\text{rank}[RAS] = \text{rank } A$, ибо $\text{rank } A \geq \text{rank}[AS] \geq \text{rank}[ASS^{-1}] = \text{rank } A$, откуда $\text{rank } A = \text{rank}[AS]$.

Аналогичные преобразования проводим для R и получаем $\text{rang } A = \text{rang } \Lambda$, а для матрицы Λ утверждение очевидно.

Лемма 1. Для того чтобы симметричная матрица P размера $n \times n$ была идемпотентной ранга r , необходимо и достаточно, чтобы r ее собственных значений равнялись 1, а $n-r$ — нулю.

Доказательство. Пусть $P^2 = P$ и $\text{rang } P = r$. Тогда из $P\vec{\xi} = \lambda\vec{\xi}$ следует $\lambda\vec{\xi}^T\vec{\xi} = \vec{\xi}^T P\vec{\xi} = \vec{\xi}^T P^2\vec{\xi} = (P\vec{\xi})^T \times (P\vec{\xi}) = \lambda^2\vec{\xi}^T\vec{\xi}$, откуда получаем $\lambda(\lambda-1) = 0$, т. е. все собственные значения равны 0 или 1. Но у симметричной матрицы P $\text{rang } P$ равен числу ненулевых собственных значений, поэтому ровно r собственных значений равны 1, а $n-r$ — нулю.

Докажем обратное утверждение. Пусть у P r собственных значений равны 1, а $n-r$ — нулю; без ограничения общности предположим, что отличны от 0 первые r собственных значений. Найдется ортогональная матрица C такая, что

$$C^T P C = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \Lambda$$

или $P = C \Lambda C^T$, откуда получаем $P^2 = C \Lambda C^T C \Lambda C^T = C \Lambda^2 C^T = C \Lambda C^T = P$ и $\text{rang } P = r$.

Лемма 2. Пусть P — проекционная матрица. Тогда $\text{rang } P = \text{tr } P$.

Доказательство. В силу леммы 1 если $\text{rang } P = r$, то r собственных значений P равны 1, а $n-r$ — нулю. Поскольку $\text{tr } P = \sum \lambda_i = r$, то $\text{rang } P = \text{tr } P$.

Лемма 3. Пусть P идемпотентна. Тогда $I - P$ также идемпотентна.

Доказательство.

$$(I - P)^2 = I - 2P + P^2 = I - 2P + P = I - P.$$

Применим это к МНК-оценкам. Легко проверить, что матрица $P = X(X^T X)^{-1} X^T$ — проекционная, следовательно $I_n - P$ — также проекционная. Но она определяет квадратичную форму

$$Q_1 = \vec{\varepsilon}^T \vec{\varepsilon} = \vec{y}^T (I_n - P) (I_n - P) \vec{y} = \vec{y}^T (I_n - P) \vec{y} = \\ = (X\vec{\theta} + \vec{\varepsilon})^T (I_n - P) (X\vec{\theta} + \vec{\varepsilon}) = \vec{\varepsilon}^T (I_n - P) \vec{\varepsilon},$$

поскольку $(I_n - P)X = X - X(X^T X)^{-1} X^T X = 0$. Записав $\vec{\varepsilon}^T (I_n - P) \vec{\varepsilon}$ в виде

$$\sum_i (I_n - P)_{ii} \varepsilon_i^2 + \sum_{i, j: i \neq j} \varepsilon_i \varepsilon_j (I_n - P)_{ij}$$

и используя независимость и равенство дисперсий компонент ε , получим $E[\vec{\varepsilon}^T (I_n - P) \vec{\varepsilon}] = \sigma^2 \text{tr}[I_n - P] = \sigma^2(n - \text{tr} P) = \sigma^2(n - \text{tr}[X(X^T X)^{-1} X^T]) = \sigma^2(n - \text{tr}[X^T X (X^T X)^{-1}]) = \sigma^2(n - \text{tr} I_k) = \sigma^2(n - k)$ (использованы свойства а) и б) следа матрицы). Отсюда получаем, что $E\left(\frac{Q_1}{n - k}\right) = \sigma^2$, т. е. $Q_1/(n - k)$ — несмещенная оценка для σ^2 . Величина $Q_1 = \vec{\varepsilon}^T \vec{\varepsilon}$ называется остаточной суммой квадратов (*RSS* — residual sum of squares).

Глава 7

ЛИНЕЙНЫЙ РЕГРЕССИОННЫЙ АНАЛИЗ В СЛУЧАЕ НОРМАЛЬНОСТИ. АНАЛИЗ ОСТАТКОВ

До сих пор на распределение ошибок ε в линейной модели были наложены весьма слабые ограничения: а) $E\varepsilon = 0$, б) $D\varepsilon_i < \infty$. При этих условиях доказана оптимальность МНК-оценок в классе всех линейных несмещенных оценок. Однако если рассмотреть класс всех несмещенных оценок, МНК-оценка, вообще говоря, не является наилучшей. Если же ограничить рассмотрение нормальным распределением ошибок, метод наименьших квадратов приводит к (асимптотически) эффективным оценкам, так как он совпадает с методом МП. Действительно, логарифм ФП для нормального распределения равен

$$\ln L(\vec{y}|\vec{\theta}) = \text{const} - \frac{1}{2} (\vec{y} - E\vec{y})^T \Sigma^{-1} (\vec{y} - E\vec{y});$$

$$E\vec{y} = X\vec{\theta}, \text{cov}(\vec{y}, \vec{y}) = \Sigma \quad (1)$$

и максимизация логарифмической ФП эквивалентна минимизации квадратичной формы $(\vec{y} - X\vec{\theta})^T \Sigma^{-1} \times \times (\vec{y} - X\vec{\theta})$ метода наименьших квадратов.

Более того, в многомерном аналоге теоремы Крамера-Рао показывается [45; 87], что условие достижимости нижней границы ковариационной матрицы оценок параметров записывается в виде

$$\partial \ln L / \partial \vec{\theta} = K(\vec{\theta}) (\vec{\hat{\theta}} - \vec{\theta}),$$

где $K(\vec{\theta})$ — матрица, не зависящая от \vec{y} ; оценка $\vec{\hat{\theta}}$, для которой выполнено это условие, является *эффeктивной*. В нашем случае для МНК-оценки $\vec{\hat{\theta}}$ имеем

$$\begin{aligned} \partial \ln L / \partial \vec{\theta} &= -\frac{1}{2} \partial / \partial \vec{\theta} \{ (\vec{y} - X\vec{\theta})^T \Sigma^{-1} (\vec{y} - X\vec{\theta}) \} = \\ &= X^T \Sigma^{-1} \vec{y} - X^T \Sigma^{-1} X \vec{\theta} = X^T \Sigma^{-1} X (\vec{\hat{\theta}} - \vec{\theta}). \end{aligned}$$

В дальнейшем ограничимся случаем некоррелированных одинаковых ошибок, т. е. будем считать \vec{y} многомерной нормальной с. в.: $\vec{y} \sim N(X\vec{\theta}, \sigma^2 I_n)$; чаще всего σ^2 неизвестно. Из вида МНК-оценки $\vec{\hat{\theta}} = (X^T X)^{-1} X^T \vec{y}$ следует нормальность $\vec{\hat{\theta}} \sim N(\vec{\theta}, (X^T X)^{-1} \sigma^2)$. Рассмотрим теперь квадратичные формы Q , Q_1 и Q_2 . Поскольку $\vec{\varepsilon} \sim N_n(\vec{0}, \sigma^2 I_n)$, то с. в. $Q/\sigma^2 = \vec{\varepsilon}^T \vec{\varepsilon} / \sigma^2 \sim \chi_n^2$. Случайная величина $\vec{\hat{\theta}} - \vec{\theta} \sim N_k(\vec{0}, (X^T X)^{-1} \sigma^2)$, поэтому, как известно из свойств нормального и χ^2 -распределений [68; 87], $(\vec{\hat{\theta}} - \vec{\theta})^T [(X^T X)^{-1} \sigma^2]^{-1} (\vec{\hat{\theta}} - \vec{\theta}) = (\vec{\hat{\theta}} - \vec{\theta})^T X^T X (\vec{\hat{\theta}} - \vec{\theta}) / \sigma^2 = Q_2 / \sigma^2 \sim \chi_k^2$. Наконец, $Q_1 = (\vec{y} - X\vec{\hat{\theta}})^T (\vec{y} - X\vec{\hat{\theta}}) = \vec{\varepsilon}^T (I_n - X(X^T X)^{-1} X^T) \vec{\varepsilon} = \vec{\varepsilon}^T (I_n - P) \vec{\varepsilon}$, причем, как доказано ранее, матрица $(I_n - P)$ проекционная. В лемме 1 гл. 6 показано, что число собственных значений проекционной матрицы, равных 1, совпадает с рангом

матрицы, т. е. равно $n-k$ для $I_n - P$. Вектор $\vec{\varepsilon}/\sigma \sim \sim N_n(0, I_n)$. Диагонализуя матрицу $I_n - P$ невырожденным ортогональным преобразованием $C: \Lambda = C(I_n - P)C^T$ и умножая $\vec{\varepsilon}$ слева на $C: \vec{\varepsilon} = C\vec{\varepsilon}$, мы сохраняем Q_1 без изменений. При этом

$$\text{cov}(\vec{\varepsilon}, \vec{\varepsilon}) = C \text{cov}(\vec{\varepsilon}, \vec{\varepsilon}) C^T = \sigma^2 C I_n C^T = \sigma^2 I_n,$$

так что

$$Q_1/\sigma^2 = \sum_{j=1}^n \lambda_j \vec{\varepsilon}_j^2/\sigma^2 = \sum_{j=1}^{n-k} \vec{\varepsilon}_j^2/\sigma^2 \sim \chi_{n-k}^2,$$

так как Λ — диагональная матрица с $n-k$ «единицами».

Наконец, легко видеть, что из $(I_n - P)X = 0$ следует

$$\begin{aligned} E[\vec{\theta}(\vec{y} - X\vec{\theta})^T] &= E[(X^T X)^{-1} X^T \vec{y} \vec{y}^T (I_n - P)^T] = \\ &= E[(X^T X)^{-1} X^T (\vec{y} - X\vec{\theta} + X\vec{\theta}) (\vec{y} - X\vec{\theta} + X\vec{\theta})^T (I_n - P)^T] = \\ &= (X^T X)^{-1} X^T \{E(\vec{y} - X\vec{\theta})(\vec{y} - X\vec{\theta})^T\} (I_n - P)^T + \\ &+ E[(X^T X)^{-1} X^T X\vec{\theta}(\vec{y} - X\vec{\theta})^T (I_n - P)^T] + \\ &+ E[(X^T X)^{-1} X^T (\vec{y} - X\vec{\theta}) \vec{\theta}^T X^T (I_n - P)^T] + \\ &+ E[(X^T X)^{-1} X^T X\vec{\theta}\vec{\theta}^T X^T (I_n - P)^T] = \\ &= (X^T X)^{-1} X^T \text{cov}(\vec{y}, \vec{y}) (I_n - P)^T = \\ &= \sigma^2 (X^T X)^{-1} X^T (I_n - P)^T = 0. \end{aligned}$$

Поскольку $E(\vec{y} - X\vec{\theta}) = \vec{0}$, то $\text{cov}(\vec{\theta}, \vec{y} - X\vec{\theta}) = 0$, что означает независимость нормальных с. в. $\vec{\theta}$ и $\vec{y} - X\vec{\theta}$.

Из вышесказанного следует, что величины $\vec{\theta}$ и Q_1 , Q_2 и Q_1 статистически независимы.

Доказанное позволяет строить интервальные оценки для регрессионных коэффициентов, а также про-

верить гипотезы о модели, описывающей наблюдения. Так, из независимости Q_1 и $\hat{\theta}$ следует независимость Q_1 и $\hat{\theta}_i$ для любого $i = \overline{1, k}$. Тогда величина

$$\frac{(\hat{\theta}_i - \theta_i)}{\sigma \sqrt{(X^T X)_{ii}^{-1}}} \bigg/ \frac{\sqrt{Q_1}}{\sigma \sqrt{n-k}} = \frac{(\hat{\theta}_i - \theta_i) \sqrt{n-k}}{\sqrt{(X^T X)_{ii}^{-1} Q_1}}$$

имеет распределение Стьюдента St_{n-k} , так как числитель распределен $N(0, 1)$, $Q_1/\sigma^2 \sim \chi^2_{n-k}$ и они независимы. Тогда, задавшись $0 < \gamma < 1$ и найдя $t_{n-k, \gamma}$ из уравнения

$$\int_{-t_{n-k, \gamma}}^{t_{n-k, \gamma}} f_{n-k}(\tau) d\tau = \gamma$$

(f_{n-k} — ф. п. в. для ф. р. St_{n-k}), построим $100 \cdot \gamma$ -процентный доверительный интервал T_γ :

$$\hat{\theta}_i - t_{n-k, \gamma} \sqrt{\frac{(X^T X)_{ii}^{-1} Q_1}{n-k}} \leq \theta_i \leq \hat{\theta}_i + t_{n-k, \gamma} \sqrt{\frac{(X^T X)_{ii}^{-1} Q_1}{n-k}}$$

с вероятностью γ накрывающий истинное значение i -й компоненты θ .

Если рассматривать всю совокупность компонент θ , то для них доверительная область может быть построена исходя из того, что

$$\frac{Q_2}{k\sigma^2} \bigg/ \frac{Q_1}{(n-k)\sigma^2} = \frac{(n-k)Q_2}{kQ_1} \quad (4)$$

имеет распределение Снедекора — Фишера [44; 71; 87]. Как и выше, задавшись размером γ и найдя из уравнения

$$\int_0^{F_{k, n-k, \gamma}} dF_{k, n-k}(\tau) = \gamma$$

значение $F_{k, n-k, \gamma}$ ($F_{k, n-k}$ — ф. п. в. для распределения Снедекора — Фишера $Sn(k, n-k)$), получим до-

верительную область, расписав Q_2 . Тогда случайный эллипсоид R_T :

$$(\hat{\theta} - \vec{\theta})^T X^T X (\hat{\theta} - \vec{\theta}) = \frac{k Q_1 F_{k, n-k, \gamma}}{n-k}, \quad (5)$$

накрывает истинное значение $\vec{\theta}$ с вероятностью $P\{\hat{\theta} \in R_T\} = \gamma$.

Нередко требуется на основании наблюдений сделать выбор между более простой

$$E\vec{y} = X^{(1)}\vec{\theta}^{(1)} \quad (6)$$

и более сложной

$$E\vec{y} = X\vec{\theta} \quad (7)$$

моделями, где $\vec{\theta}^{(1)}$ — k_1 -мерный, $\vec{\theta}$ — k -мерный векторы ($k_1 < k$), X представляет собой $X^{(1)}$ с добавлением ($k - k_1$) вектор-столбцов длины n . Задачу различения (7) и (6) можно поставить как задачу проверки параметрической гипотезы \mathcal{H} :

$$\begin{aligned} \theta_{k_1+1} = \theta_{k_1+2} = \dots = \theta_k = 0, \quad \sigma^2 > 0, \\ -\infty < \theta_i < \infty, \quad i = \overline{1, k_1}. \end{aligned} \quad (8)$$

Для проверки \mathcal{H} применим критерий ОП. Функция правдоподобия

$$L(\vec{y} | \vec{\theta}) = (2\pi\sigma^2)^{-n/2} \exp \left\{ -\frac{1}{2\sigma^2} (\vec{y} - X\vec{\theta})^T (\vec{y} - X\vec{\theta}) \right\} \quad (9)$$

имеет безусловный экстремум в точке, определяемой уравнениями

$$\begin{aligned} \partial L / \partial \vec{\theta} &= L \partial \ln L / \partial \vec{\theta} = \\ &= -\frac{L}{2\sigma^2} \partial / \partial \vec{\theta} \{ (\vec{y} - X\vec{\theta})^T (\vec{y} - X\vec{\theta}) \} = 0, \end{aligned} \quad (10)$$

$$\begin{aligned} \partial L / \partial \sigma^2 &= -\frac{nL}{2\sigma^2} + \\ &+ L \partial / \partial \sigma^2 \left\{ -\frac{1}{2\sigma^2} (\vec{y} - X\vec{\theta})^T (\vec{y} - X\vec{\theta}) \right\} = 0, \end{aligned} \quad (11)$$

которые имеют решение

$$\hat{\theta} = (X^T X)^{-1} X^T \vec{y}, \quad \hat{\sigma}^2 = \frac{1}{n} (\vec{y} - X\hat{\theta})^T (\vec{y} - X\hat{\theta}). \quad (12)$$

При этом значение L в точке, задаваемой соотношением (12), равно $(2\pi\hat{\sigma}^2 e)^{-n/2}$. Аналогично, максимизируя (9) при условии (8), найдем

$$\max L(\vec{y} | \hat{\theta}^{(1)}) = (2\pi\hat{\sigma}_{(1)}^2 e)^{-n/2},$$

где

$$\hat{\sigma}_{(1)}^2 = \frac{1}{n} (\vec{y} - X^{(1)}\hat{\theta}^{(1)})^T (\vec{y} - X^{(1)}\hat{\theta}^{(1)}). \quad (13)$$

Искомая статистика отношения правдоподобия

$$l = (\hat{\sigma}^2 / \hat{\sigma}_{(1)}^2)^{n/2} \quad (14)$$

может быть преобразована в

$$\tau = (l^{-n/2} - 1) = (\hat{\sigma}_{(1)}^2 - \hat{\sigma}^2) / \hat{\sigma}^2. \quad (15)$$

Это преобразование обеспечивает монотонность τ по l . Можно показать [45; 68], что при истинности \mathcal{H} в (15) стоит отношение суммы квадратов $k-k_1$ независимых $N(0, 1)$ с. в. к независимой сумме квадратов $n-k$ также независимых и $N(0, 1)$ с. в., т. е. τ есть отношение независимых χ^2 -величин с $k-k_1$ и $n-k$ степенями свободы. Поэтому, если \mathcal{H} верна, величина $F = \tau(n-k)/(k-k_1)$ имеет распределение Снедекора — Фишера $Sn(k-k_1, n-k)$. Задавшись вероятностью γ и найдя 100· γ -процентную точку $F_{k-k_1, n-k, \gamma}$ распределения Снедекора — Фишера, отвергнем \mathcal{H} , если $F > F_{k-k_1, n-k, \gamma}$.

К рассмотренному случаю можно свести также проверку гипотезы [79] $\theta_{k_1+1} = \theta_{k_1+1}^{(0)}, \dots, \theta_k = \theta_k^{(0)}$ (достаточно при $s > k_1$ перейти к параметрам $\theta_s = \theta_s - \theta_s^{(0)}$, $s = k_1+1, k$) и общей линейной гипотезы $C\vec{\theta} = \vec{\gamma}_0, \vec{\gamma}_0$ — вектор длины $k_1 < k$ (подробнее см. [68]).

При использовании метода наименьших квадратов для проверки адекватности модели экспериментальным данным, как мы видели, применяются суммы квадратов отклонений (невязка). Нередки ситуации,

когда модель хорошо приближает наблюдения по критерию невязки, но есть большие области значений аргумента, где отклонения модели от наблюдаемых значений имеют один знак, например + + + + ... + + — — — — ... — . Такие случаи могут указывать на неадекватность модели и требуют специального рассмотрения: могут ли систематические отклонения объясняться случайностью или же это маловероятно. Проверка случайности отклонений при произвольном виде (непрерывного) распределения остатков $\hat{\varepsilon} = y - X\hat{\theta}$ производится с помощью свободных от распределения критериев, например критерия серий Вальда — Вольфовитца [45]. Серия определяется как последовательность элементов одного вида, непосредственно перед которой и после которой стоят элементы другого вида. Если M — число элементов первого вида, N — число элементов второго вида, R — число серий во всей последовательности, то можно показать, что

$$E(R) = 1 + \frac{2MN}{M+N}, \quad D(R) = \frac{2MN(2MN - M - N)}{(M+N)^2(M+N-1)}; \quad (16)$$

распределение R затабулировано [34; 135] и при $M, N > 10$ хорошо аппроксимируется нормальным $N(E(R), \sqrt{D(R)})$. В нашей задаче M — число положительных, N — число отрицательных остатков (предполагается, что нулей немного и они поровну распределены между M и N). Тогда значимость отклонений числа серий от $E(R)$ (а также сильное неравенство M и N) указывает на неслучайный характер остатков.

Однако эффективность критерия серий невелика. Поэтому обычно используют другие критерии, например ранговые [17; 33; 45]. Пусть получены остатки $\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n$. Имеем последовательность их номеров $1, 2, \dots, n$ и последовательность рангов $i_1, i_2, \dots, i_n, i_j$ — номер остатка $\hat{\varepsilon}_j$ в вариационном ряде $\hat{\varepsilon}_{(1)} \leq \hat{\varepsilon}_{(2)} \leq \dots \leq \hat{\varepsilon}_{(n)}$. Ранговый коэффициент корреляции Спирмена определяется как

$$R = \left[\frac{1}{n} \sum_{j=1}^n j i_j - \left(\frac{n+1}{2} \right)^2 \right] / \left(\frac{n^2-1}{12} \right). \quad (17)$$

Если отклонения $\hat{\varepsilon}_i$ носят случайный характер, то R близок к 0; распределение статистики (17) затабулировано и не зависит от закона распределения $\hat{\varepsilon}_i$. Значимость отличия R от 0 указывает на неучтенную в регрессионной модели систематическую составляющую в $E\hat{y}$. Более сложные критерии основаны на подсчете числа инверсий в рангах (например, t -критерий Кендалла, критерий Уилкоксона [33; 45]).

Наряду с применением к остаткам статистических критериев широко распространено графическое исследование остатков [34; 68]. Оно опирается на свойства

остатков $\hat{\varepsilon} = (I_n - P)\hat{y}$:

а) $\text{cov}(\hat{\varepsilon}, X\theta) = E[\hat{y}^T (I_n - P) X\theta] = 0$;

б) $E\hat{\varepsilon} = 0$;

в) $D\hat{\varepsilon}_i = \sigma^2 (I_n - P)_{ii}$;

г) если модель имеет вид $Ey_i = \theta_0 + \theta_1 X_{i1} + \theta_2 X_{i2} + \dots + \theta_{k-1} X_{i,k-1}$, то из первого нормального

уравнения системы (8) из гл. 6 следует $\sum_{i=1}^n \hat{\varepsilon}_i = 0$.

Строят графики зависимости «стьюдентизированных» остатков

$$d_i = \varepsilon_i / [Q_1 (1 - P_{ii})^{1/2}]$$

от \hat{y}_i , от номера наблюдения i (или от времени, если наблюдения нумеруются в порядке измерений), от каждого из регрессоров x_1, \dots, x_k . Графики остатков в первом и третьем случаях должны в идеале ложиться внутри полосы постоянной ширины (в силу свойства а)). Клинообразность или криволинейность полосы остатков, линейный или криволинейный тренд указывают на неадекватность модели и возможные коррективы [68]. На графике зависимости от времени (от номера измерения i), кроме того, выявляется коррелированность остатков (если число наблюдений n много больше числа параметров k , то уменьшение числа степеней свободы для остатков до числа степеней свободы $n - k$ невелико, поэтому корреляционные связи между остатками $\hat{\varepsilon}_i$ слабы). Обычно графическое исследование более наглядно и позволяет выявить практически все дефекты неполноты регрессионной модели [61].

ВЫЧИСЛИТЕЛЬНЫЕ АСПЕКТЫ МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ. РОБАСТНЫЙ МНК. НЕЛИНЕЙНЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

1. Сведение задачи минимизации суммы квадратов отклонений (7) из гл. 6 к решению системы нормальных уравнений (8) из гл. 6 позволило в принципе получить решение (9) из гл. 6 и исследовать его свойства. Однако при решении практических задач такой путь вряд ли может быть рекомендован. Прежде всего неясно, что делать в случае неполного ранга, $\text{rang } X < k$, ибо не определена обратная матрица к $X^T X$. В этом случае решение задачи (7) из гл. 6 не единственно и для нахождения частного решения из множества общих решений обычно применяют аппарат обобщенных обратных матриц [4; 45; 68] либо используют идентифицирующие ограничения для компонент вектора решения. Но даже в случае задачи (7) из гл. 6 с матрицей полного ранга погрешности при задании матрицы X и вычислительные погрешности, связанные с конечной разрядностью ЭВМ, приводят к трудностям при нахождении МНК-оценок. Эти вопросы подробно рассмотрены в [68; 123].

Пусть [68] экспериментальные данные приближаются полиномиальной моделью k -й степени

$$y_i = \theta_0 + \theta_1 x_i + \theta_2 x_i^2 + \dots + \theta_k x_i^k + \varepsilon_i, \quad i = \overline{1, n},$$

причем измерения x_i выполнены на равномерной сетке на $[0, 1]$. Тогда для матрицы данных $X = \{X_{ij}\} = \{x_i^j\}$ имеем при достаточно больших n

$$(X^T X)_{rs} = n \sum_{i=1}^n x_i^r x_i^s \frac{1}{n} \approx n \int_0^1 x^r x^s dx = \frac{n}{(r+s+1)}, \quad (1)$$

т. е. $X^T X$ совпадает с точностью до множителя n с матрицей $H_k = \{(r+s+1)^{-1}\}$, которая является верхней левой угловой подматрицей размера $(k+1) \times (k+1)$ матрицы Гильберта. Известно, что, скажем, при $k=9$ элементы H_k^{-1} имеют величину порядка

$3 \cdot 10^{10}$, так что ошибка порядка 10^{-10} в определении $X^T y$ приводит к ошибке порядка 3 при вычислении $\hat{\theta}$.

Вернемся к рассмотрению системы нормальных уравнений (8) из гл. 6, которую можно переписать также в виде $A\vec{\theta} = \vec{z}$, где симметричная матрица $A = X^T X$, $\vec{z} = X^T y$. Как известно, симметричную матрицу можно привести к диагональному виду ортогональным преобразованием [78] $\vec{\theta} = V\vec{\theta}^*$, $\vec{z} = V\vec{z}^*$, так что преобразованная система имеет вид

$$\lambda_i \theta_i^* = z_i^*, \quad (2)$$

причем если $\text{rank } A = k$, то $\lambda_i > 0 (i = \overline{1, k})$. Тогда компоненты вектора $\vec{\theta}^*$ определяются как $\theta_i^* = z_i^* / \lambda_i$. Если отношение максимального и минимального собственных значений $\lambda_{\max} / \lambda_{\min}$ велико, а все вычисления и измерения выполняются с одинаковой точностью, погрешности определения малых с.з. λ_i приводят к большим погрешностям при определении $\hat{\theta}$. Таким образом, даже для матрицы X полного ранга качество полученного решения может быть очень плохим.

Рассмотрим [43; 95] вначале задачу нахождения решения СЛАУ $A\vec{x} = \vec{b}$ с невырожденной ($n \times n$) матрицей A . Определим спектральную норму как

$$\|A\| = \sup_{\vec{x} \neq 0} \|\vec{Ax}\| / \|\vec{x}\|,$$

нормы векторов \vec{x} и \vec{Ax} евклидовы. Пусть A задана точно, \vec{b} известен с неопределенностью $\vec{\delta}$, $\vec{x} + \vec{\delta x}$ — решение СЛАУ $A(\vec{x} + \vec{\delta x}) = \vec{b} + \vec{\delta}$. Имеем отсюда $\vec{\delta x} = A^{-1}\vec{\delta}$, так что $\|\vec{\delta x}\| \leq \|A\|^{-1} \cdot \|\vec{\delta}\|$. Найдем относительную погрешность определения решения $\|\vec{\delta x}\| / \|\vec{x}\|$. Поскольку $\|\vec{b}\| \leq \|A\| \cdot \|\vec{x}\|$, получим $\|\vec{\delta x}\| \cdot \|\vec{b}\| \leq \|A\| \cdot \|A^{-1}\| \cdot \|\vec{x}\| \cdot \|\vec{\delta}\|$. При $\|\vec{b}\| \neq 0$ находим оценку погрешности

$$\|\vec{\delta x}\| / \|\vec{x}\| \leq \|A\| \cdot \|A^{-1}\| \cdot \|\vec{\delta}\| / \|\vec{b}\|. \quad (3)$$

Таким образом, погрешность решения определяется величиной $\|A\| \cdot \|A^{-1}\|$, причем оценка (3) является неулучшаемой [95].

Определение 1. Числом обусловленности матрицы A называется величина

$$\text{Cond}(A) = \|A\| \cdot \|A^{-1}\|. \quad (4)$$

Теорема 1. Пусть X — матрица размера $n \times k$ ранга r . Тогда существуют ортогональная матрица U размера $n \times n$, ортогональная матрица V размера $k \times k$, диагональная матрица S размера $n \times k$ с неотрицательными диагональными элементами, расположенными в невозрастающем порядке, такие, что справедливо представление

$$X = USV^T. \quad (5)$$

Доказательство этой теоремы можно найти, например, в [68; 95; 123]. Смысл матриц, входящих в (5), следующий: столбцы матрицы U образованы нормированными собственными векторами матрицы XX^T ; столбцы V образованы нормированными собственными векторами матрицы $X^T X$; диагональные элементы S , называемые сингулярными значениями, — квадратные корни из с.з. матрицы $X^T X$, расположенных в порядке невозрастания.

Теорема 1 о сингулярном разложении (5) позволяет дать определение, эквивалентное определению 1 числа обусловленности, которое переносится на прямоугольные матрицы, характерные для задачи (5) из гл. 6 линейного МНК. Пусть $s_1 \geq s_2 \geq \dots \geq s_n$ — сингулярные значения матрицы A . Нетрудно показать [95], что $\|A\| = s_1$, $\|A^{-1}\| = s_n^{-1}$. Тогда справедливо

Определение 1'. Числом обусловленности матрицы A называется величина

$$\text{Cond}(A) = s_1/s_n. \quad (4')$$

где s_1 и s_n соответственно наибольшее и наименьшее сингулярные значения матрицы A .

Очевидно, что для симметричной матрицы A ее сингулярные значения совпадают с с.з., однако введенное определение позволяет рассматривать сингулярные значения и числа обусловленности прямоугольных матриц. Из сказанного выше ясно, что

большое $\text{Cond}(X)$ приводит к большим погрешностям при определении $\hat{\theta}$.

Из (4') нетрудно убедиться, что для $A=X^T X$ $\text{Cond}(A) = (\text{Cond}(X))^2$, поэтому переход от (7) из гл. 6 к нормальным уравнениям ухудшает обусловленность задачи и сужает класс задач, решение которых возможно на ЭВМ с данной разрядностью. Поэтому последние 10—15 лет численное решение задач МНК производят исходя непосредственно из матрицы X , а не из нормальных уравнений (8) из гл. 6. Численные методы решения задач, основанные на использовании различных ортогональных преобразований, базируются на следующей теореме [123].

Теорема 2. Пусть X — матрица размера $n \times k$ и ранга $p \leq k$ и

$$X = HRV^T, \quad (6)$$

где H — ортогональная матрица размера $n \times n$; V — ортогональная матрица размера $k \times k$, R — матрица размера $n \times k$ и вида $\begin{bmatrix} R_{11} & 0 \\ 0 & 0 \end{bmatrix}$, R_{11} — матрица размера

$p \times p$ и ранга p . Пусть $\vec{g} = H^T \vec{y} = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$ $\begin{matrix} p \\ n-p \end{matrix}$.

$\vec{z} = V^T \vec{\theta} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$ $\begin{matrix} p \\ k-p \end{matrix}$, \hat{z}_1 — решение (единственное) уравнения $R_{11} z_1 = g_1$. Тогда; а) все решения задачи

минимизации $\|X\vec{\theta} - \vec{y}\|^2$ имеют вид $\hat{\vec{\theta}} = V \begin{bmatrix} \hat{z}_1 \\ z_2 \end{bmatrix}$, где z_2

произвольно; б) любое такое решение $\hat{\vec{\theta}}$ порождает вектор остатков $\vec{\varepsilon}$, удовлетворяющий соотношению

$\vec{\varepsilon} = \vec{y} - X\hat{\vec{\theta}} = H \begin{bmatrix} 0 \\ g_2 \end{bmatrix}$; в) норма $\vec{\varepsilon}$ удовлетворяет

$\|\vec{\varepsilon}\| = \|\vec{y} - X\hat{\vec{\theta}}\| = \|g_2\|$; г) единственное решение минимальной длины (нормальное решение) записывается в виде

$$\hat{\vec{\theta}} = V \begin{bmatrix} \hat{z}_1 \\ 0 \end{bmatrix}. \quad (7)$$

Доказательство. Подставляя (6) в (7) из гл. 6 и пользуясь тем, что длина вектора сохраняется при ортогональных преобразованиях, имеем $\|X\vec{\theta} - \vec{y}\|^2 = \|HRV^T\vec{\theta} - \vec{y}\|^2 = \|RV^T\vec{\theta} - H^T\vec{y}\|^2$. Используя вид R, \vec{g}, \vec{z} , заданный условиями теоремы, получим, что

$$\|RV^T\vec{\theta} - H^T\vec{y}\|^2 = \|R_{11}z_1 - g_1\|^2 + \|g_2\|^2 \quad (8)$$

для всех $\vec{\theta}$. Второе слагаемое в (8) не зависит от $\vec{\theta}$, так что минимум нормы достигается при $R_{11}z_1 = g_1$. Соответствующее решение \hat{z}_1 единственно, поскольку

ранг $R_{11} = p$. Общее решение \hat{z} имеет вид $\hat{z} = \begin{bmatrix} \hat{z}_1 \\ z_2 \end{bmatrix}$,

где z_2 произвольно. Определяя $\hat{\theta} = V\hat{z}$, получим

$$\hat{\varepsilon} = \vec{y} - X\hat{\theta} = \vec{y} - HRV^T\hat{\theta} = H(g - R\hat{z}) = H \begin{bmatrix} 0 \\ g_2 \end{bmatrix}.$$

Очевидно, что при $z_2 = 0$ длина вектора \hat{z} минимальна. При этом решение минимальной длины имеет вид (7).

Доказанная теорема позволяет строить оценки МНК, не ухудшая обусловленности матрицы X , поскольку матрицы H и V — ортогональные, так что $\text{Cond}(X) = \text{Cond}(R)$ в случае $p = k$. При неполноте ранга матрицы X теорема указывает путь построения решения задачи (7) из гл. 6, имеющего минимальную норму (нормального псевдорешения [79]).

2. При вычислении МНК-оценок используются различные виды факторизации (6) матрицы X . Так, при ортогонализации Грама — Шмидта, известной из алгебры и функционального анализа, матрица X представляется в виде [68] $X = Q_k U$, где Q_k — матрица с ортонормированными столбцами размера $n \times k$, U — верхняя треугольная матрица размера $k \times k$. Присоединяя к Q_k еще $n - k$ ортонормированных столбцов q_i , получим запись, аналогичную (6):

$$X = Q \begin{pmatrix} U \\ 0 \end{pmatrix}_{n-k}^k = QR,$$

где $Q = (Q_k : q_{k+1}, \dots, q_n)$.

Обычно при вычислениях применяют МГШ [68; 123], более точный и вычислительно устойчивый.

Видимо, более распространенными являются программы, основанные на приведении X к верхнетреугольному виду ортогональными преобразованиями отражений (преобразование Хаусхольдера) или вращений (преобразование Гивенса) [16; 43; 91]. При этом чаще используется первое ортогональное преобразование, требующее меньшего числа арифметических операций для триангуляризации X , хотя для метода Гивенса предложены модификации (см. [110]), сравнимые по скорости с методом Хаусхольдера. Обычно, чтобы избежать вычислительной неустойчивости решения в случае плохой обусловленности X , в ходе триангуляризации осуществляют некоторые перестановки столбцов, так что факторизация (6) имеет вид $X=QRP$, где P — ортогональная матрица перестановок столбцов размера $k \times k$, R — верхняя треугольная матрица, диагональные элементы которой невозрастающие по абсолютной величине. Это позволяет, задавшись некоторым допуском τ , обрывать процесс вычислений, если начиная с некоторого номера i $|R_{ii}| < \tau$, либо заменять близкую к вырожденной матрицу R на матрицу \tilde{R} , $\text{rank } R < k$, лучше обусловленную, чем R , и искать МНК-оценку для задачи неполного ранга. Соответствующие алгоритмы и программы приведены в [86; 105; 123].

Наконец, третья модификация (6) — декомпозиция по сингулярным значениям, основанная на теореме 1. Она вычислительно более трудоемка, но обеспечивает возможность подробного анализа задачи наименьших квадратов, в том числе в случае неполного ранга, а также применение регуляризирующих алгоритмов. Программы для вычисления такой SVD-декомпозиции приведены в [86; 96; 112; 123]. SVD-декомпозиция позволяет контролировать получение осмысленного решения (7) задачи (7) из гл. 6. Прежде всего, если $\text{rank } X = p < k$, последние $k-p$ сингулярные значения равны 0 (при расчете на ЭВМ с машинной точностью ϵ_1 для арифметики с плавающей точкой), и решение строится по формулам $\vec{g} = U^T \vec{y}$, $S^+ = \text{diag}\{1/s_1, \dots, 1/s_p, 0, \dots, 0\}$, $\vec{\theta} = VS^+ \vec{g}$. Если матрица X имеет полный ранг, но ее элементы измерены

с ошибкой τ (например, τ — отношение шум/сигнал), то определяется i : $s_i/s_1 \geq \tau$, $s_{i+1}/s_1 < \tau$ и с заменой S на $\tilde{S} = \text{diag}\{s_1, \dots, s_i, 0, \dots, 0\}$ строится для нее псевдообратная матрица, как описано выше. Таким образом, получается устойчивое решение.

3. Как видно из (2), (3) и описания ортогональных декомпозиций, приведенного выше, если не принимать специальных мер в случае плохой обусловленности матрицы X , тем более неполноты ее ранга,

то получаемое по МНК решение $\hat{\theta}$ неустойчиво к ошибкам задания данных и вычислительным ошибкам, а также, как правило, очень велико по норме. Эта проблема мультиколлинеарности [3; 32], связанная с наличием (точной или почти точной) линейной зависимости между столбцами матрицы данных X , остается, пока мы рассматриваем решение задачи (7) из гл. 6 в классе линейных несмещенных оценок. Отказ от требования несмещенности приводит к регуляризованным решениям, обладающим гораздо большей устойчивостью.

Так, специально для задачи наименьших квадратов были предложены [114; 115] ридж-оценки, получаемые при замене задачи (7) из гл. 6 на

$$\varphi(\vec{\theta}) = \|X\vec{\theta} - \vec{y}\|^2 + \lambda^2 \|\vec{\theta}\|^2. \quad (9)$$

Значение параметра λ выбирается визуально при решении задачи (9) исходя из того, что возрастание RSS для ридж-оценки по сравнению с МНК-оценкой должно быть незначительным, а норма полученного решения приемлемо мала. Очевидна аналогия (9) с алгоритмами регуляризации [79]. Ридж-оценка для (9) может быть легко вычислена с использованием SVD [123]. В самом деле, подставляя (5), получим

$$\varphi = \varphi(\vec{\theta}) = \|USV^T\vec{\theta} - \vec{y}\|^2 + \lambda^2 \|\vec{\theta}\|^2 = \|SV^T\vec{\theta} - \vec{g}\|^2 + \lambda^2 \|\vec{\theta}\|^2 = \|S\vec{z} - \vec{g}\|^2 + \lambda^2 \|\vec{z}\|^2,$$

где $\vec{z} = V^T\vec{\theta}$. Поскольку S — диагональная матрица, то

$$\varphi = \sum_{i=1}^k (s_i z_i - g_i)^2 + \sum_{i=k+1}^n g_i^2 + \lambda^2 \sum_{i=1}^k z_i^2.$$

Для получения ридж-оценки \vec{z}_R дифференцируем φ по z_i , $i = \overline{1, k}$ и приравняем к 0 производную. Отсюда определяем

$$z_i = \frac{g_i s_i}{s_i^2 + \lambda^2}, \quad i = \overline{1, k}, \quad \|\vec{z}_R\|^2 = \sum_{i=1}^k g_i^2 \left(\frac{s_i}{s_i^2 + \lambda^2} \right)^2. \quad (10)$$

Ридж-оценка $\vec{\theta}_R$ вычисляется как $V\vec{z}_R$, а остаточная сумма квадратов

$$RSS = \sum_{i=k+1}^n g_i^2 + \sum_{i=1}^k g_i^2 \frac{\lambda^2}{s_i^2 + \lambda^2}. \quad (11)$$

Видно, что RSS несколько больше, чем для МНК-оценки, однако норма решения $\|\vec{z}_R\|$ имеет умеренные значения, если λ достаточно велико по сравнению с последним ненулевым сингулярным значением. По формулам (10), (11) можно вычислить серию ридж-оценок, отвечающих различным λ , выполнив один раз SVD (5). Читателям рекомендуется сравнить (10), (11) с предложенными в [42] формулами вычисления квазирешений.

К задаче (9) сводится минимизация более общего выражения

$$\varphi_1(\vec{\theta}) = \|\vec{X}\vec{\theta} - \vec{y}\|^2 + \lambda^2 (\vec{\theta} - \vec{\xi})^T Q (\vec{\theta} - \vec{\xi}), \quad (12)$$

где $\vec{\xi}$ — известный вектор, близость к которому решения необходимо обеспечить, Q — заданная симметричная положительно определенная матрица. Выпишем необходимые для сведения (12) к (9) замены:

$$\vec{\theta} = \vec{\xi} + F^{-1}\vec{\theta}, \quad X = \vec{X}F^{-1}, \quad \vec{y} = \vec{y} - \vec{X}\vec{\xi},$$

где F — треугольная матрица разложения Холецкого $Q = F^T F$ [68; 91].

Таким образом, рекомендации по выбору параметра λ в (9), (12) и соответствующих ридж-оценок имеют субъективный характер. Взяв различные λ , можно получить сильно различающиеся по норме ридж-оценки. Это естественно, поскольку при формальном введении ридж-оценок никак не учитывается инфор-

мация о неопределенности задания данных. Указанного недостатка лишены другие два класса смещенных оценок для вектора $\vec{\theta}$ — байесовские и мини-максные.

При использовании байесовского подхода [65; 39; 40] вектор $\vec{\theta}$ считается случайным с известной априорной ф. п. в. $f(\vec{\theta})$. Как и прежде, $\vec{\theta}$ является ненаблюдаемым, а регистрируется (с погрешностями) его косвенное проявление \vec{y} согласно модели (5) из гл. 6. Простейший подход, приводящий к байесовским оценкам, состоит в следующем. Пусть стохастические свойства \vec{y} (при условии, что вектор параметров принял значение $\vec{\theta}$) описываются ф. п. в. $g(\vec{y}|\vec{\theta})$. Рассматривая серии опытов по регистрации \vec{y} при различных значениях $\vec{\theta}$, реализующихся в соответствии с $f(\vec{\theta})$, имеем по формуле Байеса для ф. п. в. реализации $\vec{\theta}$ (если наблюдается вектор \vec{y})

$$p(\vec{\theta}|\vec{y}) = \frac{f(\vec{\theta}) g(\vec{y}|\vec{\theta})}{\int_{R_{\theta}} g(\vec{y}|\vec{\theta}) f(\vec{\theta}) d\vec{\theta}} \quad (13)$$

Согласно принципу МП в качестве оценки $\vec{\theta}$ нужно брать наиболее вероятную, т. е. максимизирующую (13). Поскольку знаменатель, имеющий смысл нормировки, от $\vec{\theta}$ не зависит, искомая апостериорная оценка есть

$$\vec{\theta}_B = \text{Arg max}_{\vec{\theta}} f(\vec{\theta}) g(\vec{y}|\vec{\theta}); \quad (14)$$

такая оценка называется байесовской.

Например, если распределения $\vec{\theta}$ и $\vec{\epsilon}$ нормальны: $\vec{\theta} \sim N(\vec{\theta}_0, C)$, $\vec{\epsilon} \sim N(\vec{0}, \Sigma)$, то

$$p(\vec{\theta}|\vec{y}) = \text{const} \cdot \exp \left\{ -\frac{1}{2} (X\vec{\theta} - \vec{y})^T \Sigma^{-1} (X\vec{\theta} - \vec{y}) \right\} \times \\ \times \exp \left\{ -\frac{1}{2} (\vec{\theta} - \vec{\theta}_0)^T C^{-1} (\vec{\theta} - \vec{\theta}_0) \right\}, \quad (15)$$

так что максимизация (15) сводится к минимизации функционала

$$Q_B = (X\vec{\theta} - \vec{y})^T \Sigma^{-1} (X\vec{\theta} - \vec{y}) + (\vec{\theta} - \vec{\theta}_0)^T C^{-1} (\vec{\theta} - \vec{\theta}_0), \quad (16)$$

по форме напоминающего (12). Полученная байесовская оценка [58]

$$\begin{aligned} \vec{\hat{\theta}}_B = \text{Arg min}_{\vec{\theta}} Q_B = & (X^T \Sigma^{-1} X + C^{-1})^{-1} (X^T \Sigma^{-1} \vec{y} + \\ & + C^{-1} \vec{\theta}_0) \end{aligned} \quad (17)$$

в явном виде учитывает как информацию о неопределенности задания \vec{y} (матрица Σ), так и априорную информацию о локализации искомого решения (два первых момента $\vec{\theta}_0$ и C). Ковариационная матрица оценки $\vec{\hat{\theta}}_B$

$$\begin{aligned} \text{cov}[(\vec{\hat{\theta}}_B, \vec{\hat{\theta}}_B) | X, \vec{\theta}_0, \vec{\theta}] = & (X^T \Sigma^{-1} X + C^{-1})^{-1} \times \\ & \times X^T \Sigma^{-1} X (X^T \Sigma^{-1} X + C^{-1})^{-1}, \end{aligned} \quad (18)$$

а если рассматривать не фиксированное $\vec{\theta}$, а все возможные, то

$$\begin{aligned} \text{cov}[(\vec{\hat{\theta}}_B, \vec{\hat{\theta}}_B) | X, \vec{\theta}_0] = & (X^T \Sigma^{-1} X + C^{-1})^{-1} X^T \Sigma^{-1} \times \\ & \times (X C X^T + \Sigma) \Sigma^{-1} X (X^T \Sigma^{-1} X + C^{-1})^{-1}. \end{aligned} \quad (18')$$

Оценку вида (17) можно получить также, используя аппарат фильтрации [39; 40]. Подход, использующий свойства фильтра Калмана — Бьюси, описывается в гл. 14. Можно показать, что байесовское решение оптимально в среднеквадратическом смысле [39; 40; 133; 62] при точном задании априорного распределения. Если же оно известно не полностью, оптимальность имеет место лишь для весьма частных случаев.

Например, при $\vec{\theta} \sim N(\vec{\theta}_0, \sigma^2 I_k)$, где σ^2 неизвестно, оценка $\vec{\hat{\theta}}_B$ сохраняет оптимальность, если σ^2 — с. в., подчиняющаяся гамма-распределению [30].

Хотя байесовский подход явился исторически одной из первых попыток дать статистически обоснованную процедуру построения смещенных МНК-оценок, учитывающих априорную информацию о решении, область его применения ограничена. Во-первых, это связано с интерпретацией $\vec{\theta}$ как случайного вектора, что является искусственным приемом во многих задачах и вызывает обоснованные возражения.

Во-вторых, даже если стохастическая трактовка $\vec{\theta}$ справедлива, нередко вызывает затруднения задание его точного априорного распределения. Поэтому в последние годы интенсивно развивался минимаксный подход [51; 133; 52], систематически исследованный в [92]. В применении к задаче МНК он состоит в следующем. Пусть нас интересует не решение $\vec{\theta}$ задачи (5) из гл. 6, а значение некоторого оператора $B: R_k \rightarrow Z$, где Z — сепарабельное гильбертово пространство. Обычно B — линейный ограниченный оператор, так что Z — конечномерное пространство (в частности, $B=I_k$ и $z=B(\vec{\theta})=\vec{\theta}$). Пусть априорная информация о решении (5) из гл. 6 носит детерминированный характер и сводится к принадлежности $\vec{\theta}$ замкнутому множеству Θ . Задача вычисления значения B на решении (5) из гл. 6 при наличии априорной информации является основной.

Пусть пространство Y отвечает всевозможным y из (5) из гл. 6. Назовем произвольное непрерывное отображение $d: Y \rightarrow Z$ решающей процедурой для определения приближенного решения основной задачи. Пусть в пространстве Z задана вероятностная мера $P(Z)$. Определим функцию риска $R(\vec{\theta}, d)$ решающей процедуры как

$$R(\vec{\theta}, d) = E \|B(\vec{\theta}) - d(X\vec{\theta} + \varepsilon)\|^2. \quad (19)$$

К сожалению, в общем случае нельзя построить решающую процедуру с равномерно по $\vec{\theta}$ наименьшей функцией риска.

Пусть $\Omega[d] = \Omega[R(\cdot, d)]$ — положительный функционал, заданный на множестве всех допустимых (т. е. таких, для которых $z=d(X\vec{\theta} + \varepsilon)$ имеет конечный второй момент) и удовлетворяющий условиям:

1) монотонности: если для всех $\theta \in \Theta$ справедливо $R(\vec{\theta}, d_1) > R(\vec{\theta}, d_2)$, то и $\Omega[d_1] \geq \Omega[d_2]$;

2) однородности: для всех положительных α и всех допустимых решающих процедур $\Omega[\alpha R(\cdot, d)] = \alpha \Omega[R(\cdot, d)]$;

$$3) \inf_{\vec{\theta} \in \Theta} R(\vec{\theta}, d) \leq \Omega[d] = \Omega[R(\cdot, d)] \leq \sup_{\vec{\theta} \in \Theta} R(\vec{\theta}, d).$$

Назовем Ω функционалом погрешности для основной задачи, а его значение $\Omega[d]$ от функции риска решающей процедуры d — погрешностью этой процедуры. Например, при минимаксном подходе часто рассматривают функционал погрешности вида $\Omega_1[d] = \sup_{\vec{\theta} \in \Theta} R(\vec{\theta}, d)$. Будем считать решающую процедуру

с меньшим значением функционала погрешности предпочтительнее решающей процедуры с большим значением функционала, так что, например, для Ω_1 искомая оптимальная процедура определяется как $\hat{d} = \inf_d \sup_{\vec{\theta} \in \Theta} R(\vec{\theta}, d)$. Свойства получаемых при таком подходе решений исследованы в [92]. Обзор статистических методов решения некорректно поставленных задач при обработке результатов наблюдений дан в [6].

4. Наряду с вычислительными проблемами, связанными с мультиколлинеарностью данных, возникают трудности и из-за нарушения условий нормальности распределения ошибок. Уже при рассмотрении робастных оценок параметра сдвига мы видели, что выборочное среднее, будучи МНК-оценкой параметра сдвига, не является состоятельной оценкой, например, для распределения Коши, а для экспоненциального распределения с ф. п. в. $f(y) = 0,5 \exp\{-|y - \mu|\}$ уступает в эффективности выборочной медиане. Высокоэффективные оценки, устойчивые к отклонениям ошибок от нормальности, дает теория робастных оценок [88]. Ее применение к линейным моделям вида (4) из гл. 6 приводит к минимизации функционалов, отличных от квадратичного функционала МНК (7) из гл. 6.

Так, минимизация логарифма ФП для экспоненциального распределения при виде математического

ожидания y (4) из гл. 6 приводит к задаче определения параметров линейной модели из минимума функционала

$$\sum_{i=1}^n \left| y_i - \sum_{j=1}^k X_{ij} \theta_j \right| \quad (20)$$

(метод наименьших модулей [61; 62]). Однако в вычислительном отношении он сводится к последовательному решению задач МНК (7) из гл. 6. Действительно, записав (20) в виде

$$\begin{aligned} \sum_{i=1}^n \left\{ \left(y_i - \sum_{j=1}^k X_{ij} \theta_j \right)^2 / \left| y_i - \sum_{j=1}^k X_{ij} \theta_j \right| \right\} = \\ = \sum_{i=1}^n \omega_i \left(y_i - \sum_{j=1}^k X_{ij} \theta_j \right)^2. \end{aligned} \quad (20')$$

видим, что решение (20) можно организовать итерационно, определяя $\omega_i = \left| y_i - \sum_{j=1}^k X_{ij} \theta_j \right|^{-1}$ через параметры θ_j , вычисленные на предыдущем шаге, и решая задачу (7') из гл. 6 МНК для вычисления следующего приближения. Метод наименьших модулей робастен к большим остаткам, но неоправданно чувствителен к малым, так что в случае точного равенства для какого-либо i : $y_i = \sum_{j=1}^k X_{ij} \theta_j$ вычислительный процесс прерывается. Чтобы избежать этого, обычно определяют веса на $(l+1)$ -й итерации как [61]

$$\omega_i^{(l+1)} = \begin{cases} 1, & \text{если } |y_i - \widehat{y}_i^{(l)}| \leq k; \\ k / |y_i - \widehat{y}_i^{(l)}| & \text{в противном случае.} \end{cases} \quad (21)$$

Здесь $\widehat{y}_i^{(l)} = \sum_{j=1}^k X_{ij} \theta_j^{(l)}$, k — достаточно малая величина. Таким образом, минимизируется функционал вида $\sum \Psi(y - \widehat{y})$, где

$$\Psi(u) = \begin{cases} u^2, & |u| \leq k; \\ k|u|, & |u| > k. \end{cases} \quad (22)$$

Очевидно, что соотношение (22) — введенный Хубером [119] функционал для построения робастных М-

оценок. Таким образом, при подходящем выборе k , согласованном с долей «загрязнения», модификация выбора весов (21) приводит к М-оценкам параметров регрессии.

При других способах выбора весов в процедуре (20') можно получить другие робастные оценки. Так, рассматривая $u_i = (y_i - \hat{y}_i^{(l)}) / cs_i$, $c = \text{const}$, s_i — мера разброса остатков в l -й итерации, и используя бивес-веса [61] $w_i = w(u_i) = (1 - u_i^2)^2$ при $u_i^2 < 1$ и $w_i = 0$ при $u_i^2 > 1$, получим бивес-оценки параметров регрессии. Аналогичным образом можно свести к последовательности задач МНК и другие оценки, например ЭВ-оценки [3; 59].

5. Вернемся теперь к общей регрессионной модели (6) из гл. 6, в которой теперь функция $\varphi(x, \theta)$ нелинейна по θ . Для некоторых случаев, например если $\varphi(x, \theta)$ представляется в виде произведения экспонент типа $\exp\{x_i/\theta_i\}$, удастся свести (2) из гл. 6 к (4) из гл. 6 некоторым функциональным преобразованием (для произведения экспонент — логарифмированием), однако чаще всего такая линейризованная задача не является моделью МНК (5) из гл. 6, поскольку ошибки измерений входят неаддитивно [3; 7; 32]. Чаще такую линейризацию осуществить невозможно, и приходится решать задачу (2) из гл. 6 в общей постановке.

Прежде всего для нелинейной регрессии применимость МНК обоснована лишь для случая нормального распределения ошибок, ибо лишь тогда метод МП сводится к минимизации выражения вида (3) из гл. 6 и совпадает с МНК. Поэтому в дальнейшем предполагаем в модели (2) из гл. 6 нормальный закон распределения ε .

В отличие от линейного случая из нормальности распределения ошибок не вытекает нормальность распределения оценок $\hat{\theta}$ или нормальность остатков ε , поскольку зависимости $\hat{\theta} = \hat{\theta}(y)$ или $\varepsilon(y)$ теперь нелинейные. Это создает трудности при интерпретации результатов нелинейного МНК: при построении доверительных областей, проверке гипотез и т. д. [7; 32]. Поэтому основные результаты в нелинейном МНК связаны с задачей минимизации (3) из гл. 6.

Однако и эта задача имеет ряд принципиальных и вычислительных трудностей по сравнению с линейным случаем. В отличие от линейной ситуации функционал (3) из гл. 6 для общей модели не является строго выпуклым и может иметь не единственный минимум, а несколько. Возможно даже вырождение в «овраги» вдоль некоторых линий уровня. Поэтому решение задачи глобальной минимизации (3) из гл. 6 весьма трудно. Нахождение локального минимума или минимизация (3) из гл. 6 в ограниченной области изменения параметров может производиться прямым многократным вычислением значений (3) из гл. 6 и отбором $\vec{\theta}$, приводящими к наименьшим значениям функционала, с последующим поиском минимумов вблизи этих «перспективных» точек (алгоритмы типа «случайного поиска», с обучением или без [15]). Однако этот путь приводит нередко к неоправданно большим объемам вычислений.

Чаще используют градиентные и квазиградиентные методы минимизации (3) из гл. 6; достаточно полный набор их представлен в [3; 7; 31; 32; 98]. Наряду с общими методами минимизации функции многих переменных [15; 24] для минимизации функционала НК были разработаны специальные методы, учитывающие специфику (3) из гл. 6.

Так, при квадратичной аппроксимации (3) из гл. 6 (метод Ньютона) рассматривается локальное приближение Q квадратичной по $\vec{\theta}$ функцией и проводится минимизация этого аппроксимирующего функционала. Поэтому, имея на s -м шаге $\vec{\theta}^s$ — приближение к искомой оценке $\vec{\theta}$, Q_s — значение Q в точке $\vec{\theta}^s$, $\vec{\Delta\theta}$ — поправку для вычисления следующего приближения $\vec{\theta}^{s+1} = \vec{\theta}^s + \rho \vec{\Delta\theta}$, $0 < \rho \leq 1$, — регулятор шага, найдем, учитывая условие $Q'(\vec{\theta}) = 0$ в точке минимума,

$$\vec{\Delta\theta} = - (Q''(\vec{\theta}^s))^{-1} \text{grad } Q_s, \quad (23)$$

где $Q''(\vec{\theta}^s) = \left\| \frac{\partial^2 Q}{\partial \theta_i \partial \theta_j} \right\|_{\vec{\theta} = \vec{\theta}^s}$ — матрица размера $k \times k$.

Рассматривая линейризацию $\varphi(x, \vec{\theta})$ по формуле Тейлора, получим метод Ньютона — Гаусса с

$$\begin{aligned} \Delta \vec{\theta} &= -H_s^{-1} \text{grad } Q_s = \\ &= -\frac{1}{2} \left[\left(\frac{\partial \Phi}{\partial \vec{\theta}} \right)^T \left(\frac{\partial \Phi}{\partial \vec{\theta}} \right) \right]_{\vec{\theta}=\vec{\theta}^s}^{-1} \text{grad } Q_s. \end{aligned} \quad (24)$$

Вычислительно первый из методов гораздо более трудоемкий, поскольку при его реализации приходится многократно вычислять вторые производные Φ по $\vec{\theta}$. Кроме того, в методе Ньютона — Гаусса по построению обеспечена положительная полуопределенность матрицы данных в аппроксимирующей на итерациях линейной задаче МНК. Однако квадратичная аппроксимация обеспечивает лучшее приближение для Q и, как следствие, высокую скорость сходимости.

Поскольку, как отмечалось выше, в нелинейном МНК функционал не является строго выпуклым, то, меняя постановку исходной задачи, обеспечиваем выпуклость минимизируемого функционала. Так, в методе Левенберга — Марквардта [125] к положительно полуопределенной матрице H_s добавляется положительно определенная матрица $\lambda \text{diag}\{H_s\}$, так что

$$\Delta \vec{\theta} = -(H_s + \lambda_s \text{diag}\{H_s\})^{-1} \text{grad } Q_s, \quad (25)$$

$\{\lambda_s\}$ — последовательность положительных чисел. Фактически метод Левенберга — Марквардта сводится к вычислению последовательности ридж-оценок в задаче (12).

Из квазиградиентных методов упомянем алгоритм *DUD* [131], не требующий вычисления производных. Это особенно удобно, если, например, $\Phi(x, \vec{\theta})$ вычисляется из дифференциальных уравнений, так что вычисление производных по параметрам или их разностных аналогов очень трудоемко. Если в методе Ньютона — Гаусса используется приближение Φ касательными плоскостями, то в алгоритме *DUD* приближение строится секущими плоскостями, проведенными через $k+1$ значение Φ на данном и предшествующих шагах итераций.

Следует отметить, что описанные выше градиентные и квазиградиентные методы требуют хорошего знания начальных приближений параметров. В про-

тивном случае алгоритм может сойтись «не к тому» минимуму или вообще не сойтись. Перечисленные алгоритмы решают задачи безусловной минимизации. Методы условной минимизации описаны в [15; 24].

Глава 9

РЕГУЛЯРИЗИРУЮЩИЕ АЛГОРИТМЫ РЕГУЛЯРИЗОВАННЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

1. Хотя использование ридж-оценок позволило решить множество трудных в вычислительном отношении практических задач МНК, нечеткость, субъективность рекомендаций по выбору параметра λ затрудняет применение ридж-регрессии в системах полной математической обработки экспериментальных данных. Глубокое понимание существа проблемы, необходимости связи точности задания входных данных и точности вычислений с выбором решения задачи (5) из гл. 6 стало возможно с развитием теории регуляризирующих алгоритмов [72—75]. Было показано [79], что многие обратные задачи, в том числе алгебраическая задача наименьших квадратов, являются некорректно поставленными. Прежде чем давать формальное определение регуляризирующего алгоритма, поясним для задачи НК ее некорректность и основную идею метода регуляризации.

По отношению к основной вычислительной задаче линейной алгебры — решению СЛАУ $Ax = \vec{b}$ с квадратной невырожденной матрицей — задача МНК является некорректной, поскольку уравнения (4) из гл. 6 при $n > k$ являются, вообще говоря, несовместными, так что решения в классическом смысле не существует. Фактически МНК является частным случаем построения квазирешения [42]. Далее, если матрица данных в задаче МНК (5) из гл. 6 имеет неполный ранг, $\text{rang } X = p < k$, то решение, очевидно, неединственно. В самом деле, тогда столбцы X линейно зависимы, поэтому найдется такой k -вектор $\vec{\theta}_0$ ($\|\vec{\theta}_0\| > 0$), что $X\vec{\theta}_0 = 0$. Но тогда, если вектор $\vec{\theta}$ есть решение задачи минимизации (7) из гл. 6, то любой $\vec{\theta}_\gamma = \vec{\theta} + \gamma\vec{\theta}_0$, $-\infty < \gamma < +\infty$, также является решением.

Пусть теперь задача МНК для точной матрицы данных \bar{X} имеет единственное решение, но регрессоры мультиколлинеарны, так что столбцы матрицы данных \bar{X} почти линейно зависимы и в h -окрестности \bar{X} имеется матрица X неполного ранга. Обычно в задачах обработки матрица X и вектор отклика y известны неточно, с погрешностями h и δ соответственно. Но даже при точно заданной входной матрице X (например, если ее элементы — сеточные значения известных функций) конечная разрядность ЭВМ, наличие ошибок округления и их накопление в процессе вычисления решения θ приводят к неопределенности представления X на этапах вычислительной процедуры. Поэтому погрешности в матрице данных принципиально неустраняемы. В дальнейшем будем для краткости изложения говорить о погрешности задания матрицы данных и вектора отклика, подразумевая при этом как неопределенность во входных для МНК данных, так и обусловленные конечной разрядностью представления величин ошибки округления при вычислении на ЭВМ. Тогда, рассматривая различные индивидуальные системы (5) из гл. 6, отвечающие X и y , лежащим в пределах погрешности около точных данных \bar{X} , \bar{y} , получим множество возможных решений $\{\theta\}$, сколь угодно сильно различающихся по норме, так что задача МНК неустойчива к ошибкам в X и y . Этот факт имеет принципиальное значение при создании систем полной автоматизированной обработки, а также при решении задач нелинейного МНК, поскольку в обеих ситуациях обусловленность матрицы данных, используемой при решении задач линейного МНК как одного из этапов вычислительного процесса, заранее непредсказуема.

Эти рассуждения можно проиллюстрировать конкретным примером неустойчивости вычислительной задачи МНК к малым возмущениям матрицы X . Рассматривается задача для $n=k=2$, записываемая как

$$\theta_1 + 7\theta_2 = 5,$$

$$\sqrt{2}\theta_1 + \sqrt{98}\theta_2 = \sqrt{50}.$$

При решении ее на ЭВМ с машинной точностью $\varepsilon_1 = 10^{-l}$ (вследствие конечной разрядности представления чисел на ЭВМ вырожденность матрицы системы снимается) получим для $l=100, 300$ и 500 решения $\theta_1^{(100)}=0.$; $\theta_1^{(300)}=1.6$; $\theta_1^{(500)}=5.$ Поэтому индивидуальная приближенная система $(X:y)$ даже при как угодно большой точности содержит недостаточную информацию для получения устойчивого приближенного решения исходной задачи.

Возможность устойчивого решения задач МНК при массовой обработке на ЭВМ достигается за счет использования априорной информации как количественного [42], так и качественного характера [79] об искомом решении. Информация эта дает возможность отобрать решения, являющиеся более предпочтительными по некоторому критерию, что позволяет сузить множество $\{\theta\}$ (неограниченное) допустимых решений. Например, ищутся решения, удовлетворяющие (4) из гл. 6 и близкие к заданному вектору θ_0 . В частности, поскольку линейная модель (4) из гл. 6 обычно правомерна лишь в некоторой ограниченной области множества $\{\theta\}$, особое значение имеет построение нормального решения.

Таким образом, вместо отыскания решения, отвечающего данной индивидуальной системе $(X:y)$ задачи (5) из гл. 6, которое является неустойчивым к вариациям X и y , рассматривается задача построения в соответствии с априорным принципом отбора устойчивого решения для целого семейства данных.

Различные подходы к решению некорректно поставленных задач описаны в [78; 79], поэтому здесь мы лишь вкратце напомним построение РА исходя из вариационного принципа [72; 73; 78; 129].

Рассматривается операторное уравнение [79; 83]

$$Az = u, \quad z \in Z, \quad u \in U, \quad (1)$$

Z, U — метрические пространства. Пусть задача (1) нахождения z является некорректно поставленной [79], но точное уравнение $Az = \bar{u}$ имеет единственное решение $\bar{z} \in Z$. Пусть \bar{u} задано с погрешностью δ , т. е. известно такое u , что $\rho(u, \bar{u}) \leq \delta$. Регуляризирующим алгоритмом назовем оператор $R = R_\delta^a$.

сопоставляющий каждой паре (u_δ, δ) определенный элемент $z_\delta = R_\delta^\alpha(u_\delta)$, причем $\rho_z(\bar{z}, z_\delta) \rightarrow 0$ при $\delta \rightarrow 0$. Параметр $\alpha = \alpha(\delta)$ называется *параметром регуляризации* и определяет однопараметрическое семейство $\{R_\delta^\alpha\}$ регуляризирующих операторов.

Пусть $\Omega[z]$ — непрерывный неотрицательный функционал, определенный на всюду плотном на Z подмножестве Z_1 , причем:

а) $\bar{z} \in Z_1$;

б) для всякого числа $d > 0$ множество $Z_{1,d} = \{z: \Omega[z] \leq d\}$ есть компакт на Z_1 .

Такие функционалы называются *стабилизирующими*. Рассматривая множества $Q_\delta = \{z: \rho_U(Az, u_\delta) \leq \delta\}$ и $Z_{1,\delta} = Q_\delta \cap Z_1$, можно показать [79], что минимизация $\Omega[z]$ на $Z_{1,\delta}$, имеющая решением $z_\delta = \text{Arg inf } \Omega[z]$ на $Z_{1,\delta}$, дает способ построения РА. Заметим, что при $M_0 = \{z \in Z_1: \Omega[z] = \inf_{z \in Z_1} \Omega[z]\}$ и $M_0 \cap Z_{1,\delta} = \emptyset$ удастся

свести задачу минимизации $\Omega[z]$ на $Z_{1,\delta}$ к классической задаче на условный экстремум, если $\Omega[z]$ — квазимонотонный функционал (такой, что для всякого $z_0 \in Z_1$, $z_0 \notin M_0$ в любой окрестности z_0 найдется $z_1 \in Z_1: \Omega[z_1] < \Omega[z_0]$). Для квазимонотонного функционала $\Omega[z]$ такого, что $M_0 \cap Z_{1,\delta} = \emptyset$, $\inf \Omega[z]$ достигается на $z_\delta: \rho_U(Az_\delta, u_\delta) = \delta$. Поэтому задача минимизации с ограничениями (неравенствами) переходит в задачу Лагранжа: $\inf \Omega[z]$ при условии $\rho_U(Az, u_\delta) = \delta$. Решая ее методом неопределенных множителей, получаем задачу безусловной минимизации сглаживающего функционала

$$M^\alpha[z, u_\delta] = \rho^2_U(Az, u_\delta) + \alpha \Omega[z], \quad (2)$$

параметр α определяется по невязке [60]

$$\rho_U(Az, u_\delta) = \delta. \quad (3)$$

Пусть не только правая часть (1), но и оператор A заданы неточно, т. е. вместо $\{A, u\}$ из (1) известно двухпараметрическое семейство приближенных исходных данных $\{A_h, u_\delta\}$:

$$\rho_U(\bar{u}, u_\delta) \leq \delta, \quad \rho_U(Az, A_h z) \leq h \{\Omega[z]\}^{1/2}, \quad z \in Z_1,$$

$$\Omega[z] \neq 0.$$

Для этого случая показано, что среди $z \in Z_1$, удовлет-

воряющих условию $\rho_U(A_h z, u_\delta) \leq (\delta + h(\Omega[z])^{1/2})$, нижняя грань $\Omega[z]$ на Z_1 достигается для z_1 , $\gamma = (h, \delta)$, такого, что

$$\rho_U(A_h, u_\delta) = \delta + h\{\Omega[z]\}^{1/2} \quad (4)$$

(в общем случае, если точное уравнение $Az = \bar{u}$ не имеет решения, то в (4) добавляется слагаемое $\mu = \inf_z \rho_U(A_h z, u_\delta)$ — мера несовместности [28]). Это позволяет свести задачу минимизации $\Omega[z]$ на множестве допустимых решений к задаче минимизации сглаживающего функционала

$$M^\alpha[z, u_\delta, A_h] = \rho^2_U(A_h z, u_\delta) + (\alpha - h^2)\Omega[z] + 2\delta h\{\Omega[z]\}^{1/2}, \quad (5)$$

причем параметр регуляризации α можно определить по обобщенному методу невязки [28; 83] из условия

$$\rho_U(A_h, z_1^\alpha, u_\delta) = h\{\Omega[z_1^\alpha]\}^{1/2} - \delta, \quad (6)$$

где z_1^α — решение (5). Можно также построить семейства PA исходя из обобщенного принципа невязки [29; 83].

2. Вернемся теперь от общих операторных уравнений к задаче (5) из гл. 6 МНК. В этом случае $\rho^2_U(Az, u_\delta) = \|X\bar{\theta} - \bar{y}\|^2$. В качестве $\Omega[z]$ естественно взять квадрат нормы уклонения решения от а priori заданного вектора $\bar{\theta}_0$: $\|\bar{\theta} - \bar{\theta}_0\|^2$. Тогда (2) принимает вид

$$\varphi(\bar{\theta}) = \|X\bar{\theta} - \bar{y}\|^2 + \alpha\|\bar{\theta} - \bar{\theta}_0\|^2, \quad (7)$$

причем α определяется по невязке

$$\|X\bar{\theta} - \bar{y}\|^2 = \delta^2. \quad (8)$$

Как видим, (7) и (9) из гл. 8 совпадают, если положить $\bar{\theta}_0 = 0$, $\lambda^2 = \alpha$. Однако если для (9) из гл. 8 есть большой произвол выбора λ , то условие (8) указывает конструктивный алгоритм определения α .

В случае неточно заданной матрицы данных в модели МНК, определяя погрешность через спектраль-

ную норму, введенную выше: $\|\bar{X} - X_h\| \leq h$, запишем вместо (5), (6) задачу минимизации

$$\|X_h \bar{\theta} - \bar{y}\|^2 + (\alpha - h^2) \|\bar{\theta} - \bar{\theta}_0\|^2 + 2\delta h \|\bar{\theta} - \bar{\theta}_0\| \quad (9)$$

с определением α из условия для решения (9) $\bar{\theta}^\alpha$

$$\|X_h \bar{\theta}^\alpha - \bar{y}\| - h \|\bar{\theta}^\alpha - \bar{\theta}_0\| = \delta. \quad (10)$$

В случае неточно заданной матрицы данных (9), (10) определяют регуляризованный метод наименьших квадратов (РМНК). Величину δ^2 в (8), (10) для равноточных измерений в модели (5) из гл. 6 можно оценить как $n\sigma^2$.

В [50] была отмечена связь РМНК со статистическими задачами конъюэнтного анализа [1; 3; 45] (метода ортогональных проекций [32; 53]). В задачах конъюэнтного анализа исследуются линейные функциональные зависимости вида (4) из гл. 6. Однако в отличие от МНК не только зависимая переменная, но и независимые известны с погрешностями [45]:

$$\left. \begin{aligned} \xi_i^{(l)} &= x_i^{(l)} + \delta_i^{(l)} \\ y_i &= \varphi_i + \varepsilon_i \end{aligned} \right\} \begin{aligned} l &= \overline{1, n}; \\ j &= \overline{1, k}, \end{aligned} \quad (11)$$

причем $\delta_i^{(l)}$ и ε_i статистически независимы. Задача исследователя — оценить коэффициенты $\theta_1, \dots, \theta_k$ представления $\varphi = \theta_1 x^{(1)} + \dots + \theta_k x^{(k)}$ по наблюдаемым величинам $\{\xi_i^{(l)}, y_i\}$; обычно предполагается нормальность ошибок $\varepsilon_i, \delta_i^{(l)}$.

Для решения этой задачи в конъюэнтном анализе предложены два метода. Первый из них в простейшем случае равноточных измерений $D\delta_i^{(l)} = \sigma_j^2, j = \overline{1, k}; D\varepsilon_i = \sigma_0^2$ приводит к минимизации выражения

$$\|X_h \bar{\theta} - \bar{y}\|^2 - n \sum_{j=1}^k \sigma_j^2 \theta_j^2; \quad (12)$$

в частном случае $\sigma_j^2 = \sigma^2 = \text{const}$ имеем

$$\|X_h \bar{\theta} - \bar{y}\|^2 - n\sigma^2 \|\bar{\theta}\|^2, \quad (12')$$

здесь $X_h = \{\xi_i^{(l)}\}$. Если теперь рассмотреть частный случай модели конъюэнтного анализа (11), когда

зависимая переменная y измеряется точно или с малой по сравнению с $\delta^{(l)}$ погрешностью ε , то естественно положить $\sigma_0^2=0$. Тогда в (9) погрешность правой части $\delta=0$, так что имеем вместо (9) задачу минимизации

$$\|X_h \vec{\theta} - \vec{y}\|^2 + (\alpha - h^2) \|\vec{\theta} - \vec{\theta}_0\|^2. \quad (9')$$

Поскольку спектральная норма матрицы равна ее наибольшему сингулярному значению [95; 123], нетрудно проверить, что в рамках модели (11) математическое ожидание квадрата спектральной нормы матрицы погрешностей $h^2 = E\{\|X_h - X\|^2\} = n\sigma^2$, так что при $\vec{\theta}_0 = \vec{0}$ (9') и (12') отличаются только в члене $\alpha \|\vec{\theta}\|^2$. Но последний, естественно, возникает в (12), (12'), если предполагать плохую обусловленность матрицы X_h . Таким образом, в статистической интерпретации РМНК может быть связан с моделью конфлюэнтного анализа (11), если в последней матрица X_h измерений независимых переменных плохо обусловлена.

При втором методе решения задач конфлюэнтного анализа применение метода максимального правдоподобия приводит к задаче на собственные значения для матрицы $X_h^T X_h$, что эквивалентно построению *SVD* (5) из гл. 8 для матрицы X_h . Параметры θ_i , $i = \overline{1, k}$, определяются как компоненты собственного вектора матрицы $X_h^T X_h$, отвечающие наименьшему сингулярному значению s_k . В случае плохой обусловленности X_h или сравнимости s_k с погрешностью задания X_h получаемые этим методом параметры θ_i характеризуются большими ошибками, так что и в этом случае необходима регуляризация в исходной задаче конфлюэнтного анализа.

3. Подход к построению регуляризованных методов наименьших квадратов, позволяющий строить не только устойчивые приближенные решения, но и реализующую их матрицу данных, предложен в [81; 82]. Для упрощения обозначений будем опускать «стрелку» над векторами.

Для (4) из гл. 6 рассмотрим приближенную систему вида

$$\tilde{X}\theta = \tilde{y}, \quad (13)$$

задаваемую расширенной матрицей $(X:\bar{y})$ и классом $(X:y)$ эквивалентных по точности систем. Пусть $\{X\}$ и $\{y\}$ — соответствующие эквивалентные классы. Тогда приближенная система $\Sigma = \{(X:y)\}$ задается как совокупность систем

$$X\theta = y, \quad X \in \{X\}, \quad y \in \{y\}, \quad (X:y) \in \Sigma.$$

Так, для евклидовых норм и точностей μ и δ имеем

$$\Sigma = \Sigma_{\delta\mu} : \{X\} = \{X : \|X - \tilde{X}\| \leq \mu\},$$

$$\{y\} = \{y : \|y - \tilde{y}\| \leq \delta\},$$

$$\|X\|^2 = \sum_{i=1}^n \sum_{j=1}^k X_{ij}^2, \quad \|y\|^2 = \sum_{i=1}^n y_i^2.$$

При этом не предполагается разрешимость $(X:\bar{y})$.

Назовем приближенную систему Σ -совместной, если найдется $(X:y) \in \Sigma$ такая, что решение задачи $X\theta = y$ существует; соответствующий k -вектор назовем допустимым решением Σ .

Пусть теперь $\Theta_\Sigma = \{\theta\}$ — совокупность всех допустимых решений Σ . Все элементы Θ_Σ — эквивалентные по точности решения. В важном частном случае $\Sigma = \Sigma_{\delta\mu}$ обозначим $\Theta_{\delta\mu} = \Theta_\Sigma$.

Функционал $\Omega(\theta) = \|\theta - \theta_0\|$ назовем функционалом сложности (здесь $\theta_0 \in E_k$ — заданный в постановке задачи элемент). Без ограничения общности можно положить $\theta_0 = 0$.

Под нормальным решением системы Σ , как и выше, будем понимать решение с минимальной нормой

$$\bar{\theta}_\Sigma = \text{Arg inf } \|\theta\|, \quad \theta \in \Theta_\Sigma.$$

Тогда для системы $\Sigma_{\delta\mu}$ нормальное решение определяется как решение задачи N_0 .

Задача N_0 . Найти $\bar{\theta}_{\delta\mu} = \text{Arg inf } \|\theta\|$, $\theta \in \Theta_{\delta\mu} = \{\theta : \|y - X\theta\| = 0, (X:y) \in \Sigma_{\delta\mu}\}$.

Нетрудно видеть, что справедлива

Лемма 1. Необходимыми и достаточными условиями существования допустимого решения θ задачи $X\theta = y$, $(X:y) \in \Sigma_{\delta\mu}$ являются $\|\tilde{y} - X\theta\| \leq \delta$, $X \in \{X : \|X - \tilde{X}\| \leq \mu\}$.

Из леммы (1) следует, что задаче N_0 нахождения нормального решения $\bar{\theta}_{\delta, \mu}$ эквивалентна

Задача N_0' . Найти $\bar{\theta}_{\delta, \mu}$ такое, что $\|\bar{\theta}_{\delta, \mu}\| = \eta(\delta, \mu) = \inf \|\theta\|$ на $\Theta_{\delta, \mu} = \{\theta: \|\bar{y} - X\theta\| \leq \delta, \|\bar{X} - X\| \leq \mu\}$.

Величину $\eta(\delta, \mu)$ назовем *порогом сложности* задач N_0, N_0' , поскольку не существует допустимых решений θ для $\Sigma_{\delta, \mu}$ при $\|\theta\| < \eta(\delta, \mu)$.

Класс $\Sigma = \{(X; y)\}$ называется *компактным*, если $\{X\}$ и $\{y\}$ компактны в своих метриках. Нетрудно видеть, что $\Sigma_{\delta, \mu}$ компактна в смысле указанного определения.

Теорема 1. Для всякой совместной и компактной системы Σ существуют нормальное решение $\bar{\theta}_\Sigma$ и система $(\bar{X}; \bar{y})$, его реализующая:

$$\|\bar{y} - \bar{X}\bar{\theta}_\Sigma\| = 0, (\bar{X}; \bar{y}) \in \Sigma.$$

Основная лемма. Система уравнений относительно $X: X\theta = y$ при заданных $\theta \in E_k, \|\theta\| \neq 0, y \in E_n$ разрешима относительно X . Решение \bar{X} этой системы, минимальное по норме, единственно и задается формулой

$$\bar{X} = y\theta^T / \|\theta\|^2 \quad (\text{или } X_{ij} = y_i \theta_j / \|\theta\|^2) \quad (14)$$

$$\|\bar{X}\| = \|y\| / \|\theta\|. \quad (15)$$

Доказательство. Подставив в $X\theta = y$ (14) и записав $\|\theta\|^2$ как $\theta^T \theta$, видим, что \bar{X} удовлетворяет системе. Пусть $\bar{\bar{X}}$ — другое решение. Поскольку нормы y и X согласованы, то $\|y\| \leq \|\bar{\bar{X}}\| \|\theta\|$, т. е. $\|\bar{\bar{X}}\| \geq \|\bar{X}\|$. Осталось доказать единственность решения. Пусть $\bar{\bar{X}}_1$ — такое решение, что $\|\bar{\bar{X}}_1\| = \|\bar{X}\|$ и $\bar{\bar{X}}_1 \neq \bar{X}$. Рассмотрим $\bar{\bar{X}} = (\bar{\bar{X}}_1 + \bar{X})/2$. Очевидно, что $\bar{\bar{X}}$ — также решение. Поскольку $\|\bar{X}\| < \|\bar{\bar{X}}_1\| = \|\bar{X}\|$ в силу неравенства Коши—Буняковского, то получаем противоречие с доказанной минимальностью $\|\bar{X}\|$. Значит, наше допущение неверно и $\bar{\bar{X}}_1 = \bar{X}$.

Лемма 2. Пусть заданы $0 \in E_k, \|\theta\| \neq 0, \rho \in E_n, \|\rho\| \neq 0$, и класс матриц $\Delta_\mu = \{\Delta: \|\Delta\| \leq \mu\}$. Тогда существует и единственна $\bar{\Delta} \in \Delta$ такая, что

$$\bar{\Delta}\theta = \rho \quad (16)$$

при максимальном значении α . $\bar{\Delta}$ и α определяются соотношениями

$$\bar{\Delta} = \mu \frac{\rho \theta^T}{\|\rho\| \|\theta\|}, \quad \|\bar{\Delta}\| = \mu, \quad \alpha = \mu \frac{\|\theta\|}{\|\rho\|}. \quad (17)$$

Доказательство. Выполнение (16) для $\bar{\Delta}$ и α , определенных в (17), проверяется непосредственной подстановкой; также легко вычисляется $\|\bar{\Delta}\| = \mu$.

Пусть $\alpha > 0$, $\Delta' = \bar{\Delta}/\alpha$, $\bar{\Delta} \in \Delta_\mu$. Задача (16) для Δ' запишется как $\Delta' \theta = \rho$. Матрица Δ' , являющаяся решением с минимальной нормой, согласно основной лемме существует, единственна и равна

$$\bar{\Delta}/\alpha = \Delta' = \rho \theta^T / \|\theta\|^2, \quad \|\Delta'\| = \|\rho\| / \|\theta\| = \|\bar{\Delta}\| / \alpha \leq \mu / \alpha.$$

Отсюда имеем $\alpha \leq \mu \|\theta\| / \|\rho\|$. Поскольку, как показано выше, для $\bar{\Delta}$ и α , определенных в (17), выполняется (16), то максимально возможное $\alpha = \mu \|\theta\| / \|\rho\|$. Подставляя это α в Δ' , находим, что отвечающая ему матрица $\bar{\Delta}$ единственна и определяется (17).

Докажем теперь утверждение, аналогичное (4): если $(X: \bar{y}) \in \Sigma_{\delta\mu}$ реализует нормальное решение $\bar{\theta}_{\delta\mu}$, $\|\bar{\theta}_{\delta\mu}\| > 0$, $\|\bar{y} - X \bar{\theta}_{\delta\mu}\| = 0$, то \bar{y} и X лежат на границе $\{y\}$ и $\{X\}$:

$$\|\bar{y} - \tilde{y}\| = \delta, \quad \|X - \tilde{X}\| = \mu.$$

Лемма 3. Пусть $\hat{\theta} \neq 0$ принадлежит E_k и X лежит внутри $\{X\}$: $\|X - \tilde{X}\| \leq \mu_1 < \mu$. Тогда существуют X^* и θ^* такие, что $X \hat{\theta} = X^* \theta^*$, $\|\theta^*\| < \|\hat{\theta}\|$, $\|X^* - X\| < \mu$.

Доказательство. Определим $\theta^* = (1 - \zeta) \hat{\theta}$ и $X^* = X + \Delta$. Пусть $X^* \theta^* = (1 - \zeta) (X \hat{\theta} + \Delta \hat{\theta}) = X \hat{\theta}$. Тогда матрица Δ определяется из соотношения

$$\Delta \hat{\theta} = \frac{\zeta}{1 - \zeta} X \hat{\theta} = \frac{\zeta}{1 - \zeta} \hat{y}, \quad \hat{y} = X \hat{\theta}.$$

Найдя согласно основной лемме

$$\Delta = \frac{\zeta}{1 - \zeta} \frac{\hat{y} \hat{\theta}^T}{\|\hat{\theta}\|^2},$$

имеем при достаточно малом ξ $\|\Delta\| \leq \mu - \mu_1$, так что $\|X^* - \bar{X}\| < \mu$. Если же $X\bar{\theta} = 0$, то можно взять $\theta^* = 0$.

Следствие. Если $\|\bar{y} - \bar{X}\bar{\theta}\| = 0$, $\|\bar{\theta}\| \neq 0$ и $\|\bar{X} - \bar{X}\| = \mu_1 < \mu$, то существуют X^* и θ^* такие, что $\|\bar{y} - X^*\theta^*\| = 0$, $\|\theta^*\| < \|\bar{\theta}\|$, $\|X^* - \bar{X}\| < \mu$.

В дальнейшем предполагается, что нормальное решение $\bar{\theta}_{\delta, \mu} \neq 0$, что достигается при $\|\bar{y}\| > \delta$.

Теорема 2. Если $(\bar{X}; \bar{y})$ реализует нормальное решение $\bar{\theta}_{\delta, \mu}$ системы $\Sigma_{\delta, \mu}$, то $\|\bar{X} - \bar{X}\| = \mu$.

Доказательство методом «от противного» следует из леммы 3.

Теорема 3. Если $(\bar{X}; \bar{y})$ реализует $\bar{\theta}_{\delta, \mu}$, то $\|\bar{y} - \bar{X}\bar{\theta}_{\delta, \mu}\| = \delta$.

Идея доказательства такая же, как и для теоремы 2. В самом деле, если $\|\bar{y} - \bar{X}\bar{\theta}_{\delta, \mu}\| = \delta_1 < \delta$, то существует ε -окрестность $U_\varepsilon(\bar{\theta}_{\delta, \mu})$ такая, что для всякой точки $\theta \in U_\varepsilon(\bar{\theta}_{\delta, \mu})$ значение $\|\bar{y} - \bar{X}\theta\| < \delta$, т. е. $\theta \in \Theta_{\delta, \mu}$. Точка $\theta_\xi^* = (1 - \xi)\bar{\theta}_{\delta, \mu} \in U_\varepsilon(\bar{\theta}_{\delta, \mu})$ при достаточно малом $\xi > 0$ и $\|\theta_\xi^*\| < \|\bar{\theta}_{\delta, \mu}\|$, что доказывает неверность исходного предположения.

Таким образом, задаче N_0 (или N_0') эквивалентна

Задача N_1 . Найти $\bar{\theta}_{\delta, \mu}$ такое, что $\|\bar{\theta}_{\delta, \mu}\| = \eta(\delta, \mu) = \inf \|\theta\|$ на $\Theta_{\delta, \mu}^1 = \{\theta : \|\bar{y} - \bar{X}\theta\| = \delta, \|\bar{X} - \bar{X}\| = \mu\}$.

Рассмотрим вариационную задачу: при заданных \bar{y} и $\bar{\theta}$ найти матрицу X , реализующую $\inf \|\bar{y} - X\bar{\theta}\|$ на $\{X : \|X - \bar{X}\| \leq \mu\}$. Обозначим $\rho(\theta) = \bar{y} - X\theta$, $\bar{\rho}(\theta) = \bar{y} - \bar{X}\theta$, откуда $\inf \|\bar{y} - X\theta\| = \inf \|\rho(\theta)\|$ и $\rho(\theta) = \bar{\rho}(\theta) - \Delta\theta$, где $\Delta = X - \bar{X}$. Введем также $v(\theta) = \|\rho(\theta)\| - \mu\|\theta\| = \|\bar{y} - X\theta\| - \mu\|\theta\|$ и $\bar{v}(\theta) = \|\bar{\rho}(\theta)\| - \mu\|\theta\| = \|\bar{y} - \bar{X}\theta\| - \mu\|\theta\|$.

Теорема 4. Если $\bar{v}(\theta) \leq 0$ или $\|\bar{\rho}(\theta)\| = \beta\mu\|\theta\|$, $0 < \beta \leq 1$, то разрешимо уравнение

$$\bar{y} - X\theta = 0, \quad X \in \{X\}, \quad (18)$$

причем X , наименее по норме уклоняющееся от \bar{X} , есть $X = \bar{X} + \hat{\Delta}$, где

$$\hat{\Delta} = \frac{\bar{\rho}(\theta)\theta^T}{\|\theta\|^2} = \beta\mu \frac{\bar{\rho}(\theta)\theta^T}{\|\bar{\rho}(\theta)\|\|\theta\|}, \quad \|\hat{\Delta}\| = \beta\mu. \quad (19)$$

Доказательство. Уравнение (18) равносильно соотношению $\tilde{\rho}(\theta) - \Delta\theta = 0$, и в силу основной леммы его минимальное по норме решение

$$\hat{\Delta} = \frac{\tilde{\rho}(\theta)\theta^T}{\|\theta\|^2}, \quad \|\hat{\Delta}\| = \frac{\|\tilde{\rho}(\theta)\|}{\|\theta\|} = \beta\mu.$$

Если же $\tilde{\rho}(\theta) = 0$, то $\tilde{y} - X\theta = 0$ и тогда $X = X$, $\|\hat{\Delta}\| = 0$.

Следствие. Нормальное решение $\bar{\theta}_{\delta\mu}$ системы $\Sigma_{\delta\mu}$ удовлетворяет условию $\tilde{v}(\bar{\theta}_{\delta\mu}) > 0$.

Это показывается от противного: если бы $\tilde{v}(\bar{\theta}_{\delta\mu}) \leq 0$, то $\bar{\theta}_{\delta\mu}$ удовлетворяло бы уравнению $X\bar{\theta}_{\delta\mu} = \tilde{y}$, $X \in \{X\}$, но по теореме 3 для нормального решения $\|\tilde{y} - X\bar{\theta}_{\delta\mu}\| = \delta > 0$.

Теорема 5. Если $\tilde{v}(\theta) > 0$, т. е. $\|\tilde{\rho}(\theta)\| > \mu\|\theta\|$, то $\inf \|\tilde{y} - X\theta\|$ на $\{X\}$ реализуется для единственной матрицы $X = X + \hat{\Delta}$, где

$$\hat{\Delta} = \mu \frac{\tilde{\rho}(\theta)\theta^T}{\|\tilde{\rho}(\theta)\|\|\theta\|}, \quad \|\hat{\Delta}\| = \mu,$$

причем $\|\hat{\rho}(\theta)\| = \|\tilde{\rho}(\theta)\| - \mu\|\theta\|$ (при этом $\hat{\rho}(\theta) = \tilde{y} - X\theta = \tilde{\rho}(\theta) - \Delta\theta = \tilde{\rho}(\theta)[1 - \mu\|\theta\|/\|\tilde{\rho}(\theta)\|]$, $0 < (1 - \mu\|\theta\|/\|\tilde{\rho}(\theta)\|) < 1$ и коллинеарен $\hat{\Delta}\theta$).

Справедливость теоремы 5 вытекает из приведенных ниже лемм 4-6.

Лемма 4. Если $X' = \text{Arg} \inf \|\tilde{y} - X\theta\|$ на $\{X\}$, то $\rho'(\theta)$ коллинеарен $\tilde{\rho}(\theta)$.

Доказательство. Пусть $\rho'(\theta) = \tilde{\rho}(\theta) - \Delta'\theta$ не коллинеарен $\tilde{\rho}(\theta)$. Тогда $\|\rho'(\theta)\| > \|\tilde{\rho}(\theta)\| - \|\Delta'\theta\| \geq \|\tilde{\rho}(\theta)\| - \mu\|\theta\| = \|\hat{\rho}(\theta)\|$, $\|\Delta'\theta\| \leq \mu\|\theta\|$. Это доказывает, что X' не реализует на $\{X\}$ $\inf \|\tilde{\rho}(\theta)\|$, что противоречит условию. Следовательно, предположение о неколлинеарности неверно, и лемма доказана.

Лемма 5. Матрица $X = X + \hat{\Delta}$ реализует $\inf \|\tilde{y} - X\theta\|$ на $\{X\}$.

Доказательство. Пусть $X' = X + \Delta'$ — какая-либо матрица, реализующая $\inf \|\tilde{\rho}(\theta)\|$ на $\{X\}$ (существование таких матриц следует из компактности $\{X\}$). Тогда $\rho'(\theta)$ коллинеарен $\tilde{\rho}(\theta)$, а также $\Delta'\theta$.

поскольку $\Delta'\theta = \tilde{\rho}(\theta) - \rho'(\theta)$, $\|\rho'(\theta)\| = \|\tilde{\rho}(\theta)\| - \|\Delta'\theta\| \geq \|\tilde{\rho}(\theta)\| - \mu\|\theta\| = \|\tilde{\rho}(\theta)\|$, $\|\Delta'\theta\| \leq \mu\|\theta\| \leq \|\tilde{\rho}(\theta)\|$.

С другой стороны, X' реализует $\inf \|\rho(\theta)\|$ на $\{X\}$, т. е. $\|\rho'(\theta)\| \leq \|\tilde{\rho}(\theta)\|$, так что $\|\rho'(\theta)\| = \|\tilde{\rho}(\theta)\|$. Таким образом, $X = X' + \Delta$ реализует $\inf \|\rho(\theta)\|$ на $\{X\}$.

Лемма 6. Существует и единственная матрица

$$X = \text{Arg} \inf \|\rho(\theta)\| \text{ на } \{X\}.$$

Доказательство. Заметим сначала, что если $X'' = X' - \Delta''$ реализует $\inf \|\rho(\theta)\|$ на $\{X\}$, то $\|\Delta''\| = \mu$. По лемме 4 из $\|\rho''(\theta)\| = \|\tilde{\rho}(\theta)\|$ следуют равенства $\|\Delta''\theta\| = \|\widehat{\Delta}\theta\|$ и $\|\Delta''\theta\| = \mu\|\theta\|$, так как по определению $\widehat{\Delta} \|\widehat{\Delta}\theta\| = \mu\|\theta\|$, откуда $\|\Delta''\| \geq \mu$. Однако $X'' \in \{X\}$, поэтому $\|\Delta''\| \leq \mu$. Значит, имеет место равенство $\|\Delta''\| = \mu$.

Пусть существует $X'' \neq X$ такая, что $X'' = \text{Arg} \inf \|\rho(\theta)\|$. Тогда и $X' = 0,5(X'' + X)$ реализует $\inf \|\rho(\theta)\|$ на $\{X\}$. В самом деле, $\|\rho'(\theta)\| = \|\tilde{y} - X'\theta\| = 0,5\|(\tilde{y} - X''\theta) + (\tilde{y} - X\theta)\| = 0,5\|\rho''(\theta) + \tilde{\rho}(\theta)\|$. Векторы $\rho''(\theta)$ и $\tilde{\rho}(\theta)$ коллинеарны по лемме 4 и $\|\rho''(\theta)\| = \|\tilde{\rho}(\theta)\|$. Отсюда $\rho''(\theta) = \tilde{\rho}(\theta)$ (если бы $\rho''(\theta) = -\tilde{\rho}(\theta)$, то $\rho'(\theta) = 0$). Итак, $\|\rho'(\theta)\| = \|\tilde{y} - X'\theta\| = \|\rho''(\theta)\| = \|\tilde{\rho}(\theta)\|$ и в силу замечания в начале доказательства леммы $\|\Delta'\| = \|\widehat{\Delta}\| = \|\Delta''\|$. Однако если $\|\Delta'\| = \|\Delta''\| = \mu$ и $\Delta' \neq \Delta''$, то $\|\Delta'\| \leq \|\widehat{\Delta}\| = \|\Delta''\| = \mu$, что противоречит установленному выше равенству. Значит, $X = X''$.

Можно доказать также лемму 6, опираясь на леммы 4 и 2. Запишем $\tilde{y} - X\theta = \tilde{\rho}(\theta) - \Delta\theta$. Пусть \tilde{y} и θ фиксированы. По лемме 4 $\Delta\theta$ коллинеарен $\rho(\theta)$, так что поиск $X = \text{Arg} \inf \|\rho(\theta)\|$ на $\{X\}$ эквивалентен построению Δ , реализующего $\Delta\theta = \alpha\rho(\theta)$. Поскольку $\|\tilde{\rho}(\theta)\| > \mu\|\theta\|$, то $|\alpha| \leq \mu \frac{\|\theta\|}{\|\tilde{\rho}(\theta)\|} < 1$. Поэтому $\|\tilde{y} - X\theta\| = \|\tilde{\rho}(\theta)\| (1 - \alpha)$ минимальна при максимальном α . Это максимальное α и соответствующее ему Δ даются формулами (17) леммы 2, причем они определены единственным образом:

$$\widehat{\Delta} = \mu \frac{\tilde{\rho}(\theta)\theta^T}{\|\tilde{\rho}(\theta)\|\|\theta\|}, \quad \alpha = \mu \frac{\|\theta\|}{\|\tilde{\rho}(\theta)\|}, \quad \|\widehat{\Delta}\| = \mu.$$

Это доказывает лемму 6, а вместе с тем и завершает доказательство теоремы 5.

Теорема 5 указывает способ построения матрицы X , реализующей $\inf \|\tilde{y} - X\theta\|$. Теперь займемся нахождением нормального решения.

Задача N_2 . Найти $\bar{\theta}_2$, такой, что $\|\bar{\theta}_2\| = \eta(\delta, \mu) = \inf \|\theta\|$ на $\Theta_{\delta, \mu}^2 = \{\theta : \|\tilde{y} - X\theta\| - \mu \|\theta\| = \delta\}$.

Лемма 7. Если задан вектор θ такой, что $\|\tilde{\rho}(\theta)\| - \mu \|\theta\| = \delta$, то $\theta \in \Theta_{\delta, \mu}^1$.

Доказательство. Согласно теореме 5, взяв $X = \hat{X} + \hat{\Delta}$, где $\hat{\Delta} = \mu \frac{\tilde{\rho}(\theta)\theta^T}{\|\tilde{\rho}(\theta)\| \|\theta\|}$, $X \in \{X\}$, $\|\hat{\Delta}\theta\| = \mu \|\theta\|$, имеем $\|\tilde{y} - X\theta\| = \|\tilde{\rho}(\theta)\| - \mu \|\theta\| = \delta$, $\|X - \hat{X}\| = \mu$, т. е. $\theta \in \Theta_{\delta, \mu}^1$.

Следствие. Всякий вектор $\bar{\theta}_2$, являющийся решением задачи N_2 , удовлетворяет также условию $\bar{\theta}_2 \in \Theta_{\delta, \mu}^1$, и если $\bar{\theta}_1$ — решение задачи N_1 , то $\|\bar{\theta}_1\| \leq \|\bar{\theta}_2\|$.

Лемма 8. Если $\bar{\theta}_1$ — нормальное решение задачи N_1 и \bar{X} — матрица, его реализующая, то она совпадает с $X = \text{Arg} \inf \|\tilde{y} - X\bar{\theta}_1\|$ (на $\{X\}$):

$$X = \bar{X} \quad (20)$$

и

$$\|\tilde{\rho}(\bar{\theta}_1)\| = \|\tilde{y} - \bar{X}\bar{\theta}_1\| = \|\tilde{y} - \bar{X}\bar{\theta}_1\| - \mu \|\bar{\theta}_1\| = \delta, \quad (21)$$

т. е. $\bar{\theta}_1 \in \Theta_{\delta, \mu}^2$.

Доказательство. По определению $\|\tilde{\rho}(\bar{\theta}_1)\| = \|\tilde{y} - \bar{X}\bar{\theta}_1\| \leq \|\tilde{y} - X\bar{\theta}_1\| = \|\tilde{\rho}(\bar{\theta}_1)\| = \delta$. Пусть $\|\tilde{\rho}(\bar{\theta}_1)\| = \|\tilde{\rho}(\bar{\theta}_1)\| - \mu \|\bar{\theta}_1\| = \delta_1 < \delta$, т. е. $\bar{\theta}_1 \in \Theta_{\delta_1, \mu}^2$. Для решения задачи N_2 при $\delta = \delta_1$, равного $\bar{\theta}_2(\delta_1)$, имеем $\|\bar{\theta}_2(\delta_1)\| = \inf_{\theta \in \Theta_{\delta_1, \mu}^2} \|\theta\| \leq \|\bar{\theta}_1\|$. Однако $\|\bar{\theta}_2(\delta)\| \leq \|\bar{\theta}_2(\delta_1)\|$, поскольку

$\delta > \delta_1$ (см. [80]). Итак, $\|\bar{\theta}_1\| \geq \|\bar{\theta}_2(\delta_1)\| > \|\bar{\theta}_2(\delta)\|$, что противоречит лемме 7. Если же $\|\tilde{y} - \bar{X}\bar{\theta}_1\| = \|\tilde{y} - X\bar{\theta}_1\| = \delta$, то из единственности решения X вариационной задачи следует справедливость (20) и (21).

Следствие. Всякое решение $\bar{\theta}_1$ задачи N_1 удовлетворяет условию $\|\bar{\theta}_2\| \leq \|\bar{\theta}_1\|$.

Теорема 6. Множества решений задач N_1 и N_2 совпадают.

Справедливость утверждения вытекает из лемм 7, 8 и следствий к ним.

Обратимся к доказательству единственности нормального решения системы $\Sigma_{\delta\mu}$ (задача N_0).

Как было показано выше, решения задач N_0 (или N_0'), N_1 и N_2 совпадают. Обозначим через $[N_2]$ задачу, сопряженную с N_2 .

Задача $[N_2]$. Найти $\bar{\theta}$, для которого

$$\bar{v}(\bar{\theta}) = \|\bar{y} - X\bar{\theta}\| - \mu\|\bar{\theta}\| = \delta(\eta, \mu) = \inf_{\theta} \bar{v}(\theta)$$

на $\Theta_\eta = \{\theta: \|\theta\| = \eta\}$.

Лемма 9. Множества решений задач N_2 и $[N_2]$ совпадают при $\eta = \eta(\delta, \mu)$.

Напомним, что задачи, множества решений которых совпадают, называют *взаимными*. В работе [80] рассмотрены две сопряженные экстремальные задачи с ограничениями типа неравенств. Пусть в пространстве R заданы непрерывные функции $\psi(z)$ и $\omega(z)$. Тогда имеем две сопряженные задачи.

Задача M . Найти на множестве $Z_\delta = \{z: \psi(z) \leq \delta\}$ элемент \bar{z}_δ такой, что $\omega(\bar{z}_\delta) = \inf \omega(z)$ (на Z_δ); обозначим $\omega(\bar{z}_\delta) = \bar{\eta}(\delta)$.

Задача $[M]$. Найти на множестве $\Omega_\eta = \{z: \omega(z) \leq \eta\}$ элемент \bar{z}_η такой, что $\psi(\bar{z}_\eta) = \inf \psi(z)$ (на Ω_η); обозначим $\psi(\bar{z}_\eta) = \bar{\delta}(\eta)$.

В [80] сформулированы условия, при которых задачи M и $[M]$ являются взаимными. В частности, показано, что они являются таковыми, если R — конечномерное евклидово пространство, $\omega(z) = \|z\|$, а $\psi(z)$ — произвольная непрерывная функция; при этом в задаче $[M]$ $\eta = \bar{\eta}(\delta)$. Поэтому утверждение леммы 9 следует из принципа взаимности.

Теорема единственности. Если приближенная система $\Sigma_{\delta\mu}$ совместна, то существует и единственно нормальное решение $\bar{\theta}_{\delta\mu}$ этой системы (см. задачу N_0).

Доказательство. Всякое нормальное решение системы $\Sigma_{\delta\mu}$ является решением задач N_0 , N_1 , N_2 и $[N_2]$. Задача $[N_2]$ — задача отыскания $\bar{\theta}: \|\bar{y} - X\bar{\theta}\| - \mu\|\bar{\theta}\| \leq \|\bar{y} - X\theta\| - \mu\|\theta\|$ при $\|\theta\| = \eta(\delta, \mu)$. Отсюда $\bar{\theta}$ является решением также следующей задачи для индивидуально заданной системы $(X: \bar{y})$.

Задача $[n_0]$. Найти $\bar{\theta}$, для которого $\|\bar{y} - X\bar{\theta}\| = \inf \|\bar{y} - X\theta\|$ при $\|\theta\| = \eta(\delta, \mu)$. Эта задача эквивалентна следующей взаимной задаче о нахождении нормального решения $\bar{\theta}$.

Задача n_0 . Найти $\bar{\theta}$, для которого $\|\bar{\theta}\| = \inf \|\theta\|$ при условии $\|\bar{y} - X\theta\| = \delta + \mu\eta(\delta, \mu) = \delta_1$.

Единственность такого решения для линейного оператора хорошо известна, так как если $\theta' \neq \theta''$ — решения задачи n_0 , то $\|\theta'\| = \|\theta''\|$, и для $\theta' \neq \theta''$ и $\bar{\theta} = 0,5(\theta' + \theta'')$ имеем $\|\bar{\theta}\| < \|\theta'\| = \|\theta''\| = \eta(\delta)$ и $\|\bar{\rho}(\bar{\theta})\| = \|\bar{y} - X\bar{\theta}\| < 0,5(\|\bar{y} - X\theta'\| + \|\bar{y} - X\theta''\|) = \delta_1$, что противоречит определению $\eta(\delta, \mu)$.

Задача N_2 позволила найти нормальное решение $\bar{\theta}_{0\mu}$ задачи $\Sigma_{0\mu}$ и доказать его единственность. Покажем, что с помощью $\bar{\theta}_2$ можно явно выписать $(X:\bar{y})$, реализующую $\bar{\theta}_{0\mu}$.

В лемме 8 показано, что \bar{X} , реализующая решение $\bar{\theta}_1$ задачи N_1 , совпадающее с нормальным решением $\bar{\theta}_{0\mu}$, тождественна \bar{X} , т. е.

$$\bar{X} = \bar{X} + \bar{\Delta} = \bar{X} = \bar{X} + \bar{\Delta}, \quad \bar{\Delta} = \mu \frac{\tilde{\rho}(\bar{\theta}_2) \bar{\theta}_2^T}{\|\tilde{\rho}(\bar{\theta}_2)\| \|\bar{\theta}_2\|}, \quad \|\bar{\Delta}\| = \mu.$$

Для \bar{y} получаем $\bar{y} = X\bar{\theta}_2 = \bar{y} - y'$, где

$$y' = \bar{y} - \bar{X} \bar{\theta}_2 = \bar{y} - \bar{X} \bar{\theta}_2 - \mu \frac{\tilde{\rho}(\bar{\theta}_2)}{\|\tilde{\rho}(\bar{\theta}_2)\|} \|\bar{\theta}_2\|,$$

или

$$y' = \tilde{\rho}(\bar{\theta}_2) \left(1 - \mu \frac{\|\bar{\theta}_2\|}{\|\tilde{\rho}(\bar{\theta}_2)\|} \right),$$

так что $\|\bar{y} - \bar{y}\| = \|\tilde{\rho}(\bar{\theta}_2)\| - \mu \|\bar{\theta}_2\| = \delta$, откуда следует

Теорема 7. Индивидуальная система $(\bar{X}:\bar{y})$, реализующая нормальное решение $\bar{\theta}_{0\mu}$ приближенной системы $\Sigma_{0\mu}$, определяется через $\bar{\theta}_2$ — решение задачи N_2 — с помощью формул

$$\bar{X} = \bar{X} + \bar{\Delta}, \quad \bar{\Delta} = \mu \frac{\tilde{\rho}(\bar{\theta}_2) \bar{\theta}_2^T}{\|\tilde{\rho}(\bar{\theta}_2)\| \|\bar{\theta}_2\|}, \quad \|\bar{\Delta}\| = \mu, \quad (22)$$

$$\bar{y} = \bar{y} - y', \quad y' = \tilde{\rho}(\bar{\theta}_2) \left(1 - \mu \frac{\|\bar{\theta}_2\|}{\|\tilde{\rho}(\bar{\theta}_2)\|} \right), \quad \|\bar{y} - \bar{y}\| = \delta.$$

Пусть дана СЛАУ

$$\begin{aligned} \bar{X}\theta = \bar{y}, \quad \bar{X} = \{\bar{X}_{ij}\}, \quad \bar{y} = \{y_i\}, \quad \theta = \{\theta_j\}, \\ i = \overline{1, n}, \quad j = \overline{1, k}. \end{aligned} \quad (23)$$

Считаем, что система разрешима, т. е. $\bar{\Theta} = \{\theta : \bar{X}\theta = \bar{y}\} \neq \emptyset$. Нормальным решением (23) назовем $\bar{\theta} \in \bar{\Theta} : \|\bar{\theta}\| \leq \|\theta\|$ на $\bar{\Theta}$. Очевидна единственность нормального решения для всякой разрешимой системы.

Пусть $\Sigma_{\delta\mu}$ дает приближенную информацию относительно (23) с точностью δ, μ :

$$\|\bar{X} - \tilde{X}\| \leq \mu, \quad \|\bar{y} - \tilde{y}\| \leq \delta, \quad \|X - \tilde{X}\| \leq \mu, \quad \|y - \tilde{y}\| \leq \delta. \quad (24)$$

Тогда $\bar{\theta}_{\delta\mu}$ — нормальное решение $\Sigma_{\delta\mu}$ — дает устойчивое приближение к $\bar{\theta}$ — нормальному решению (23).

Теорема 8. Для любой степени точности $\varepsilon > 0$ существуют $\delta(\varepsilon) > 0$ и $\mu(\varepsilon) > 0$ такие, что $\bar{\theta}_{\delta\mu}$ отклоняется от $\bar{\theta}$ не более чем на $\varepsilon : \|\bar{\theta}_{\delta\mu} - \bar{\theta}\| \leq \varepsilon$ при $0 < \delta < \delta(\varepsilon), 0 < \mu < \mu(\varepsilon)$.

Доказательство. Очевидно, что $\bar{\theta} \in \bar{\Theta}_{\delta\mu}$ при любых δ и μ , следовательно, $\|\bar{\theta}_{\delta\mu}\| \leq \|\bar{\theta}\|$ при всех δ, μ . Отсюда вытекает возможность выбора подпоследовательностей $\delta_n \rightarrow 0, \mu_n \rightarrow 0$, таких, что $\bar{\theta}_{\delta_n \mu_n} \rightarrow \bar{\theta}$, а также

$\tilde{X}_{\mu_n} \rightarrow \bar{X}, \tilde{y}_{\delta_n} \rightarrow \bar{y}$. Отсюда $\|\bar{\theta}\| \leq \|\bar{\theta}\|$ и $\bar{X}\bar{\theta} = \bar{y}$.

В силу определения нормального решения задачи (23) получаем, что $\|\bar{\theta}\| = \|\bar{\theta}\|$, а в силу единственности определения $\bar{\theta}$ имеем $\bar{\theta} = \bar{\theta}$, откуда следует справедливость теоремы 8.

Обратимся к рассмотрению метода определения $\bar{\theta}_{\delta\mu}$, вытекающего из приведенных результатов. Было установлено, что решение $\bar{\theta}_{\eta\mu}$ задачи $[N_2] : \tilde{v}(\bar{\theta}_{\eta\mu}) = \delta(\eta, \mu) = \inf(\|\tilde{y} - \tilde{X}\theta\| - \mu\|\theta\|)$ на $\Theta_\eta = \{\theta : \|\theta\| = \eta\}$ и решение $\bar{\theta}_\eta$ задачи $[n_0] : \|\tilde{y} - \tilde{X}\bar{\theta}_\eta\| = \delta(\eta) = \inf\|\tilde{y} - \tilde{X}\theta\|$ на $\Theta_\eta = \{\theta : \|\theta\| = \eta\}$ совпадают для всех η и $\mu : \bar{\theta}_{\eta\mu} = \bar{\theta}_\eta$ и $\delta(\eta, \mu) = \delta(\eta) - \mu\eta$. Находя корень уравнения $\delta(\eta') - \mu\eta' = \delta$, получаем $\theta_{\delta\mu} = \theta_{\delta(\eta')}$, где $\delta(\eta') = \delta' + \mu\eta'$. Эти рассуждения иллюстрирует рис. 3.

Таким образом, задача N_0 (или N_2) свелась к нахождению корня уравнения $\delta(\eta') - \mu\eta' = \delta'$, в которое входит $\delta(\eta)$. Эта функция встречается при решении:

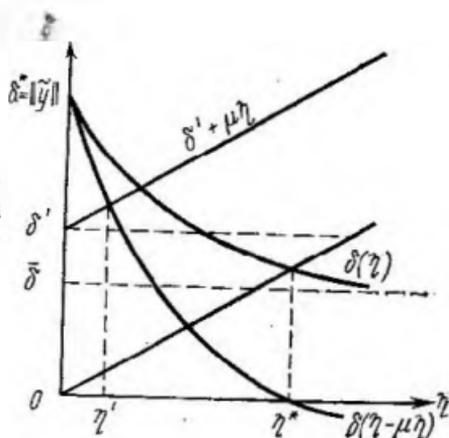


Рис. 3. δ — мера несопоставимости задачи n_0 , η^* — абсцисса пересечения кривой $\delta(\eta)$ с прямой $\mu\eta$

хорошо известной задачи для СЛАУ с невозмущенной X .

Очевидно, что решения задачи n_0 :

$$\|\bar{\theta}\| = \inf \|\theta\| \text{ на } \Theta_{\delta_1} = \{\theta : \|\bar{y} - X\theta\| = \delta_1\}$$

и задачи n'_0 :

$$\|\bar{\theta}\|^2 = \inf \|\theta\|^2 \text{ на } \Theta_{\delta_1} = \{\theta : \|\bar{y} - X\theta\|^2 = \delta_1^2\}$$

совпадают. Последняя задача, сводящаяся к минимизации квадрата нормы вектора при квадратичных ограничениях, обсуждалась выше в п. 2 и может быть решена, например, методом неопределенных множителей Лагранжа. Имеется развитое программное обеспечение для решения задач этого класса [9; 83].

В заключение отметим, что дальнейший анализ и примеры применения РМК читатель может найти в [84; 85].

Глава 10

ВВЕДЕНИЕ В ПЛАНИРОВАНИЕ РЕГРЕССИОННЫХ ЭКСПЕРИМЕНТОВ

1. До сих пор число точек измерений и сетка аргумента предполагались заданными. Тогда математическая статистика (например, регрессионный анализ)

дает рецепт извлечения максимальной информации из измерений или позволяет использовать не эффективные, но более простые или более универсальные (робастные оценки) методы. Однако точность получаемых выводов зависит не только от методики статистической обработки наблюдений, но и от методики измерений, способа проведения эксперимента. За счет более рационального выбора точек измерений, числа замеров в каждой точке возможно значительное сокращение количества измерений или, при том же числе измерений, существенное повышение надежности выводов. В свою очередь, это означает более эффективное использование аппаратуры, снижение затрат на экспериментальные исследования, что очень важно для дорогостоящих экспериментов, например в ядерной физике, космических исследованиях.

Вопросы эти рассматриваются в теории планирования экспериментов, которая основывается на математической статистике и методах решения экстремальных задач. В планировании экспериментов различаются две постановки: 1) планирование экстремальных экспериментов [63], т. е. поиск условий, при которых изучаемый процесс, эксперимент удовлетворяет некоторому критерию оптимальности (например, нужно подобрать температуру, давление, концентрацию реагентов, при которых достигается максимальный выход реакции), 2) планирование экспериментов по выяснению механизма явлений (т. е. поиск математической модели данного процесса). При проведении и обработке исследовательских экспериментов актуальные задачи второго типа; они рассмотрены, например, в работе [94], схеме изложения которой мы в основном и следуем.

Перепишем некоторые из формул гл. 6 в обычно принятых в теории планирования эксперимента обозначениях [56]. В случае линейной модели (5) из гл. 6 результат наблюдения в точке x_j может быть представлен как

$$y_j = \vec{\theta}^T \vec{f}(x_j) + \varepsilon_j, \quad (1)$$

где $\vec{f}(x_j) = (X_{j1}, X_{j2}, \dots, X_{jk})^T$. Пусть $\vec{f}(x_j)$ — известные функции, так что целью эксперимента является

не выбор модели из нескольких конкурирующих, а лишь определение параметров модели — вектора $\vec{\theta}$. Предполагаем, что y_j имеют дисперсию σ_j^2 и независимы. Тогда, как показано в (11) из гл. 6, НЛНО параметров $\theta_1, \dots, \theta_k$ является МНК-оценка

$$\hat{\vec{\theta}} = \Phi^{-1} \vec{Y}, \quad (2)$$

где

$$\Phi = \sum_{j=1}^n \omega_j \vec{f}(x_j) \vec{f}^T(x_j), \quad \vec{Y} = \sum_{j=1}^n \omega_j \vec{f}(x_j) y_j, \quad \omega_j = \sigma_j^{-2}, \quad (3)$$

причем ковариационная матрица для $\hat{\vec{\theta}}$ определяется (12) из гл. 6, или

$$D(\hat{\vec{\theta}}) = \text{cov}(\hat{\vec{\theta}}, \hat{\vec{\theta}}) = \Phi^{-1}. \quad (4)$$

Матрица Φ называется информационной матрицей Фишера. Она является положительно полуопределенной и представима в виде

$$\Phi = FF^T, \quad F \{ \sqrt{\omega_j} \vec{f}(x_j) \} \quad (5)$$

(в случае $k=n$ F — квадратная матрица, и из $|\det F| > 0$ следует $\Phi > 0$). Ниже для упрощения обозначений нередко опускается значок вектора там, где это не вызывает недоразумений.

2. В гл. 6 мы видели, что $c^T \hat{\vec{\theta}}$ является оптимальной оценкой линейной формы $c^T \vec{\theta}$, а компоненты вектора $\hat{\vec{\theta}}$ имеют минимальную дисперсию в классе линейных несмещенных оценок. Однако возможно сравнение оценок и по другим критериям, например:

1) $\hat{\vec{\theta}}$ лучше $\tilde{\vec{\theta}}$, если $D(\hat{\vec{\theta}}) = D(\tilde{\vec{\theta}}) + d$, где матрица $d > 0$;

2) $\hat{\vec{\theta}}$ лучше $\tilde{\vec{\theta}}$, если $\det D(\hat{\vec{\theta}}) = |D(\hat{\vec{\theta}})| > |D(\tilde{\vec{\theta}})|$.

Заметим, что поскольку $D(\hat{\vec{\theta}}), D(\tilde{\vec{\theta}}) \geq 0$ — симметричные матрицы, то из $D(\hat{\vec{\theta}}) \geq D(\tilde{\vec{\theta}})$ следует $|D(\hat{\vec{\theta}})| \geq |D(\tilde{\vec{\theta}})|$. Поэтому оценки, частично упорядоченные по первому критерию, сохраняют тот же порядок по второму.

Нетрудно показать, что МНК-оценка оптимальна и по этим двум критериям. Для этого достаточно перейти к преобразованным величинам $\eta_j = y_j \sqrt{w_j}$, $j = \overline{1, n}$, для которых регрессионная задача имеет вид

$$\vec{\eta} = F^T \vec{\theta} + \vec{\varepsilon}, \quad \text{cov}(\varepsilon_i, \varepsilon_j) = \delta_{ij}. \quad (6)$$

Ее решение $\vec{\theta} = (FF^T)^{-1} F \vec{\eta}$ совпадает с (2). Для модели (6) доказательство оптимальности по первому критерию аналогично доказательству [130] свойства 3 МНК-оценок (гл. 6). Из эквивалентности решений (6) и (1) следует оптимальность МНК-оценок для общей линейной модели с независимыми наблюдениями по критерию 1. Из сделанного выше замечания следует оптимальность оценок НК и по второму критерию.

3. Определение 1. Экспериментом \mathcal{E}_l называется совокупность величин

$$y_{11}^{(l)}, y_{12}^{(l)}, \dots, y_{1r_1}^{(l)}, y_{21}^{(l)}, \dots, y_{2r_2}^{(l)}, \dots, y_{n_1}^{(l)}, \dots, y_{nr_n}^{(l)}, \quad (7)$$

$$\begin{array}{ccc} \omega_1^{(l)} & \omega_2^{(l)} & \omega_n^{(l)} \\ x_1^{(l)} & x_2^{(l)} & x_n^{(l)} \end{array}$$

т. е. совокупность точек измерения $\{x_1^{(l)}, \dots, x_n^{(l)}\}$, погрешностей измерения в этих точках и значений наблюдаемых величин (предполагается, что y измеряется в точке $x_j^{(l)}$ r_j раз).

Два эксперимента \mathcal{E}_1 и \mathcal{E}_2 различны ($\mathcal{E}_1 \neq \mathcal{E}_2$), если у них различаются хотя бы две величины, определенные в (7).

Можно показать [45; 56; 68; 94], что оценка (2) справедлива и в случае повторных измерений в точках, как в (7), однако в этом случае

$$\Phi \sum_{j=1}^n \omega_j r_j \vec{f}(x_j) \vec{f}^T(x_j), \quad \vec{Y} = \sum_{j=1}^n \sum_{i=1}^{r_j} \omega_j \vec{f}(x_j) y_{ji}. \quad (3')$$

Если, отвлекаясь от других соображений, сравнивать эксперименты по степени локализованности оптимальных оценок (как мы видели, ими являются для НЛНО оценки МНК), то распространенные критерии

качества эксперимента следующие: \mathcal{E}_1 предпочтительнее \mathcal{E}_2 , если:

- а) $|D\hat{\theta}(\mathcal{E}_1)| < |D\hat{\theta}(\mathcal{E}_2)|$;
- б) $D(\hat{\theta}(\mathcal{E}_1)) < D(\hat{\theta}(\mathcal{E}_2))$;
- в) $\text{tr}[D(\hat{\theta}(\mathcal{E}_1))] < \text{tr}[D(\hat{\theta}(\mathcal{E}_2))]$;
- г) $\max_{\alpha} D_{\alpha\alpha}(\hat{\theta}(\mathcal{E}_1)) < \max_{\alpha} D_{\alpha\alpha}(\hat{\theta}(\mathcal{E}_2))$.

(8)

Однако может случиться, что осуществление эксперимента, лучшего по одному из критериев а)–г), требует сложной процедуры измерений. Эти факторы можно учесть, введя функцию потерь

$$R[\mathcal{E}] = \tau + \psi[D(\hat{\theta}(\mathcal{E}))], \quad (9)$$

где τ — затраты на проведение эксперимента \mathcal{E} , ψ — функционал заданного вида, характеризующий выигрыш от повышения точности оценивания $\hat{\theta}$. Нередко он линеен: $\psi[D(\hat{\theta})] = c\mathcal{L}[D(\hat{\theta})]$, где c — постоянная, \mathcal{L} — линейный функционал, определяемый критерием качества эксперимента, например, совпадающий с одной из величин, введенных в (8). В простейшем случае затраты c_j на измерения в точке x_j не зависят от номера измерения, а τ равен сумме затрат во всех точках:

$$\tau = \sum_{j=1}^n c_j r_j.$$

Введем теперь основные для дальнейшего определения планов.

Определение 2. Совокупность величин

$$x_1, \dots, x_n; r_1, \dots, r_n \left(\sum_{j=1}^n r_j = N \right) \quad (10)$$

называется планом эксперимента $\mathcal{E}(N)$.

Определение 3. Нормированным планом $\varepsilon(N)$ называется совокупность

$$x_1, \dots, x_n; p_1, \dots, p_n \quad (p_j = r_j/N, \sum p_j = 1). \quad (11)$$

Определения 2, 3 разумны, поскольку матрица $D(\hat{\theta}) =$

$=\Phi^{-1}$ не зависит от значений величины y или от значений параметров. При этом

$$\Phi(\varepsilon(N)) = N \sum_{j=1}^n p_j \omega_j f(x_j) f^T(x_j) = N\Phi(\varepsilon(N)),$$

так что

$$D(\varepsilon(N)) = N^{-1}D(\varepsilon(N)). \quad (12)$$

Тогда для заданного нормированного плана $\varepsilon(N)$ в простейшем случае функция потерь имеет вид

$$R[N, \varepsilon(N)] = N \sum_{j=1}^n c_j p_j + cN^{-k} |D(\varepsilon(N))| \quad (13)$$

в случае а), а в остальных случаях

$$R[N, \varepsilon(N)] = N \sum_{j=1}^n c_j p_j + cN^{-1} \mathcal{L}[D(\varepsilon(N))]. \quad (13')$$

Теперь планирование эксперимента сводится к экстремальной задаче поиска $\min_{N, \varepsilon(N)} R[N, \varepsilon(N)]$. Если в

(13') N задано, а $c_1 = c_2 = \dots = c_n = c_0$, то $R[N, \varepsilon(N)] = Nc_0 + cN^{-1} \mathcal{L}[D(\varepsilon(N))]$ и надо минимизировать $\mathcal{L}[D(\varepsilon(N))]$ (аналогично в случае (13)).

Чтобы применять аппарат решения экстремальных задач, использующий математический анализ, обобщим понятие нормированного плана.

Определение 4. Непрерывным нормированным планом ε называется совокупность величин

$$x_1, \dots, x_n; \quad p_1, \dots, p_n \quad (\sum p_i = 1), \quad (14)$$

где $0 < p_i < 1$ — любые вещественные числа.

Понятно, что если N велико, оптимальный непрерывный нормированный план с большой точностью может быть приближен дискретным нормированным планом $\varepsilon(N)$.

Определение 5. Пусть X — некоторая замкнутая область, $x \in X$ — ее точки и на X заданы функция $\lambda(x) \geq 0$ и нормированная на 1 мера $\xi(x)$, а также пусть

$$\Phi(\varepsilon) = \int_X \lambda(x) f(x) f^T(x) d\xi(x), \quad (15)$$

при абсолютной непрерывности $\xi(x)$

$$\Phi(\varepsilon) = \int_X \lambda(x) f(x) f^T(x) p(x) dx, \quad \int p(x) dx = 1. \quad (15')$$

Тогда совокупность $\{X, \xi(x), \lambda(x)\}$ называется обобщенным непрерывным планом.

Легко понять, что $\xi(x)$ обобщает понятие количества (доли) измерений $r_j(p_j)$, а $\lambda(x)$ — понятие весов w_j . В дальнейшем будем считать обычно, что $\lambda(x) = \sigma^2 \gamma(x)$, где γ — известная функция; для простоты выкладок будем полагать $\sigma^2 = 1$.

4. Вернемся к функции потерь (9) в простейшем случае (13), (13'). При фиксированном плане ε в случае (13) имеем

$$\frac{\partial R[N, \varepsilon]}{\partial N} = c_0 - kcN^{-(k+1)} |D(\varepsilon)| = 0, \quad (16)$$

так что оптимальное $N_{\text{опт}} = [kc |D(\varepsilon)| / c_0]^{\frac{1}{k+1}}$, а в случае (13') необходимое условие экстремума приводит к

$$N_{\text{опт}} = [c \mathcal{L}[D(\varepsilon)] / c_0]^{1/2},$$

при этом, чтобы такое определение N было разумным, $|D(\varepsilon)| \gg c_0$, $c \mathcal{L}[D(\varepsilon)] \gg c_0$. Тогда, как легко видеть, функция потерь, отвечающая оптимальному N , связана прямой зависимостью с соответствующим критерием (8) оптимальности. Так, подставляя найденное N в (13) (при этом опустим знак $[\cdot]$ целой части числа), имеем

$$R[N_{\text{опт}}, \varepsilon] = c_0^{\frac{k}{k+1}} (kc |D(\varepsilon)|)^{\frac{1}{k+1}} + \\ + (c |D(\varepsilon)|)^{\frac{1}{k+1}} \left(\frac{c_0}{k} \right)^{\frac{k}{k+1}} \quad (17)$$

(аналогично в случае (13), но там $R[N_{\text{опт}}, \varepsilon] \sim \mathcal{L}[D(\varepsilon)]^{1/2}$). Таким образом, задача оптимизации функции потерь может быть решена в два этапа: 1) ищем план эксперимента, удовлетворяющий одному из условий оптимальности (8); 2) для найденно-

го оптимального плана находим оптимальное число измерений $N_{\text{опт}}$. При этом величина $R[N, \epsilon]$ в оптимальной точке, как это видно из (17), зависит только от ϵ .

Определение 6. 1). Планы, являющиеся наилучшими в смысле (8, а), называются D -оптимальными. 2). Планы, минимизирующие максимум дисперсии оценки функции регрессии

$$\begin{aligned} \max_x d(x, \epsilon) &= \max_x f^T(x) D(\epsilon) f(x) = \\ &= \max_x E [f^T(x) \bar{\theta} - f^T(x) \theta]^2, \end{aligned} \quad (18)$$

называются минимаксными, или G -оптимальными. 3). Планы, минимизирующие $k^{-1} \text{tr}[D(\epsilon)]$, называются A -оптимальными. 4). Планы, минимизирующие $\max_{\alpha} D_{\alpha\alpha}(\epsilon)$, называются минимаксными в пространстве параметров.

Для дальнейшего нам понадобится [94] известная из линейной алгебры

Теорема (Каратеодори). Пусть Σ — подмножество n -мерного линейного пространства, Σ^* — его выпуклая оболочка, т. е. множество точек вида $s^* =$

$$= \sum_{i=1}^M \alpha_i s_i, \quad \alpha_i \geq 0, \quad \sum_{i=1}^M \alpha_i = 1, \quad s_i \in \Sigma. \quad \text{Всякую точку}$$

$s^* \in \Sigma^*$ можно представить как

$$s^* = \sum_{i=1}^{n+1} \alpha_i s_i, \quad \alpha_i \geq 0, \quad \sum_{i=1}^{n+1} \alpha_i = 1, \quad s_i \in \Sigma. \quad (19)$$

Докажем теперь основную теорему о планах эксперимента.

Теорема 1. 1). Для всякого плана ϵ матрица $\Phi(\epsilon)$ — симметричная положительно полуопределенная. 2). Если спектр плана (т. е. точки x_1, \dots, x_n) состоит менее чем из k точек, то матрица Φ вырожденная. 3). Множество матриц $\{\Phi(\epsilon)\}$, отвечающее всем нормированным планам, выпукло. Если $f(x)$ и $\lambda(x)$ непрерывны, то $\{\Phi(\epsilon)\}$ замкнуто. 4). Для любого плана ϵ найдется план $\bar{\epsilon}$, спектр которого содержит не более $k(k+1)/2+1$ точек, такой, что $\Phi(\bar{\epsilon}) = \Phi(\epsilon)$.

Доказательство. 1). Симметричность Φ следует из (15), а положительная полуопределенность — из

$$\begin{aligned} z^T \Phi(\varepsilon) z &= \int \lambda(x) z^T f(x) f^T(x) z d\xi(x) = \\ &= \int \lambda(x) [z^T f(x)]^2 d\xi(x) \geq 0. \end{aligned}$$

2). План с конечным числом точек может быть представлен суперпозицией точечных планов (т. е. сосредоточенных в одной точке x_j)

$$\Phi(\varepsilon) = \sum_{j=1}^n p_j \lambda(x_j) f(x_j) f^T(x_j) = \sum_{j=1}^n \Phi_j(x_j) p_j.$$

Как известно из линейной алгебры, $\text{rank } \Phi(\varepsilon) \leq \leq \sum \text{rank } \Phi_j = n$, так что при $n < k$ матрица $\Phi(\varepsilon)$ вырожденная.

3). Пусть ε_1 и ε_2 — два произвольных обобщенных нормированных плана, заданных на X . Их линейной комбинацией $\varepsilon = (1-\alpha)\varepsilon_1 + \alpha\varepsilon_2$, $0 \leq \alpha \leq 1$, является нормированный план с мерой $\xi(x) = (1-\alpha)\xi_1(x) + \alpha\xi_2(x)$, т. е. множество планов выпукло. Из (15) легко видеть, что $\Phi(\varepsilon) = (1-\alpha)\Phi(\varepsilon_1) + \alpha\Phi(\varepsilon_2)$ является информационной матрицей плана ε , т. е. $\{\Phi(\varepsilon)\}$ выпукло. Замкнутость $\{\Phi(\varepsilon)\}$ следует из замкнутости X и непрерывности $f(x)$ и $\lambda(x)$.

4). В силу симметричности $\Phi(\varepsilon)$ полностью определяется своими $k(k+1)/2$ элементами, т. е. вектором из $E_{k(k+1)/2}$. Из определения $\Phi(\varepsilon)$ следует, что множество векторов, определяющих $\Phi(\varepsilon)$, — выпуклая оболочка векторов, отвечающих информационным матрицам $\Phi(x_j)$ точечных планов. Из теоремы Каратеодори следует справедливость 4).

Практическим следствием из этой теоремы является то, что бессмысленно проводить измерения менее чем в k точках и более чем в $n_0 = k(k+1)/2 + 1$ точках. Поэтому увеличение количества точек измерений до чисел, больших n_0 , не дает выигрыша в точности оценок.

5. Введенные выше 4 определения оптимальных планов приводят, вообще говоря, к различным стратегиям проведения эксперимента. Эта неэквивалентность остается в силе, пока мы оперируем с дискрет-

ными планами. Однако при переходе к непрерывным планам удается установить ряд теорем эквивалентности [56; 94], в частности доказать эквивалентность D -оптимальных и минимаксных планов в случае равноточных измерений (т. е. при $\lambda(x) \equiv 1$).

Лемма 1. Пусть $\Phi(\varepsilon)$ неособенная. Тогда: 1) взвешенная сумма дисперсий оценки поверхности отклика $d(x, \varepsilon)$, взятая по всем точкам плана ε , равна числу неизвестных параметров k :

$$\sum_{j=1}^n p_j \lambda(x_j) d(x_j, \varepsilon) = k \quad (20)$$

или (более обще)

$$\int_X \lambda(x) d(x, \varepsilon) d\xi(x) = k; \quad (20')$$

2) минимальное значение $\max_{x \in X} \lambda(x) d(x, \varepsilon)$ не может быть меньше k :

$$\max_{x \in X} \lambda(x) d(x, \varepsilon) \geq k. \quad (21)$$

Доказательство. 1). Чтобы установить справедливость первого утверждения, заметим, что интеграл, стоящий в (20'), выражается через след матрицы, и воспользуемся свойством коммутативности следа (гл. 6):

$$\begin{aligned} \int_X \lambda(x) d(x, \varepsilon) d\xi(x) &= \int \lambda(x) f^T(x) \Phi^{-1}(\varepsilon) f(x) d\xi(x) = \\ &= \int \lambda(x) \operatorname{tr} [f^T(x) \Phi^{-1}(\varepsilon) f(x)] d\xi(x) = \\ &= \operatorname{tr} [\Phi^{-1}(\varepsilon) \int \lambda(x) f(x) f^T(x) d\xi(x)] = \\ &= \operatorname{tr} [\Phi^{-1}(\varepsilon) \Phi(\varepsilon)] = \operatorname{tr} I_k = k. \end{aligned}$$

2). Используя доказанное выше и теорему о среднем, имеем

$$\int_X \lambda(x) d(x, \varepsilon) d\xi(x) = d^* \int_X d\xi(x) = d^* = k,$$

где

$$\min_x \lambda(x) d(x, \varepsilon) \leq d^* \leq \max_x \lambda(x) d(x, \varepsilon).$$

откуда следует справедливость второго утверждения.

Лемма 2. Функция $\ln|\Phi(\varepsilon)|$ строго вогнута.

Доказательство. Было показано, что $\{\Phi(\varepsilon)\}$ — выпуклое множество. Согласно определению вогнутости [64], нужно показать, что для $\Phi = (1-\alpha)\Phi_1 + \alpha\Phi_2$, $\Phi_1 \neq \Phi_2$, $\Phi_1, \Phi_2 > 0$ — матрицы планов, $0 < \alpha < 1$, справедливо

$$\ln|\Phi| > (1-\alpha)\ln|\Phi_1| + \alpha\ln|\Phi_2|. \quad (22)$$

Рассматривая Φ_1 и Φ_2 как матрицы коэффициентов положительно определенных квадратичных форм и используя тот факт, что две такие формы одновременно приводятся к диагональному виду, можем утверждать, что найдется матрица G такая, что

$$G^T \Phi_1 G = I_k, \quad G^T \Phi_2 G = D = \text{diag}\{d_1, \dots, d_k\}.$$

Матрица G невырождена, поэтому смысл неравенства (22) не изменится, если перейти от Φ, Φ_1, Φ_2 к $G^T \Phi G, I_k, D$ соответственно. Поэтому достаточно доказать лемму для $\Phi_1' = I_k, \Phi_2' = D$. Для них (22) имеет форму $\ln|(1-\alpha)I_k + \alpha D| > (1-\alpha)\ln|I_k| + \alpha\ln|D| = \alpha\ln|D|$. Матрица $(1-\alpha)I_k + \alpha D$ диагональная, так что последнее неравенство перепишем в виде

$$\sum_{i=1}^k \ln(1-\alpha + \alpha d_i) > \alpha \sum_{i=1}^k \ln|d_i|. \quad (22')$$

Рассматривая i -е члены слева и справа, видим, что имеет место неравенство $\ln(1-\alpha + \alpha d_i) > \alpha \ln d_i$, выражающее строгую вогнутость $\ln x$ на $[1, d_i]$. Суммируя такие неравенства по i , получим (22'), а следовательно, и (22).

Лемма 3. Пусть ε_1 и ε_2 — планы с информационными матрицами $\Phi(\varepsilon_1)$ и $\Phi(\varepsilon_2)$, $\Phi(\varepsilon)$ — информационная матрица плана $\varepsilon = (1-\alpha)\varepsilon_1 + \alpha\varepsilon_2$, $0 \leq \alpha \leq 1$. Тогда

$$\frac{\partial}{\partial \alpha} \ln|\Phi(\varepsilon)| = \text{tr}[\Phi^{-1}(\varepsilon) \{\Phi(\varepsilon_2) - \Phi(\varepsilon_1)\}]. \quad (23)$$

Доказательство. Очевидно, что $\Phi(\varepsilon) = (1-\alpha)\Phi(\varepsilon_1) + \alpha\Phi(\varepsilon_2)$. Продифференцировав $\ln|(1-\alpha)\Phi(\varepsilon_1) + \alpha\Phi(\varepsilon_2)|$ и используя формулу [56] $\frac{\partial}{\partial t} \ln|A(t)| = \text{tr} \left[A^{-1} \frac{\partial A}{\partial t} \right]$, получим

$$\begin{aligned} \frac{\partial}{\partial \alpha} \ln|\Phi(\varepsilon)| &= \text{tr} \left[\Phi^{-1}(\varepsilon) \frac{\partial \Phi(\varepsilon)}{\partial \alpha} \right] = \\ &= \text{tr} [\Phi^{-1}(\varepsilon) \{\Phi(\varepsilon_2) - \Phi(\varepsilon_1)\}]. \end{aligned}$$

Теорема 2 (Кифер, Вольфовитц). *Следующие утверждения эквивалентны:*

- 1) план ε^* максимизирует $|\Phi(\varepsilon)|$;
- 2) план ε^* минимизирует $\max_{x \in X} \lambda(x) d(x, \varepsilon)$;
- 3) $\max_{x \in X} \lambda(x) d(x, \varepsilon^*) = k$.

Информационные матрицы планов, удовлетворяющих 1) — 3), совпадают. Любая линейная комбинация таких планов также удовлетворяет 1) — 3).

Доказательство. Из 1) следует 2). Пусть $\varepsilon^* = \text{Arg} \max |\Phi(\varepsilon)|$, а ε — произвольный план. Рассмотрим $\varepsilon = (1-\alpha)\varepsilon^* + \alpha\varepsilon$, $0 \leq \alpha \leq 1$. По лемме 3

$$\begin{aligned} \frac{\partial}{\partial \alpha} \ln|\Phi(\varepsilon)|_{\alpha=0} &= \text{tr} [\Phi^{-1}(\varepsilon^*) \{\Phi(\varepsilon) - \Phi(\varepsilon^*)\}]_{\alpha=0} = \\ &= \text{tr} [\Phi^{-1}(\varepsilon^*) \Phi(\varepsilon)] - k. \end{aligned}$$

В силу определения ε^* эта производная меньше или равна 0. Пусть ε — точечный план со спектром $x \in X$. Тогда

$$\begin{aligned} \text{tr} [\Phi^{-1}(\varepsilon^*) \Phi(\varepsilon)] - k &= \\ &= \text{tr} [\Phi^{-1}(\varepsilon^*) \lambda(x) f(x) f^T(x)] - k = \\ &= \lambda(x) d(x, \varepsilon^*) - k \leq 0 \end{aligned}$$

(снова воспользовались коммутативностью следа). Но по (21) для любого плана, в том числе для ε^* , $\max_{x \in X} \lambda(x) d(x, \varepsilon^*) \geq k$, т. е. на самом деле имеет место равенство. Поэтому D -оптимальный план минимизирует (21).

Из 2) следует 1). Предположим, что ε^* , минимизирующий $\max \lambda(x) d(x, \varepsilon)$, не D -оптимален. Тогда найдется такой план ε , что

$$\begin{aligned} \frac{\partial}{\partial \alpha} \ln |(1 - \alpha) \Phi(\varepsilon^*) + \alpha \Phi(\varepsilon)|_{\alpha=0} = \\ = \text{tr} [\Phi^{-1}(\varepsilon^*) \Phi(\varepsilon)] - k > 0. \end{aligned} \quad (24)$$

По теореме 1 любой план ε можно представить как суперпозицию не более чем $k(k+1)/2 + 1 = n_0$ точечных планов $\varepsilon(x_j)$. Поэтому будем считать, что спектр ε состоит из конечного числа точек. Тогда $\text{tr} [\Phi^{-1}(\varepsilon^*) \Phi(\varepsilon)] - k = \sum_{j=1}^{n_0} p_j \lambda(x_j) d(x_j, \varepsilon^*) - k$. Поскольку ε^* удовлетворяет 2), $\lambda(x) d(x, \varepsilon^*) \leq k$, следовательно,

$$\sum_{j=1}^{n_0} p_j \lambda(x_j) d(x_j, \varepsilon^*) - k \leq k \sum_{j=1}^{n_0} p_j - k = 0. \quad (25)$$

Неравенства (24) и (25) совместны, лишь если план ε^* , удовлетворяющий 2), D -оптимален.

Эквивалентность 1) и 3), 2) и 3). Выше мы попутно доказали, что из 1) и 2) следует 3). Обратное утверждение: из 3) следует 2) вытекает из леммы 1. В силу эквивалентности 1) и 2) утверждения 1) и 3) также эквивалентны.

Докажем теперь последние утверждения теоремы. Пусть ε_1 и ε_2 — D -оптимальные планы с матрицами $\Phi(\varepsilon_1)$ и $\Phi(\varepsilon_2)$, причем $\Phi(\varepsilon_1) \neq \Phi(\varepsilon_2)$. Рассмотрим $\Phi(\varepsilon) = (1 - \alpha) \Phi(\varepsilon_1) + \alpha \Phi(\varepsilon_2)$, $0 < \alpha < 1$. По лемме 2

$$\ln |\Phi(\varepsilon)| > (1 - \alpha) \ln |\Phi(\varepsilon_1)| + \alpha \ln |\Phi(\varepsilon_2)|. \quad (26)$$

Но в силу D -оптимальности $|\Phi(\varepsilon_1)| = |\Phi(\varepsilon_2)| \geq |\Phi(\varepsilon)|$, что противоречит (26). Следовательно, неверно исходное предположение и $\Phi(\varepsilon_1) = \Phi(\varepsilon_2)$.

Следствие. В точках оптимального плана ε^* функция $\lambda(x) d(x, \varepsilon^*)$ достигает максимального значения k .

Предположим, что найдется $x \in \varepsilon^*$ такая, что $\lambda(x)d(x, \varepsilon^*) < k$. Тогда в силу третьего утверждения теоремы

$$\sum_{j=1}^{n_0} p_j \lambda(x_j) d(x_j, \varepsilon^*) < \sum_{j=1}^{n_0} p_j k = k,$$

что противоречит лемме 1.

Теорема 2 и следствие из нее облегчают нахождение D -оптимальных и минимаксных планов. Примеры нахождения конкретных оптимальных планов, в частности, для модели полиномиальной или тригонометрической регрессии, читатель может найти в [56; 94]. Там же рассмотрены итерационные процедуры построения оптимальных планов.

Глава 11

ПРИМЕНЕНИЕ МЕТОДА МОНТЕ-КАРЛО В АНАЛИЗЕ И ПЛАНИРОВАНИИ ЭКСПЕРИМЕНТОВ

При решении ряда практических задач результаты математической теории планирования экспериментов оказываются недостаточными. Прежде всего, это обусловлено тем, что точные оптимальные планы, основанные на аналитических выражениях для информационной матрицы Фишера и связанных с ней величин, характеризующих тот или иной критерий оптимальности, удается теоретически исследовать на экстремум лишь в редких случаях. Чаще оптимизация производится численно, обычно итерационными методами. Поэтому даже в линейном случае процедура построения оптимального плана довольно громоздка [56; 94]. Однако в современных экспериментах зависимость функции отклика от параметров, как правило, нелинейна, так что вычислительные трудности еще возрастают.

Во-вторых, нередко задачи планирования режима измерений, конфигурации экспериментальной аппаратуры отличаются от обычно рассматриваемых в математической теории планирования эксперимента [56; 94]. Так, в спектрометрических измерениях зависимости числа регистрируемых частиц от энергии обыч-

но производится группировка отсчетов в интервале энергий $[E_i, E_i + \Delta E_i]$ (внутри i -го спектрального канала). Поэтому результатом эксперимента фактически является гистограмма. В этом случае задача планирования эксперимента заключается в выборе расположения спектральных каналов, их ширины ΔE_i и общего числа отсчетов, необходимых для достижения желаемой точности выводов при интерпретации результатов.

Задачи такого типа могут быть решены с помощью метода статистических испытаний (метода Монте-Карло) [26; 36; 37; 70]. Под методом Монте-Карло в широком смысле понимается любой способ решения задачи, при котором используются случайные или псевдослучайные числа. Обычно он сводится к тому, что моделируются на ЭВМ случайные величины с заданным законом распределения, вычисляются их выборочные характеристики (например, моменты, эмпирическая ф. р.), которые используются в качестве приближений к соответствующим теоретическим значениям. Обычно выборочные характеристики связаны с выборочным средним, поскольку выборочное среднее, как мы видели в гл. 5, 6, является эффективной оценкой при нормальном законе распределения и наилучшей несмещенной оценкой (НЛНО) в классе линейных оценок.

Правомерность метода Монте-Карло вытекает из закона больших чисел в различных его формах [25; 93]. Пусть ξ_1, \dots, ξ_n — независимые одинаково распределенные с.в. Простейшая форма закона больших чисел следует из неравенства Чебышева, которое справедливо при условии существования конечных м.о. $E\xi_k = a < \infty$ и дисперсии $D\xi_k = \sigma^2 < \infty$. Пусть

$s_n = \sum_{k=1}^n \xi_k$ и $\epsilon > 0$ произвольно. Тогда справедливо неравенство Чебышева

$$P\{|s_n/n - a| \leq \epsilon\} \geq 1 - \sigma^2/n\epsilon^2. \quad (1)$$

Поэтому для всякого $\epsilon > 0$ и всякого $\delta > 0$ найдется n_0 такое, что для всех $n > n_0$ $P\{|s_n/n - a| \leq \epsilon\} \geq 1 - \delta$.

Возможна более слабая формулировка закона в форме Хинчина, в которой при условии существования м.о. $E\xi_k = a < \infty$ справедливо соотношение

$\lim_{n \rightarrow \infty} P\{|s_n/n - a| > \epsilon\} = 0$. Согласно усиленному закону больших чисел в форме Колмогорова при условии существования конечного м.о. при $n \rightarrow \infty$ имеем $P\{s_n/n \rightarrow a\} = 1$. Как видно, условия применимости метода Монте-Карло весьма общие и выполняются в большинстве практических задач.

Скорость сходимости выборочных величин к их математическим ожиданиям оценивается исходя из других теорем теории вероятностей, чаще всего из центральной предельной теоремы. Так, если конечны дисперсия σ^2 и $\beta_3 = E|\xi_k - a|^3 < \infty$, то

$$\left| P\{s_n - an > \sigma\sqrt{nx}\} - (2\pi)^{-1/2} \int_x^\infty e^{-u^2/2} du \right| \leq \frac{c_0 \beta_3}{\sigma^3 \sqrt{n}},$$

где c_0 — постоянная, оцениваемая неравенствами $0,9051 > c_0 \geq (2\pi)^{-1/2}$ [107]. Таким образом, скорость сходимости метода Монте-Карло довольно мала ($\approx n^{-1/2}$). Однако существуют методы преобразования с.в., оставляющие неизменными их математические ожидания, но уменьшающие дисперсии [36; 37; 46; 70], что позволяет улучшить сходимость монте-карловских величин.

Рассмотрим теперь источники случайных чисел, необходимых для реализации метода Монте-Карло. Старейшим способом является использование таблиц случайных чисел, когда числа, следующие определенному закону распределения, обычно равномерному, заранее «запасаются впрок». Получают их с помощью некоторых физических устройств или процессов, обеспечивающих требуемые свойства. Недостатком метода таблиц является громоздкость их хранения, что требует при использовании на ЭВМ больших объемов памяти.

Другой способ получения случайных чисел — использование специальных физических датчиков (например, основанных на радиоактивном распаде или флуктуациях сигналов в электронных устройствах), связанных с ЭВМ. Этот метод не требует большого объема памяти, но применяется редко, поскольку не обеспечивает возможности повторного воспроизведения данной последовательности с.в., а главное, чрез-

вычайно трудно обеспечить стабильность генерации чисел с заданным распределением.

Общераспространенным в настоящее время является способ генерации на ЭВМ так называемых псевдослучайных чисел. При этом получается не случайным, а детерминированным образом совокупность чисел, свойства которых аналогичны свойствам выборки с заданным законом распределения. Обычно исходной задачей является получение равномерно распределенных псевдослучайных чисел, из которых с помощью стандартных способов имеем тот или иной закон распределения. Как правило, с помощью рекуррентных соотношений для целых чисел вычисляется $m_{i+k} = M(m_{i+k-1}, \dots, m_i)$, где k фиксировано, $i = 1, 2, \dots$, M — целочисленная функция целочисленных аргументов, $0 < M < P$, $P > 0$ — целое, числа m_1, \dots, m_k неотрицательны, не превосходят P и задаются в начале счета. Последовательность чисел $\alpha_1, \alpha_2, \dots$, удовлетворяющих дискретному равномерному распределению [33] на $(0, 1)$, получается как $\alpha_i = m_i/P$, $i = 1, 2, \dots$. Если число P достаточно велико, скажем, близко к максимальному целому, представимому на данной ЭВМ, последовательность $\{\alpha_i\}$ хорошо приближает непрерывное равномерное распределение.

Широко распространенным является метод сравнений Лемера [124] с рекуррентными формулами вида

$$m_{i+k} = \left(\sum_{j=0}^{k-1} b_j m_{i+j} + \theta \right) \pmod{P}, \quad i = 1, 2, \dots \quad (2)$$

Его простейшая модификация при $k=1$ и $\theta=0$ — мультипликативный метод вычисления произведения по данному модулю

$$m_i = L m_{i-1} \pmod{P}, \quad i = 1, 2, \dots, \quad (2')$$

где P взаимно просто с L . Например, неплохие свойства имеет последовательность с $L=5^{17}$, $P=2^{42}$, $m_0=1$.

Получаемые псевдослучайные числа проверяются по совокупности критериев согласия с заданной ф.р. (например, Пирсона или Колмогорова — Смирнова), отсутствия корреляционных связей между членами, однородности и т. д. [26]. Полная проверка по всей

совокупности тестов весьма трудоемка. Поэтому в зависимости от решаемой задачи нередко ограничиваются лишь частью тестов, например, проверкой значений нескольких первых моментов и проверкой по критерию χ^2 . Нужно также иметь в виду, что при детерминированном способе (2), (2') задания последовательностей и ограниченности количества несовпадающих целых чисел, меньших P , с некоторого момента начинается повторение последовательности. Таким образом, качество псевдослучайных чисел характеризует также максимальное число неповторяющихся чисел (длина серии), так что способ генерации и константа в рекуррентных соотношениях выбираются в соответствии с числом n монте-карловских реализаций с. в. ξ . Различные вычислительные алгоритмы генерации равномерных псевдослучайных чисел описаны в [47], хорошая программа на фортране приведена в [96].

Получение псевдослучайных чисел с заданной ф. р. из равномерных основано на преобразовании Н. В. Смирнова.

Пусть ξ — решение уравнения

$$\int_{-\infty}^{\xi} dF_{\xi}(t) = \alpha \quad (\text{т. е. } \xi = F_{\xi}^{-1}(\alpha)), \quad (3)$$

где $\alpha \sim R(0, 1)$. Тогда $\xi \sim F_{\xi}(x)$.

Доказательство. Поскольку $F_{\xi}(x)$ неубывающая, то $P\{\xi \leq x\} = P\left\{\int_{-\infty}^{\xi} dF_{\xi}(t) \leq \int_{-\infty}^x dF_{\xi}(t)\right\}$. По условию ξ — решение (3), так что $P\{\xi \leq x\} = P\{\alpha \leq F_{\xi}(x)\}$. α является равномерно распределенной $R(0, 1)$. Поэтому $P\{\alpha \leq F_{\xi}(x)\} = F_{\xi}(x)$, откуда $P\{\xi \leq x\} = F_{\xi}(x)$, т. е. с. в. ξ имеет ф. р. $F_{\xi}(x)$, что и требовалось доказать.

Из этого утверждения вытекают, в частности, способы генерации дискретных и непрерывных с. в. В дискретном случае, если с. в. принимает конечное число значений, вычисляют вероятности $p_i = P\{\xi = x_i\}$, $i = 1, k$, разбивают отрезок $(0, 1)$ на доли $\Delta_1, \Delta_2, \dots, \Delta_k$, длина $\Delta_i = p_i$, и рассматривают процесс попаданий с. в. $\alpha \sim R(0, 1)$ на $\Delta_1, \dots, \Delta_k$. Если $\alpha \in \Delta_i$, полагают $\xi = x_i$. Легко видеть, что так определенная

с. в. имеет заданную функцию вероятностей p_i . Если же число принимаемых ξ значений счетно, то обычно группируют вместе значения x_i , вероятность которых мала, в событие $y_{k+1} = \bigcup_i \{\xi = x_i\}$ и рассматривают, как и выше, конечное число событий x_1, \dots, x_k, y_{k+1} . При осуществлении события y_{k+1} по некоторому правилу присваивают ξ одно из значений x_i . В случае непрерывного закона распределения $F_\xi(x)$ — непрерывная монотонно неубывающая. Поэтому решение (3) может быть в принципе легко найдено, например, методом Ньютона.

Однако наряду с этим общим способом построения псевдослучайных чисел с заданными свойствами для широкоиспользуемых распределений разработаны специфические способы генерации. Так, можно показать, что число N , удовлетворяющее условию

$$\prod_{k=0}^N \alpha_k < \exp\{-\lambda\} < \prod_{k=0}^{N-1} \alpha_k,$$

имеет распределение Пуассона; здесь $\alpha_0, \alpha_1, \dots$ — независимые с. в. с распределением $R(0, 1)$ [37]. Для этого докажем, что $\xi_n = \prod_{k=0}^n \alpha_k$ имеет ф. п. в. $f_n(x) = (-\ln x)^n (n!)^{-1}$, $0 < x \leq 1$. Доказательство проведем по индукции. При $n=0$ утверждение очевидно. Пусть оно выполнено при $n=m$. Тогда

$$\begin{aligned} F_{m+1}(x) &= P\{\xi_{m+1} \leq x\} = P\{\alpha_{m+1} \xi_m \leq x\} = \\ &= \frac{1}{m!} \int_0^1 P\{\alpha_{m+1} \xi_m \leq x \mid \xi_m = y\} \times (-\ln y)^m dy = \\ &= \frac{1}{m!} \left(\int_0^x (-\ln y)^m dy + \int_0^1 \frac{x}{y} (-\ln y)^m dy \right). \end{aligned}$$

Отсюда ф. п. в.

$$\begin{aligned} f_{m+1}(x) = F'_{m+1}(x) &= \frac{1}{m!} \left\{ (-\ln x)^m - \int_x^1 \frac{(-\ln y)^m}{y} dy - \right. \\ &\left. - (-\ln x)^m \right\} = (-\ln x)^{m+1} / (m+1)!, \end{aligned}$$

что и требовалось. Вернувшись к с. в. N , имеем эквивалентность событий $\{N = m\}$ и $\{\xi_{m-1}\alpha_m < e^{-\lambda}\} \cap \{\xi_{m-1} > e^{-\lambda}\}$. Поэтому $P\{N = m\} = \frac{1}{(m-1)!} \int_0^1 (-\ln y)^{m-1} \times$
 $\times P\{\xi_{m-1}\alpha_m < e^{-\lambda} \cap \xi_{m-1} > e^{-\lambda} | \xi_{m-1} = y\} dy = e^{-\lambda}/(m-1)! \times$
 $\times \int_{e^{-\lambda}}^1 \frac{(-\ln y)^{m-1}}{y} dy = \lambda^m/m! e^{-\lambda}$.

Из непрерывных распределений наиболее употребительно нормальное. Часто генерация последовательности нормальных псевдослучайных чисел основана на центральной предельной теореме: если $\alpha_1, \dots, \alpha_m$ — независимые $R(0, 1)$ числа, то

$s_m = \sum_{i=1}^m \alpha_i$ при больших m имеет нормальное

$N(m/2, \sqrt{m/12})$ распределение. Поэтому величина $\xi = (s_m - m/2)/\sqrt{m/12} \sim N(0, 1)$. Во многих случаях удовлетворительный результат достигается уже при $m = 12$. При меньших m возможно улучшение качества случайных чисел специальными поправками при суммировании [26]. Другим распространенным способом получения нормальных псевдослучайных чисел является использование преобразования с. в. [87]:

$$\xi = \sqrt{-2 \ln \alpha_1} \cos 2\pi\alpha_2, \quad \eta = \sqrt{-2 \ln \alpha_1} \sin 2\pi\alpha_2, \quad (4)$$

$\alpha_1, \alpha_2 \sim R(0, 1)$ и независимы. Тогда $\xi, \eta \sim N(0, 1)$ и независимы. Справедливость этого утверждения проверяется по формуле для ф. п. в. функции от случайной величины [87], т. е. фактически сводится к замене переменных при интегрировании. Такие распространенные распределения, как χ^2 , Стьюдента, Снедекора — Фишера, могут быть вычислены согласно их определению через независимые нормальные с. в. [44; 87]. Впрочем, можно их моделировать, исходя непосредственно из равномерного закона, поскольку для $\alpha \sim R(0, 1)$ с. в. $\xi = \sqrt{-2 \ln \alpha} \sim \chi_1^2$, а χ_m^2 — распределение, воспроизводимое по m [44; 87].

Чаще всего метод Монте-Карло сводится к тому, что многократно (N раз) моделируется с. в. ξ с за-

данным законом распределения (ее значения называются реализациями), вычисляются интересующие нас функции от с.в. и их выборочные статистические характеристики (м.о., дисперсия, закон распределения и т. д.) по всем N реализациям. Тогда, как мы видели, с ростом N достигается все более точная аппроксимация соответствующих теоретических характеристик, причем точность приближения может быть оценена согласно (1) или по центральной предельной теореме. Типичным примером такого применения метода Монте-Карло является вычисление интегралов, объемов, площадей.

Пусть тело V сложной формы целиком уместается внутри куба объема Q . Моделируя попадание в V равномерно распределенных в Q с.в. и подсчитывая

общее число $N_V = \sum_{i=1}^N I_V(i)$, $I_V(i) = 1$ при $\xi_i \in V$,

$I_V(i) = 0$ при $\xi_i \in Q - V$, получим приближенный объем V как QM_V/N ; при достаточно большом N достигается удовлетворительная точность вычисления объема. Наряду с интегрированием метод Монте-Карло применяется при решении дифференциальных и интегральных уравнений, интерполяции, нахождении нулей функции, их экстремумов и т. д. [36; 37; 70].

При проектировании экспериментов и способов их математической обработки более важны иные два аспекта применения метода Монте-Карло. Это, прежде всего, имитация реальных физических процессов путем моделирования на ЭВМ отвечающих им с.в. и процессов [46]. Так, например, при расчете нейтронных реакторов, детекторов частиц и других установок и приборов в ядерной физике использование «классических» методов типа решения уравнений в частных производных наталкивается на принципиальные и вычислительные трудности, связанные со сложной картиной отражения и рассеяния частиц на ядрах вещества. В то же время происходящие процессы (рассеяние, ионизация, фотоэффект, образование пар и т. д.) имеют стохастическую природу и теоретически хорошо изучены. Поэтому, моделируя на ЭВМ судьбу отдельной частицы при ее взаимодействии с веществом и осуществляя достаточно большое число испытаний Монте-Карло, можно спроектировать уст-

ройство с нужными свойствами или определить характеристики имеющегося прибора [22]. К этому же типу решаемых методом Монте-Карло задач проектирования и планирования экспериментов относится упомянутая выше задача выбора оптимальной конфигурации спектральных каналов, если рассматривается сравнительно небольшой набор возможных конфигураций.

Второй аспект применения метода Монте-Карло в обработке экспериментов — исследование статистических свойств величин, получаемых при том или ином способе обработки. Так, в теории робастных оценок аналитические выражения для их эффективности имеют асимптотический характер и известны не для всех оценок (например, трудно получить их для итерационных робастных оценок). В этой ситуации эффективным средством исследования оказывается метод Монте-Карло [10; 23; 90; 102; 111]. M раз моделируются выборки объема l с заданным законом распределения; для каждой выборки вычисляются оценка, рекомендуемая нормальной теорией, и робастная оценка. Строится выборочное среднее для каждой из этих оценок по M выборкам и выборочная дисперсия. Эффективность робастной оценки вычисляется как отношение выборочных дисперсий по (2), (2') из гл. 5.

При других способах обработки результатов экспериментов используют квазиреальный эксперимент, состоящий в многократном моделировании на ЭВМ данных с присущим им распределением ошибок, применении к ним подлежащего исследованию способа математической обработки, вычислении выборочных характеристик по реализациям: среднего и дисперсии. Он является эффективным средством анализа алгоритмов и программ. Так, в нелинейном МНК линеаризованное приближение описывает лишь локальный участок функции отклика, тогда как для оценки точности результата, построения доверительных областей, проверки гипотез нужно рассматривать отклик в достаточно большой окрестности точки минимума функционала МНК. Искомые статистические характеристики могут быть получены осуществлением на ЭВМ квазиреального эксперимента [7; 32]. Те же проблемы возникают при применении регулярирую-

щих алгоритмов, особенно в нелинейных задачах. Таким образом, метод Монте-Карло является эффективным средством контроля качества математической обработки.

Возможна также проверка «несмещенности» используемых численных методов, того, что они не вносят систематических искажений в результаты. Так, имея измерения физической величины при различных значениях аргумента и зная ядро интегрального уравнения, связывающего искомую и измеряемую функции, методом регуляризации [78; 79] восстанавливаем искомую функцию. Проводя обратный анализ, численно интегрируем регуляризованное решение, умноженное на известное ядро; результат интегрирования многократно (N раз) возмущаем случайными ошибками с заданной ф.р. Сравнивая исходную экспериментальную кривую и реализации, оцениваем, насколько адекватно решение описывает экспериментальные данные. Далее, применяя к каждой «экспериментальной» кривой, смоделированной в реализации, регуляризирующий алгоритм, получим N «реализаций» восстановленной искомой функции. Тогда вычисление среднего по N позволяет оценить несмещенность процедуры решения, выборочные дисперсии в точках вычисления искомой функции характеризуют статистические ошибки, а ковариационная матрица, построенная по этим «реализациям», описывает меру зависимости между значениями в соседних точках.

Глава 12

ПРИМЕНЕНИЕ ФАКТОРНОГО АНАЛИЗА ДЛЯ ОЦЕНКИ ТОЧНОСТИ ДАННЫХ

Как мы видели при рассмотрении процедур выявления сбоев на экспериментальных кривых, нередко результат измерений можно трактовать как нормально распределенный вектор с коррелированными компонентами. Напомним, почему это так.

В режиме поточечного измерения экспериментальной кривой $y(t_i)$, (t_i) — сетка значений аргумента, $i=1, n$, нужное количество отсчетов набирается длительным измерением в фиксированной точке t_{i_0} , пока не накопится необходимая статистика, затем

выставляется значение t_{i+1} регулируемого параметра, проводятся контроль и регулировка режима экспериментальной установки и измерительной аппаратуры, набирается статистика в этой точке и т. д. Благодаря независимости поддержания режима работы аппаратуры в различных точках t_i измерения получаются некоррелированными, но они же имеют значительно большую погрешность, чем та, которая может вызываться статистическими свойствами изучаемых величин, так как установление режима аппаратуры внутри некоторой области допустимых значений параметров каждый раз происходит с ошибкой. Кроме того, много усилий и времени требуют контроль и поддержание необходимого режима аппаратуры для осуществления идентичности процесса измерений.

В последние годы в связи с развитием многоканальных счетных систем получил распространение автоматический режим измерений. При его осуществлении контроль состояния аппаратуры производится перед началом серии измерений; сетка $\langle t_i \rangle$ пробегается многократно за малое время τ ; информация, снятая в i -й точке, суммируется в i -м канале счетной системы; пробегание сетки $\langle t_i \rangle$ происходит до получения необходимого числа отсчетов. Идентичность процесса измерений достигается за счет быстрого перехода от точки к точке, так что состояние аппаратуры не успевает существенно измениться. Кривая получается измеренной с меньшими ошибками, чем при первой методике, лучше очерченной по форме. Однако медленное изменение параметров аппаратуры, происходящее неконтролируемым образом (аппаратурный дрейф), приводит к согласованным последствиям в близких точках, т. е. измерения оказываются коррелированными. Ковариационная матрица Q , описывающая вторые моменты измеряемого случайного вектора, недиагональна, что делает необходимым использование аппарата многомерного статистического анализа [5].

Если бы Q была известна, можно было бы осуществить переход к некоррелированным величинам. Действительно, поскольку y предполагается невырожденной n -мерной с. в., то Q — положительно определенная симметричная матрица, для которой спра-

ведлива факторизация Холецкого $Q=FF^T$, F — нижняя треугольная матрица, $\det F \neq 0$, так что существует F^{-1} . Перейдя к $\vec{z}=F^{-1}\vec{y}$, получим вектор с некоррелированными компонентами, поскольку $\text{cov}(\vec{z}, \vec{z}) = =F^{-1}Q(F^{-1})^T=I$ [5; 87].

Однако обычно Q неизвестна. Доказано [5; 87], что оценкой максимального правдоподобия для нее служит выборочная ковариационная матрица

$$A = \frac{1}{N} \sum_{j=1}^N (\vec{y}_j - \vec{\bar{y}})(\vec{y}_j - \vec{\bar{y}})^T;$$

здесь N — объем выборки, $\vec{\bar{y}} = \frac{1}{N} \sum_{j=1}^N \vec{y}_j$ — выбо-

рочное среднее. Преобразования \vec{y} , переводящие его в вектор с некоррелированными компонентами, рассматриваются в методе главных компонент и существенно используют матрицу A [3; 5].

Однако размерности измеряемых векторов в экспериментах нередко таковы, что число подлежащих оценке элементов ковариационной матрицы $n(n+1)/2$ порядка числа реализаций \vec{y} , а это влечет за собой низкую точность оценок. Использование A как оценки Q в этом случае может приводить к грубым ошибкам при анализе наблюдений, проверке гипотез, восстановлении решения методом наименьших квадратов и т. д.

Между тем сама специфика задачи подсказывает путь уменьшения числа оцениваемых параметров. Ошибки, связанные с аппаратным дрейфом, обусловлены действием небольшого числа причин (например, колебаний напряжения питания, нагрева аппаратуры и т. д.); сами изучаемые величины, как правило, некоррелированы при различных t_i . Поэтому в качестве модели измеряемых величин можно предложить

$$y^{(i)} = \sum_{r=1}^{r_0} L_{ir} x^{(r)} + \epsilon^{(i)}, \quad (1)$$

где i — номер компоненты вектора \vec{y} , r — номер причины (фактора), $\vec{\epsilon}$ — составляющая \vec{y} , связанная с самим изучаемым эффектом, $\text{cov}(\epsilon^{(i)}, \epsilon^{(j)}) = \sigma_i^2 \delta_{ij}$, $L = \|L_{ir}\|$ — неизвестная матрица коэффициентов. В такой постановке (1) совпадает с основной моделью факторного анализа [41; 54; 97; 101].

Факторный анализ, первоначально развивавшийся как средство обработки опытов в психологии, в последнее время широко используется для описания процессов в металлургии, в химии и т. д. Основное предположение факторного анализа состоит в том, что ковариационная матрица с. в. \vec{y} представляется как сумма диагональной матрицы и матрицы меньшего ранга k ($k < n$):

$$Q = LL^T + V, \quad (2)$$

а для самой \vec{y} справедливо представление

$$\vec{y} = L\vec{f} + \vec{\epsilon} + \vec{a} \quad (E\vec{y} = \vec{a}), \quad (3)$$

где $L = \|L_{ir}\|$ — матрица размера $n \times k$ и ранга k , $V = \text{diag}\{v_1, \dots, v_n\}$, \vec{f} — k -мерный случайный вектор, $\vec{\epsilon}$ — n -мерный вектор с некоррелированными компонентами; \vec{f} и $\vec{\epsilon}$ статистически независимы. В дальнейшем предполагается нормальность всех векторов.

Подобная задача возникает, например, когда нужно объяснить сложные результаты психологических тестов действием некоторых общих для всех людей причин (память, внимательность, быстрота реакции и т. д.); то, что остается кроме этого, характеризует личность испытуемого. При этом значения элементов матрицы L_{ir} характеризуют интенсивность r -го фактора для i -го индивидуума.

Вектор \vec{f} предполагается распределенным $N(0, I_k)$ и называется вектором простых факторов, $\vec{\epsilon} \sim N(0, V)$, L — матрица факторных нагрузок. В факторном анализе определению по наблюдениям y_1, \dots, y_n подлежат факторные нагрузки L_{ir} и остаточные дисперсии v_i . Заметим, что есть неоднозначность в выборе L (и факторов \vec{f}): если C — ортогональная матрица,

то, взяв $L' = LC$ и $\vec{f}' = C^T \vec{f}$, мы не изменим ни (2), ни (3), которые являются инвариантами ортогонального преобразования. Поэтому, получив в каком-то базисе факторные нагрузки L_{ir} , мы можем осуществить вращение, чтобы разумно интерпретировать результаты. Выбор C во многом субъективен, что долгое время препятствовало признанию факторного анализа со стороны статистиков.

Для нахождения L и V имеется много различных методов, различающихся эффективностью и простотой реализации [97]. Рассмотрим нахождение оценок на примере метода МП [54], который дает эффективные оценки, но громоздок при решении практических задач и обычно применяется на мощных ЭВМ.

В условиях нормальности логарифм ФП записывается в виде (\vec{a} предполагается известным)

$$F = \text{const} - \frac{N}{2} \ln |Q| - \frac{N}{2} \sum_{i,j} A_{ij} q^{ij} = \\ = \text{const} - \frac{N}{2} \ln |Q| - \frac{N}{2} \text{tr} (AQ^{-1}).$$

Максимизируем F по L_{ir} и v_i . Решение ищется в таком базисе, чтобы $J = L^T V^{-1} L$ была диагональна. Кроме того, потребуем, чтобы диагональные элементы J были расположены в порядке убывания; это фиксирует L с точностью до выбора знака.

Частная производная от $\ln |Q|$ по L_{ir} запишется как

$$2 \sum_j L_{jr} Q_{ij} / |Q| = 2 \sum_j L_{jr} q^{ij},$$

где Q_{ij} — алгебраическое дополнение q_{ij} в $|Q|$. Аналогично

$$\frac{\partial \ln |Q|}{\partial v_i} = \frac{Q_{ii}}{|Q|} = q^{ii}.$$

Сумма в F может быть записана как $\sum_{u,v} A_{uv} Q_{uv} / |Q|$, и для нее частная производная по L_{ir} имеет вид \int

$$2 \sum_{\substack{u, w, j, \\ u \neq i, w \neq j}} A_{uw} L_{jr} Q_{uw, ij} / |Q| - \\ - \sum_{u, w} A_{uw} Q_{uw} \cdot 2 \sum_j L_{jr} Q_{ij} / |Q|^2,$$

где $Q_{uw, ij}$ — алгебраическое дополнение q_{ij} в Q_{uw} . По теореме Якоби $Q_{uw, ij} = Q_{uw} Q_{ij} - Q_{uj} Q_{iw}$, откуда

$$-2 \sum_{u, w, j} A_{uw} L_{jr} Q_{uj} Q_{iw} / |Q|^2 = -2 \sum_{u, w, j} L_{jr} q^{ju} A_{uw} q^{wi}.$$

Частная производная суммы по v_i в F записывается в виде

$$- \sum_{u, w} q^{iu} A_{uw} q^{wi}.$$

Видно, что

$$\frac{\partial F}{\partial L_{ir}} = -N \left(\sum_j L_{jr} q^{ji} - \sum_{j, u, w} L_{jr} q^{ju} A_{uw} q^{wi} \right),$$

а это — элемент r -строки и j -го столбца матрицы $L^T Q^{-1} - L^T Q^{-1} A Q^{-1}$, умноженный на $-N$, в то время

как $\frac{\partial F}{\partial v_i} = -\frac{N}{2} \left(q_{ii} - \sum_{u, w} q^{iu} A_{uw} q^{wi} \right)$, т. е. диагональ-

ному элементу матрицы $Q^{-1} - Q^{-1} A Q^{-1}$, умноженному на $-N/2$.

Приравнявая производные 0, получим систему уравнений МП

$$L^T Q^{-1} - L^T Q^{-1} A Q^{-1} = 0, \quad (4)$$

$$\text{diag}\{Q^{-1} - Q^{-1} A Q^{-1}\} = 0. \quad (5)$$

Преобразуем (4), (5) к более удобному для решения виду. Умножив (4) справа на Q , получим

$$L^T - L^T Q^{-1} A = 0. \quad (6)$$

Умножив (5) слева на диагональную матрицу $V = Q - LL^T$ и используя (6), получим

$$\text{diag}\{I - A Q^{-1}\} = 0. \quad (7)$$

Умножив (7) справа на V и используя (6), придем к соотношению

$$\text{diag}\{Q-A\}=0. \quad (7')$$

Отсюда следует, что оценкой для q_{ii} является A_{ii} и

$$\hat{\sigma}_i = A_{ii} - \sum_{r=1}^k \hat{L}_{ir}^2. \quad (8)$$

Далее воспользуемся тождеством

$$Q^{-1} = V^{-1} - V^{-1}L(I+J)^{-1}L^T V^{-1}, \quad (9)$$

которое проверяется умножением его на $Q = LL^T + V$. Умножение (9) слева на L^T приводит к

$$L^T Q^{-1} = (I+J)^{-1} L^T V^{-1}. \quad (10)$$

Подставив (10) в (6), получим $L^T = (I+J)^{-1} L^T V^{-1} A$, что эквивалентно равенству $L^T = J^{-1}(L^T V^{-1} A - L^T)$, или

$$L^T = J^{-1} L^T V^{-1} (A - V). \quad (11)$$

Из (11) следует, что $H = L^T V^{-1} (A - V) V^{-1} L = J^2$ и, следовательно, диагональна.

Элементы матрицы I в (11) являются собственными значениями для $V^{-1}(A - V)$, а строки L^T — собственными вектор-строками, т. е., задавшись начальными приближениями для L и V , можно организовать итерационный процесс. Из (11) находится следующее приближение для L как решение задачи на собственные значения; оно подставляется в (8) для определения следующего приближения V и т. д. Найдя при фиксированном k матрицы L и V , можно рассмотреть критерии для проверки гипотезы о том, что k факторов описывают экспериментальные данные; если это не так, то надо увеличить k , и т. д.

Вернемся к задаче [77; 89] оценивания ковариационной матрицы Q , о которой шла речь выше. Будем предполагать, что с.в. y задана на декартовом произведении пространств $\Omega = \Omega_1 \times \Omega_2$, где Ω_1 описывает вероятностный характер самого изучаемого процесса, Ω_2 — стохастическую природу аппаратного дрейфа; элементы $\omega_2 \in \Omega_2$ — функции времени. Будем предполагать, что при фиксированных параметрах

аппаратуры измеряемая величина обладает независимыми компонентами, т. е. вектор $\vec{y}(\omega_1|\omega_2)$ имеет некоррелированные компоненты. Обозначая условное м. о. (при заданном ω_2) через E_{ω_2} , а дисперсию — через D_{ω_2} , получим

$$E(y_i y_j) = \int_{\omega_2} E_{\omega_2}(y_i y_j) dP_{\omega_2} = \overline{E_{\omega_2}(y_i y_j)} = \overline{E_{\omega_2} y_i E_{\omega_2} y_j},$$

где обозначено \bar{z} — усреднение z по мере P_{ω_2} .

Используя независимость компонент $\vec{y}(\omega_1|\omega_2)$, элементы Q можно представить в виде

$$q_{ij} = \overline{E_{\omega_2} y_i E_{\omega_2} y_j} - \overline{E_{\omega_2} y_i} \overline{E_{\omega_2} y_j} \quad (i \neq j),$$

$$q_{ii} = \overline{D_{\omega_2} y_i} + D(\overline{E_{\omega_2} y_i}).$$

Запишем $E_{\omega_2} y_i$ как $E_{\omega_2} y_i = a_i x_i$, выделив в явном виде математическое ожидание с. в. y_i . Тогда случайная величина $x_i = x_i(\omega_2)$ характеризует случайные колебания параметров аппаратуры относительно средних величин; заметим, что $\bar{x}_i = 1$ ($i = 1, n$). Предположим, что x нормальна. Подставляя $E_{\omega_2} y_i$ в q_{ij} , получим $q_{ij} = a_i a_j \text{cov}(x_i, x_j)$, $i \neq j$, $q_{ii} = D_{\omega_2} y_i + a_i^2 D x_i$, т. е. дисперсия y_i распадается на 2 слагаемых, первое из которых описывает статистические ошибки измерений, а второе (а также недиагональные члены Q) характеризует влияние дрейфа на точность измерений.

Тогда для \vec{y} имеем представление

$$\vec{y} = \vec{a} + \Lambda \vec{x}_0 + \vec{\varepsilon}, \quad (12)$$

где $\Lambda = \text{diag}\{a_1, \dots, a_n\}$, \vec{x}_0 и $\vec{\varepsilon}$ — независимые n -мерные нормальные величины: $\vec{x}_0 = \vec{x} - \vec{1} \sim N(\vec{0}, X)$, $X = \text{cov}(\vec{x}, \vec{x})$, $\vec{\varepsilon} \sim N(\vec{0}, V)$, $V = \text{diag}\{\overline{D_{\omega_2} y_i}\}$. Справедливость (12) следует из того, что слева и справа в нем стоят величины, распределенные $N(\vec{a}, Q)$. Параметризация (12) пока что не облегчила задачу, так как с. в. \vec{x}_0 и $\vec{\varepsilon}$ по-прежнему имеют размерность n .

Используем тот факт, что дрейф происходит медленнее, чем пробегание сетки. Будем предполагать, что с вероятностью, близкой к 1, произвольная реализация \vec{x}_0 с. в. x_0 является медленно меняющейся функцией аргумента i . Поэтому \vec{x}_0 может быть аппроксимирована конечным числом членов разложения по некоторой ортогональной системе функций $\{\varphi_i^{(s)}\}$:

$$x_{0i} = \sum_{s=1}^k u_s \varphi_i^{(s)},$$

$k < n$, или в векторном виде

$$\vec{x}_0 = \Phi \vec{u} \quad (\Phi = \|\varphi_i^{(s)}\|, \quad i = \overline{1, n}; \quad s = \overline{1, k}), \quad (13)$$

где \vec{u} — k -мерная нормальная с. в. с невырожденной ковариационной матрицей. Как было указано выше, \vec{u} может быть преобразована к нормальному вектору \vec{j} с некоррелированными компонентами. Потребуем, чтобы $\vec{j} \sim N(0, I_k)$. Тогда получим для \vec{y} факторное представление

$$\vec{y} = \vec{a} + L\vec{j} + \vec{\epsilon} = \vec{a} + \Lambda \Phi H \vec{j} + \vec{\epsilon}, \quad (14)$$

где H — матрица размера $k \times k$, осуществляющая невырожденное преобразование $\vec{u} = H\vec{j}$, \vec{j} и $\vec{\epsilon}$ независимы. Для ковариационной матрицы Q получим искомую формулу

$$Q = LL^T + V = \Lambda \Phi H H^T \Phi^T \Lambda + V. \quad (15)$$

Поскольку Λ и матрица Φ базисных функций предполагаются известными, для определения Q нужно найти $k^2 + n$ неизвестных параметров — элементов H и V . Если $k \ll n$, то это позволяет существенно уменьшить число оцениваемых параметров. Заметим, что присущая факторному анализу неоднозначность определения матрицы L факторных нагрузок с точностью до правого умножения на ортогональную матрицу C для нас не играет роли, поскольку Q инвариантна относительно таких вращений.

Применяя метод МП подобно рассмотренному выше, получим систему

$$L^T Q^{-1} - L^T Q^{-1} A L = 0, \quad (16)$$

$$\text{diag}\{Q^{-1} - Q^{-1} A Q^{-1}\} = 0. \quad (17)$$

Используя базис, в котором диагональная матрица

$$J = L^T V^{-1} L = H^T \Phi^T \Lambda V^{-1} \Lambda \Phi H, \quad (18)$$

приведем систему (16), (17) к виду, более удобному для численного решения. Получим после ряда преобразований

$$J H^T = H^T R, \quad (19)$$

$$V = \text{diag}\{A + L L^T - 2 L L^T Q^{-1} A\}, \quad (20)$$

где

$$R = \Phi^T \Lambda V^{-1} (A - V) V^{-1} \Lambda \Phi (\Phi^T \Lambda V^{-1} \Lambda \Phi)^{-1}.$$

Система (19), (20) может быть решена итерационными методами.

В важном частном случае регистрации почти пуассоновских событий имеет место соотношение $D_{\omega_i} y_i = \mu E_{\omega_i} y_i, i = \overline{1, n}$, где $\mu > 0$ — постоянная, близкая к 1, описывающая отклонения от пуассоновости. Тогда $V = \mu \Lambda$ и вместо (20) имеем кубическое по μ уравнение.

Глава 13

УЧЕТ СИСТЕМАТИЧЕСКИХ ИСКАЖЕНИЙ

Помимо случайных ошибок и грубых сбоев, всякий эксперимент несвободен от систематических погрешностей: фоновых подложек, просчетов регистрирующей аппаратуры. Поэтому для получения несмещенных оценок измеряемых величин нужно учесть систематические искажения. Рассмотрим, как решаются эти задачи, на примере фотоядерного эксперимента [20].

Измеряется выход (среднее число продуктов реакции) в зависимости от энергии: $Y = Y(t), t = \langle t_i \rangle, i = \overline{1, n}$. Исправление систематических погрешностей

сводится к устранению просчетов, учету фоновой подложки, нормировке на дозу облучения. Процессы, обуславливающие эти искажения, весьма сложны. Например, в случае фоновой подложки механизм появления фоновых излучений — тот же, что и для изучаемого процесса (т. е. неизвестен до конца). Построение достаточно полных физических, а затем и математических моделей всех этих процессов требует больших усилий, а корректный учет величин, входящих в такие модели, потребовал бы количества измерений, соизмеримого с числом наблюдений изучаемого эффекта. Поэтому стараются проводить эксперимент так, чтобы влияние этих искажений на результат было небольшим (например, чтобы отношение фон/излучаемый эффект было малым), выбирают достаточно простые, но удовлетворительно описывающие характер искажений модели, а сопутствующие измерения, необходимые для определения параметров в этих моделях, ведут на «малых» (по сравнению с числом наблюдений в основном эксперименте) статистиках.

В случае фотоядерного эксперимента исправление систематических погрешностей сводится к устранению просчетов, учету фоновой подложки, нормировке на дозу облучения по формуле

$$\hat{Y}(t) = (Y^{(n)}(t) - \Phi^{(n,c)}(t)) F(t) / D^{(c)}(t), \quad (1)$$

где $\hat{Y}(t)$ — кривая выхода после всех исправлений, $Y^{(n)}(t)$ — исправленная на просчеты кривая выхода, $\Phi^{(n,c)}(t)$ — нормированный и сглаженный фон, $D^{(c)}(t)$ — сглаженная функция дозы облучения, $F(t)$ — чувствительность дозиметра (известная функция).

1. Возможны просчеты, связанные с наличием «мертвого времени» τ у регистрирующей аппаратуры. Скажем, попадание частицы в счетчик Гейгера отмечается прохождением электрического тока через трубку с инертным газом. В готовом к работе счетчике на электроды, разделенные промежутком инертного газа, подано напряжение, близкое к напряжению пробоя этого слоя газа. При появлении внутри трубки заряженной частицы высокой энергии происходит ионизация атомов газа и развивается электрический разряд, который фиксируется соответствующим измерительным прибором. Пока ионизованные частицы не

достигнут электродов и не передадут свой заряд, ток в цепи продолжается. Если в этот промежуток времени в счетчик попадет еще одна частица, ее появление не будет зафиксировано. Аналогичные явления, связанные с появлением «мертвого времени», возникают и при работе детекторов других типов, например сцинтилляционных.

Существует много различных моделей для описания просчетов детекторов [27]. Будем предполагать, что средний промежуток времени между двумя последовательными отсчетами много больше «мертвого времени», так что вероятность просчетов мала. Тогда просчеты могут учитываться по формуле

$$EY^{(n)}(t) = EY(t) + \alpha(EY(t))^2, \quad (2)$$

где α зависит от «мертвого времени» τ регистрирующего тракта, времени измерений и характера изменения интенсивности счета со временем. При обработке эксперимента исправления производились по формуле

$$Y^{(n)}(t) = Y(t) + \alpha(Y^{(c)}(t))^2, \quad (2')$$

где $Y^{(c)}(t)$ — сглаженная кривая выхода. Коэффициент α либо считается известным (определяется из условий эксперимента), либо оценивается из сравнения $Y(t)$ с эталонной кривой $Y_M(t)$, построенной при малых нагрузках регистрирующей аппаратуры, когда просчетами можно пренебречь. Относительно $Y_M(t)$ и $Y(t)$ предполагается, что они распределены по нормальному закону с м. о. $EY(t_i) = a_i$,

$$EY_M(t_i) = r_1(EY^{(n)}(t_i) + c_1) = u_1 a_i^2 + u_2 a_i + u_3, \quad (3)$$

где $u_1 = r_1 \alpha$, $u_2 = r_1$, $u_3 = r_1 c_1$. Коэффициент r_1 учитывает неодинаковость доз облучения для Y и Y_M , c_1 характеризует космический фон. Обычно для построения кривой Y_M берется малая статистика (малая доза облучения, $r_1 \approx 0,01 \div 0,04$), поскольку ищутся небольшие поправки на просчеты и α можно определять с невысокой точностью. Предполагается, что измерения при различных энергиях некоррелированы, так что $\text{cov}(Y_M(t_i), Y_M(t_j)) = DY_M(t_i) \delta_{ij}$, $\text{cov}(Y(t_i), Y(t_j)) = DY(t_i) \delta_{ij}$ (для кривой выхода $Y(t)$ в основном эксперименте учет коррелированности компонент и переход

ж некоррелированным величинам может быть выполнен предварительно так, как описано выше в гл. 12), $\text{cov}(Y(t_i), Y_M(t_j)) = 0$.

Воспользуемся методом МП для определения параметров модели (3). Имеем функцию правдоподобия для совокупности измерений $Y_i = Y(t_i)$, $Y_{(M)i} = Y_M(t_i)$, $i = \overline{1, n}$:

$$L = \text{const} \cdot \exp \left\{ - \sum_{i=1}^n \frac{(Y_i - a_i)^2}{2DY_i} - \sum_{i=1}^n \frac{(Y_{(M)i} - u_1 a_i^2 - u_2 a_i - u_3)^2}{2DY_{(M)i}} \right\}. \quad (4)$$

Как обычно в методе МП, ищем оценки для a_i , u_1 , u_2 , u_3 , максимизируя L . Если бы a_i оценивались лишь по кривой $Y(t)$, вторая сумма в экспоненте исчезала бы и оценки максимального правдоподобия равнялись Y_i . Поскольку кривая $Y_M(t)$ снята с малой статистикой, использование совокупности измерений для оценивания a_i несущественно улучшит эти измерения. В то же время зависимость от совокупности u_1 , u_2 , u_3 , $\{a_i\}$ во второй сумме нелинейная, что приводит к нелинейности уравнений МП и необходимости организации итерационного процесса для их решения. Поэтому будем считать, что м. о. оценивается как $\tilde{a}_i = Y_i$, и решать задачу оценивания только u_1 , u_2 , u_3 . Имеем

$$\sum_{k=1}^s A_{sk} \tilde{u}_k = b_s, \quad s = \overline{1, 3},$$

где

$$A_{sk} = \sum \omega_i Y_i^{6-s-k}, \quad b_s = \sum \omega_i Y_{(M)i} Y_i^{3-s}, \quad \omega_i = (DY_{(M)i})^{-1}. \quad (5)$$

Поскольку регистрируемые продукты ядерных реакций имеют распределение Пуассона (предполагаем, что энергия t ниже порога реакции $(\gamma, 2n)$, либо вклад этой реакции в Y мал [8; 18]), для которого,

если $\xi \sim Po(\mu)$, то $D\xi = E\xi = \mu$, весовые множители ω_i равны $(EY_{(M)i})^{-1}$. Для используемых статистик отсчетов $EY_{(M)i} \gg 1$. Поэтому допустимо считать $EY_{(M)i} \approx Y_{(M)i}$, так что окончательно в (5) $\omega_i \approx Y_{(M)i}^{-1}$. Найдя из (5) $\bar{u}_1, \bar{u}_2, \bar{u}_3$, определим $\alpha = \bar{u}_1/\bar{u}_2$ и будем использовать его в (2') для учета просчетов.

2. Фон в основном обусловлен реакциями при взаимодействии пучка γ -квантов с коллиматором ускорителя, контейнером образца и предметами в зале. Кроме того, нужно учитывать фон космического излучения.

Кривая выхода фотоядерного эксперимента представляет собой монотонно возрастающую функцию энергии, по виду в целом напоминает экспоненту или ветвь параболы. Структурные особенности сечения, проявляющиеся в выходе в соответствии с (4) из гл. 1, приводят в нем к появлению особенностей типа изломов кривой или перегибов. Поэтому наличие фона обуславливает некоторое (вообще говоря, неравномерное) поднятие «идеальной» (без фона) кривой выхода.

Для оценки фоновой подложки проводятся специальные измерения фона: снимается кривая выхода без образца $\Phi_i, i=1, n$. На кривой выхода $Y(t)$ имеется участок $i=1, i_0$, отвечающий энергиям $t_i < t_0$, где t_0 — пороговая энергия реакции (γ, n); отсчеты на этом участке целиком обусловлены фоном. Поскольку кривая измеряется со статистикой, в десятки и сотни раз меньшей, чем Y_i , для учета вклада фона в выход $Y(t)$ нужно провести нормировку Φ_i к соответствующей статистике (дозе облучения). Для этого используются начальные участки кривых Y_i и Φ_i , а также специальные вспомогательные измерения доз облучения. Для лучшей точности объединим оба способа нормировки.

Результатами измерений являются $\Phi_i, Y_i (i=1, i_0)$, отношение доз облучения \bar{r} для кривых Φ_i и Y_i , полученное в дозиметрических измерениях; пусть $\bar{D}\bar{r}$ — оценка дисперсии \bar{r} (например, ее выборочная дисперсия). ФП объединенной совокупности измерений в тех же предположениях о Y, Φ и \bar{r} , что и в п. 1, имеет вид

$$L_{\Phi} = \text{const} \cdot \exp \left\{ - \sum_{i=1}^{i_0} \frac{(Y_i - a_i)^2}{2DY_i} - \sum_{i=1}^{i_0} \frac{(\Phi_i - ra_i - c_2)^2}{2D\Phi_i} - \frac{(\widehat{r} - r)^2}{2D\widehat{r}} \right\}; \quad (6)$$

здесь $r = E\hat{r}$, модель фона имеет вид $E\Phi_i = ra_i + c_2$, где c_2 учитывает космический фон и набросы аппаратуры. Аналогично анализу ФП (4), учитывая «малость» статистики Φ_i по сравнению с Y_i , оценим $\bar{a}_i = Y_i$. Для определения \widehat{r} и \bar{c}_2 имеем СЛАУ

$$\begin{aligned} \widehat{r} \left(\sum_{i=1}^{i_0} \omega_i Y_i^2 + \frac{1}{D\widehat{r}} \right) + \bar{c}_2 \sum_{i=1}^{i_0} \omega_i Y_i &= \sum_{i=1}^{i_0} \omega_i Y_i \Phi_i + \frac{\widehat{r}}{D\widehat{r}}, \\ \widehat{r} \sum_{i=1}^{i_0} \omega_i Y_i + \bar{c}_2 \sum_{i=1}^{i_0} \omega_i &= \sum_{i=1}^{i_0} \omega_i \Phi_i, \end{aligned} \quad (7)$$

где $\omega_i = (D\Phi_i)^{-1} \approx \Phi_i^{-1}$, $D\widehat{r} \approx D\bar{r}$.

Разрешив (7) относительно \widehat{r} и \bar{c}_2 , найдем нормированный на дозу облучения кривой Y_i фон

$$\Phi_i^{(H)} = \frac{\Phi_i - \bar{c}_2}{\widehat{r}}, \quad i = \overline{1, n}. \quad (8)$$

Чтобы уменьшить статистические флуктуации в $\Phi^{(H)}(t)$, как и в квадратичном члене формулы (2'), обычно проводят сглаживание этих кривых с помощью процедур типа скользящего среднего [98] или скользящих медиан [61].

В других типах экспериментов, скажем при регистрации гамма-спектров, учет фоновой подложки удастся осуществить и без проведения дополнительных измерений. Типичная картина экспериментальных кривых в экспериментах такого класса представлена на рис. 4. Ниспадающий фон (штриховая линия) является «пьедесталом» для спектральной кривой $Y(t)$ с максимумами и минимумами. Нередко математическая модель, описывающая такой фон, достаточно проста, например линейная или экспоненциальная. Параметры модели легко определить, построив по методу НК соответствующую аппроксимацию через точ-

ки локальных минимумов $Y(t)$. Далее аппроксимированный фон вычитается из $Y(t)$.

3. Функция дозы входит в формулу (1) для $Y(t)$ как множитель и существенно влияет на точность \hat{Y} . В то же время она измеряется с меньшей точностью, чем $Y(t)$, и на более редкой сетке, так что наряду с ее сглаживанием требуется интерполяция $D(t)$ на промежуточные значения аргумента.

Кривая $D(t)$ имеет весьма простую структуру (монотонное возрастание, отсутствие перегибов), напоминающая экспоненту. Поэтому ее можно аппроксимировать

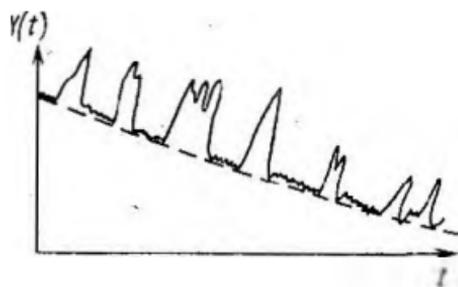


Рис. 4

подходящей нелинейной по параметрам функцией (гамма-функцией, экспонентой, причем для экспоненты, по крайней мере для начальной оценки, задача сводится к линейному МНК, как отмечено в гл. 6, 8) либо искать подходящую линейную аппроксимацию, например поли-

номом небольшой (не выше 5-й, 6-й) степени. Для построения аппроксимации в последнем случае будем использовать линейный МНК.

Однако если степень полинома выше второй, аппроксимация $D(t)$ может содержать перегибы в изучаемой области $t_i, i=1, n$. Эти особенности, связанные со статистическими флуктуациями $D(t)$ и выбором по МНК коэффициентов аппроксимирующего полинома, могут привести к появлению ложной структуры в регуляризованном решении уравнения (4) из гл. 1. Поэтому наложим ограничения на класс решений: будем искать полином степени κ $\mathcal{P}_\kappa(t)$, наименее удаленный от сглаживаемой функции $D(t)$ и не имеющий перегибов на $[t_1, t_n]$. Тогда $\mathcal{P}_\kappa(t)$ совпадает с одной из функций $\mathcal{P}_\kappa^{(0)}(t), \mathcal{P}_\kappa^{(1)}(t), \dots, \mathcal{P}_\kappa^{(\kappa-2)}(t)$, где $\mathcal{P}_\kappa^{(0)}(t)$ — полином МНК, проведенный без ограничений; $\mathcal{P}_\kappa^{(1)}(t)$ — аналогичный полином при условии, что $\mathcal{P}_\kappa^{(1)}(t)$ имеет нуль 1-го порядка на границе $[t_1, t_n], \dots, \mathcal{P}_\kappa^{(\kappa-2)}(t)$ — полином степени κ , проведенный при условии, что $\mathcal{P}_\kappa^{(\kappa-2)}(t)$ имеет нуль $(\kappa-2)$ -го порядка на:

границе этой области. Эти условия приводят к дополнительным связям параметров полинома. Так если

$$\mathcal{P}_x(t) = u_0 + u_1 t + u_2 t^2 + \dots + u_x t^x,$$

то условие равенства $\mathcal{P}_x^*(t)$ нулю при $t=t_1$ (или $t=t_n$) приводит к связи

$$2u_2 + 6u_3 t_1 + \dots + x(x-1)u_x t_1^{x-2} = 0, \quad (9)$$

откуда $u_2 = -\frac{1}{2} (6u_3 t_1 + 12u_4 t_1^2 + \dots + x(x-1)u_x t_1^{x-2})$ (аналогично для $t=t_n$). Нули более высокой кратности приводят к рассмотрению наряду с условиями типа (9) также аналогичных условий на более высокие производные $\mathcal{P}_x(t)$ в точках $t=t_1$ и $t=t_n$. Далее задача решается перебором: из совокупности $\mathcal{P}_x^{(0)}(t)$, $\mathcal{P}_x^{(1)}(t)$, ..., $\mathcal{P}_x^{(x-2)}(t)$ оставляем полиномы, не имеющие перегибов внутри $[t_1, t_n]$ (хотя бы один, $\mathcal{P}_x^{(x-2)}(t)$, заведомо удовлетворяет этому требованию), затем из них выберем полином, дающий наименьшую величину RSS .

Другой возможный путь решения подобных задач линейного МНК с ограничениями на производные аппроксимирующей функции рассмотрен в [123].

Глава 14

ПРИМЕНЕНИЕ РЕКУРРЕНТНОГО ОЦЕНИВАНИЯ И ФИЛЬТРАЦИИ ПРИ ОБРАБОТКЕ ДАННЫХ

1. Как мы уже отмечали в гл. 1, традиционные алгоритмы статистической обработки, оперирующие со всей совокупностью данных y_1, y_2, \dots, y_n после их регистрации, не всегда приемлемы. Они требуют больших затрат памяти, если объем обрабатываемой информации n велик. Кроме того, если результаты статистической обработки используются для оперативного контроля и управления процессом, алгоритмы off-line-обработки, многократно применяемые к расширяющейся совокупности данных $\{y_i\}$, зачастую не обеспечивают необходимого быстрого действия, сравнимого со скоростью протекания процесса и поступления но-

вой информации. На примере вычисления выборочно-среднего по формуле (3) из гл. 1 была описана основная идея, используемая в алгоритмах on-line-обработки: оценка на l -м шаге ищется через внесение поправки в оценку, вычисленную на $(l-1)$ -м шаге; в случае линейной оценки, как в (3) из гл. 1, поправка просто добавляется к ранее найденной оценке, причем величина поправки зависит лишь от вновь поступившего наблюдения y_l и пропорциональна невязке (разности между измеренным y_l и результатом его предсказания исходя из вычисленной на $(l-1)$ -м шаге оценки). В подобной рекуррентной процедуре каждое наблюдение участвует в вычислительном процессе только один раз, что позволяет существенно увеличить скорость обработки и сократить требуемый объем памяти.

Не претендуя на сколько-нибудь полное и систематическое изложение (читатели могут исчерпывающе ознакомиться с этим кругом вопросов в статьях и монографиях по рекуррентному оцениванию, фильтрации и стохастической аппроксимации), рассмотрим, следуя [14], обобщение модели линейного МНК (5) из гл. 6. Оно приводит к процедуре рекуррентного оценивания, известной как фильтр Калмана и широко применяемой при обработке данных. Пусть изучаемый процесс характеризуется n -мерным вектором состояний x , являющимся функцией времени, информация поступает в дискретные моменты времени, $t_0, t_1, \dots, t, \dots$, так что $x = x(t)$. Предполагается, что вектор x является случайной величиной и эволюционирует от измерения в момент времени t к измерению в момент времени $t+1$ (здесь под $t+1$ имеется в виду номер наблюдения $t+1 \doteq t_{n+1}$ при $t \doteq t_n$) по линейному закону

$$x(t+1) = A(t)x(t) + v(t), \quad (1)$$

где $A(t)$ — известная матрица, возможно, меняющаяся со временем, $v(t)$ — процесс типа белого шума с дискретным временем [87], причем

$$Ev(t) = 0, \quad Ev(t)v^T(t) = Q(t)\delta_{tt}, \quad (2)$$

$Q(t)$ — известная матрица, δ_{tt} — символ Кронекера.

Непосредственно вектор x не регистрируется, а измеряется косвенно путем регистрации вектора наблю-

дений $y(t)$, связь которого с вектором состояний $x(t)$ также предполагается линейной:

$$y(t) = C(t)x(t) + w(t), \quad (3)$$

размерность $y(t)$ равна m , $C(t)$ — известная, зависящая от времени прямоугольная матрица, $w(t)$ — процесс типа белого шума с характеристиками

$$Ew(t) = 0, \quad Ew(t)w^T(\tau) = R(t)\delta_{t\tau}, \quad Ew(t)v^T(t) = 0. \quad (4)$$

Имея наблюдения $y(t_0), \dots, y(t_s), \dots, y(t)$, можно попытаться оценить вектор $x(t)$ в момент времени t_i . Ввиду линейности модели (1), (3) оценку $x(t_i)$ будем искать в классе линейных несмещенных оценок с минимальной дисперсией, т. е. минимизирующих квадратичный функционал

$$J = E(x(t_i) - \hat{x}(t_i))^2. \quad (5)$$

Момент времени t_i может принадлежать промежутку наблюдений $[t_0, t]$ (задача интерполяции, или сглаживания), не принадлежать ему (задача экстраполяции, или прогноза), совпадать с ним. Нас интересует задача фильтрации — задача экстраполяции на шаг вперед: $t_i = t + h_i \triangleq t_1$, где h_i — интервал времени между последовательными измерениями t и следующим за ним. Как мы условились выше и как это принято в теории фильтрации, обозначим $t+1 \triangleq t_1$.

Сделаем некоторые дополнительные предположения о векторе состояний $x(t)$, часть из которых далее снимем. Пусть его начальное значение $x(t_0)$ есть n -мерный случайный вектор с нулевым м.о.: $E x(t_0) = 0$. В силу (1) и (2) имеем

$$E x(t) = 0. \quad (6)$$

Пусть задана ковариационная матрица

$$P(t_0) = E x(t_0)x^T(t_0). \quad (7)$$

Кроме того, предполагаем, что с.в. $x(t_0)$ не коррелирована с $v(t)$ и $w(t)$:

$$E x(t_0)v^T(t) = 0, \quad E x(t_0)w^T(t) = 0. \quad (8)$$

Из (3) и (4), очевидно, следует $E y(t) = 0$.

2. Получим необходимые для дальнейшего соотношения, справедливые для линейных несмещенных оце-

нок с минимальной дисперсией ошибок оценивания вектора x по косвенным измерениям в линейной модели

$$z = Dx + s, \quad (9)$$

где D — матрица размера $m \times n$, x — n -вектор, z — m -вектор наблюдения, s — случайная m -мерная по- меха. Вектор x — также с. в., причем

$$Ex = 0, \quad Exx^T = P, \quad (10)$$

$$Es = 0, \quad Ess^T = S, \quad (10')$$

$$Exs^T = 0, \quad (10'')$$

матрицы P и S известны. По наблюдениям с. в. y ищем оценку x , являющуюся: а) линейной, б) несмещенной, в) имеющей минимальную дисперсию ошибок оценивания. Для искомой оценки \hat{x} имеем в силу а) $\hat{x} = g + Gz$, где g — постоянный вектор, G — матрица. Условия (10), (10') приводят к $Ez = 0$, так что м. о. $E\hat{x} = Ex$ согласно б) дает

$$E\hat{x} = g + GEz = g = Ex = 0.$$

Итак, искомая оценка $\hat{x} = Gz$. Вид G найдем из условия в). Для i -й компоненты имеем $\hat{x}_i = g^i z$, где вектор-строка g^i есть i -я строка матрицы G . Тогда, рассматривая вектор ошибки оценивания $\tilde{x} = x - \hat{x}$, имеем для дисперсии ошибки оценивания,

$$D\tilde{x}^2 = \sum_{i=1}^n D\tilde{x}_i^2 = \sum_{i=1}^n E(x_i - \hat{x}_i)^2. \quad (11)$$

Минимальность $D\tilde{x}^2$ достигается, если каждое слагаемое в (11) минимально, т. е. нужно минимизировать по g^i выражение

$$\begin{aligned} E\tilde{x}_i^2 &= E(x_i - \hat{x}_i)^2 = E(x_i - g^i z)^2 = Ex_i^2 - \\ &- 2E(x_i z^T)(g^i)^T + g^i E(z z^T)(g^i)^T. \end{aligned} \quad (12)$$

Первое слагаемое в (12) не зависит от g^i . Необходимое условие экстремума $\partial/\partial(g^i)^T E\tilde{x}_i^2 = 0$ приводит к соотношению

$$Ex_i z^T - g^i Ezz^T = 0, \quad (13)$$

или в матричном виде

$$Exz^T - GEzz^T = 0. \quad (14)$$

Дифференцируя (13) по g^i , в силу достаточного условия минимума $\left\| \frac{\partial^2 E\bar{x}_i^2}{\partial g_i^i \partial g_j^j} \right\| > 0$ имеем матричное условие $Ezz^T > 0$, обеспечивающее минимум в точке, определяемой уравнением (14). Условие (14) — аналог уравнения Винера — Хопфа [14; 62] для случая дискретного времени. Преобразуя (14), имеем

$$0 = E(xz^T - Gzz^T) = E(x - Gz)z^T = E(\bar{x}z^T). \quad (15)$$

Умножив (15) справа на G^T , получим

$$0 = E(\bar{x}z^T)G^T = E(\bar{x}\hat{x}), \quad (16)$$

откуда

$$\begin{aligned} E(\bar{x}\bar{x}^T) &= E\bar{x}(x - \hat{x})^T = E\bar{x}x^T = E(x - Gz)x^T = \\ &= Exx^T - GEzx^T. \end{aligned} \quad (17)$$

Минимально возможная сумма дисперсий определяется следом этого выражения.

С учетом (9), (10) — (10'') имеем

$$Exz^T = E[x(x^T D^T + s^T)] = PD^T, \quad (18)$$

$$Ezz^T = E[(Dx + s)(x^T D^T + s^T)] = DPD^T + S. \quad (18')$$

Подставляя (18), (18') в (14), получим

$$G(DPD^T + S) = PD^T. \quad (19)$$

Если $DPD^T + S$ невырождена (для этого достаточно выполнения одного из условий: $S > 0$ либо $P > 0$, $n > m$ и $\text{rang } D = m$), то имеем

$$G = PD^T (DPD^T + S)^{-1}, \quad (20)$$

откуда получается оптимальная линейная несмещенная оценка с минимальной дисперсией ошибки

$$\hat{x} = PD^T (DPD^T + S)^{-1} z. \quad (21)$$

Матрица ковариации ошибки оценивания ввиду (17) равна

$$E\bar{x}\bar{x}^T = P - GDP, \quad (22)$$

или с учетом (20)

$$E\hat{x}\hat{x}^T = P - PD^T(DPD^T + S)^{-1}DP. \quad (23)$$

3. Вернемся теперь к нашей задаче оптимальной фильтрации. В первый момент t_0 имеем $y(t_0) = C(t_0)x(t_0) + \omega(t_0)$, причем $E x(t_0)x^T(t_0) = P(t_0)$ известна. Сравнивая это выражение с (9), (10), получим

$$\hat{x} = K(t_0)y(t_0), \quad (24)$$

где

$$K(t_0) = P(t_0)C^T(t_0)[C(t_0)P(t_0)C^T(t_0) + R(t_0)]^{-1}, \quad (25)$$

причем

$$\begin{aligned} \bar{P}(t_0) &\doteq E[(\hat{x}(t_0) - x(t_0))(\hat{x}(t_0) - x(t_0))^T] = \\ &= P(t_0) - K(t_0)C(t_0)P(t_0) \end{aligned} \quad (25')$$

согласно (22).

Используя дискретный аналог уравнения Винера—Хопфа (14) и (17), построим процедуру рекуррентного оценивания. Пусть наблюдения $y(t_0), \dots, y(t)$ уже использованы для построения линейной несмещенной оценки с минимальной дисперсией, т. е. оценки оптимальной в среднеквадратическом смысле. Это означает согласно (14), что

$$E[(x(t) - \hat{x}(t))(y^T(t_0), y^T(t_0+1), \dots, y^T(t))] = 0, \quad (26)$$

где $(y^T(t_0), \dots, y^T(t))$ — «длинный» вектор-строка, составленный из векторов-строк наблюдений $y^T(t_0), \dots, y^T(t)$. Пусть $x^*(t+1)$ — оптимальная оценка экстраполяции вектора x по результатам наблюдений $y(t_0); \dots, y(t)$, вычисленная в точке $t+1$. Согласно (14) для нее также

$$E[(x(t+1) - x^*(t+1))(y^T(t_0), \dots, y^T(t))] = 0. \quad (26')$$

Пусть проведено новое измерение $y(t+1)$; с его учетом можно построить оптимальную оценку $\hat{x}(t+1)$, причем согласно (14) для нее

$$\begin{aligned} E[(x(t+1) - \hat{x}(t+1))(y^T(t_0), \dots, y^T(t), y^T(t+1))] = \\ = 0. \end{aligned} \quad (26'')$$

Сравним (26'), которое должно выполняться соглас-

но условию оптимальности (14), с уже удовлетворенным по нашему рекуррентному построению условием (26). Используя (1), запишем в (26') $x(t+1) = A(t)x(t) + v(t)$. Поскольку $v(t)$ некоррелировано с $y(t_0), \dots, y(t)$ в силу (1), (2), (3), (4), (8), то имеем с учетом (26)

$$E[(A(t)x(t) - x^*(t+1))(y^T(t_0), y^T(t_0+1), \dots, y^T(t))] = 0,$$

которое в силу произвольности $y(t_0), \dots, y(t)$ выполнено при

$$x^*(t+1) = A(t)\hat{x}(t). \quad (27)$$

С учетом (1) соотношение (27) для оптимальной экстраполяции представляется вполне естественным (возмущения $v(t+1)$ носят случайный характер и не учитываются при построении прогноза $x^*(t+1)$).

Ошибка экстраполяции запишется в виде

$$x(t+1) - x^*(t+1) = A(t)x(t) + v(t) - A(t)\hat{x}(t) = A(t)\tilde{x}(t) + v(t), \quad (27')$$

а ее ковариационная матрица

$$P^*(t+1) = E[(x(t+1) - x^*(t+1))(x(t+1) - x^*(t+1))^T] = E[(A(t)\tilde{x}(t) + v(t))(A(t)\tilde{x}(t) + v(t))^T] = A(t)\bar{P}(t)A^T(t) + Q(t), \quad (28)$$

поскольку $E[v(t)\tilde{x}^T(t)] = 0$ в силу отмеченной некоррелированности $v(t)$ с $y(t_0), \dots, y(t)$ и согласно (1) и (2) с $x(t)$.

Заменяя в (27) $\hat{x}(t)$ на $x(t) - \bar{x}(t)$ и используя (1), получим

$$x^*(t+1) = x(t+1) - A(t)\bar{x}(t) - v(t). \quad (29)$$

Будем трактовать это соотношение как уравнение для определения $x(t+1)$, в котором учтена совокупность ранее проведенных наблюдений $y(t_0), \dots, y(t)$. Роль вектора помех в (29) играет величина $-v(t) - A(t)\bar{x}(t)$. Согласно (3) дополнительное измерение в момент $t+1$ имеет вид

$$y(t+1) = C(t+1)x(t+1) + w(t+1). \quad (29')$$

Объединенный «вектор наблюдений» $((x^*(t+1))^T, y^T(t+1))^T$ отвечает системе относительно $x(t+1)$

$$\begin{bmatrix} x^*(t+1) \\ y(t+1) \end{bmatrix} = \begin{bmatrix} I_n \\ C(t+1) \end{bmatrix} x(t+1) + \begin{bmatrix} -A(t)\bar{x}(t) - v(t) \\ \omega(t+1) \end{bmatrix}. \quad (29'')$$

Сравнивая (29'') с (9), видим, что матрица ковариации ошибки

$$s = \begin{bmatrix} -A(t)\bar{x}(t) - v(t) \\ \omega(t+1) \end{bmatrix}$$

имеет вид

$$S = \begin{bmatrix} P^*(t+1) & 0 \\ 0 & R(t+1) \end{bmatrix}, \quad (30)$$

где $R(t+1)$ определена в (4), $P^*(t+1)$ — в (28).

Обобщенная МНК-оценка получается согласно (21), где

$$D = \begin{bmatrix} I_n \\ C(t+1) \end{bmatrix}, \quad z = \begin{bmatrix} x^*(t+1) \\ y(t+1) \end{bmatrix}. \quad (31)$$

Несложные вычисления величин, входящих в (21), с учетом (30), (31) приводят к

$$\begin{aligned} \hat{x}(t+1) = & [(P^*(t+1))^{-1} + C^T(t+1)R^{-1}(t+1) \times \\ & \times C(t+1)]^{-1} \{ (P^*(t+1))^{-1} x^*(t+1) + \\ & + C^T(t+1)R^{-1}(t+1)y(t+1) \}. \end{aligned} \quad (32)$$

Обозначим множитель перед $y(t+1)$ через

$$\begin{aligned} K(t+1) = & [(P^*(t+1))^{-1} + C^T(t+1)R^{-1}(t+1)C(t+1)]^{-1} C^T(t+1)R^{-1}(t+1) = \\ = & \bar{P}(t+1)C^T(t+1)R^{-1}(t+1), \\ (\bar{P}(t+1))^{-1} = & (P^*(t+1))^{-1} + C^T(t+1) \times \\ & \times R^{-1}(t+1)C(t+1). \end{aligned} \quad (33)$$

Тогда, поменяв момент времени $t+1$ в (32), (33) на t , имеем

$$\hat{x}(t) = \bar{P}(t) (P^*(t))^{-1} x^*(t) + K(t) y(t), \quad (32')$$

а, умножив (33) слева на \bar{P} , получаем

$$I_n = \bar{P}(t) (P^*(t))^{-1} + K(t) C(t), \quad (33')$$

откуда $\bar{P}(t) (P^*(t))^{-1} = I_n - K(t) C(t)$. Подставляя его в (32'), имеем для обобщенной МНК-оценки

$$\hat{x}(t) = x^*(t) + K(t) \{y(t) - C(t)x^*(t)\} \quad (34)$$

или

$$\begin{aligned} \hat{x}(t+1) &= x^*(t+1) + K(t+1) \{y(t+1) - \\ &- C(t+1)x^*(t+1)\}. \end{aligned} \quad (34')$$

Получим выражения для $K(t)$ и $P(t)$, используя вид (34) обобщенной МНК-оценки и условие (26''). Заменяем в соответствии с (3) y на $Cx + w$. Тогда имеем для ошибки оценивания

$$\begin{aligned} \bar{x}(t+1) &= x(t+1) - \hat{x}(t+1) = \\ &= [I_n - K(t+1)C(t+1)] [x(t+1) - x^*(t+1)] - \\ &- K(t+1)w(t+1). \end{aligned} \quad (35)$$

Подставив (35) в (26''), рассмотрим сначала первую часть м.о., содержащую «старые» наблюдения $y^T(t_0), \dots, y^T(t)$. Видим, что $\bar{x}(t+1)$ некоррелирована со старыми наблюдениями в силу (26') и (35). Оставшаяся часть (26'') позволяет построить матрицу K . Согласно (35) и (3)

$$\begin{aligned} E\{[I_n - K(t+1)C(t+1)](x(t+1) - x^*(t+1)) - \\ - K(t+1)w(t+1)](x^T(t+1))C^T(t+1) + w(t+1)\} = \\ = 0. \end{aligned} \quad (36)$$

Величины $x(t+1)$ и $x^*(t+1)$ некоррелированы с $w(t+1)$ вследствие (1), (4), (8), (27). Используя (16), (27), (27'), (28), найдем, что ковариация между $x(t+1) - x^*(t+1)$ и $x(t+1)$ равна $P^*(t+1)$. В результате приведем (36) к виду

$$\begin{aligned} (I_n - K(t+1)C(t+1))P^*(t+1)C^T(t+1) - \\ - K(t+1)R(t+1) = 0, \end{aligned}$$

откуда

$$K(t+1) = P^*(t+1) C^T(t+1) [C(t+1) P^*(t+1) \times \\ \times C^T(t+1) + R(t+1)]^{-1}. \quad (37)$$

Матрица $K(t+1)$ называется матрицей передачи (усиления) Калмана. Наконец, умножая обе части (35) справа на $x^T(t+1)$ и вычисляя м.о. с учетом (17), имеем

$$\tilde{P}(t+1) = [I_n - K(t+1) C(t+1)] P^*(t+1). \quad (38)$$

Соотношения (27), (28), (34'), (37), (38) определяют *фильтр Калмана* с дискретным временем [14; 62]. При этом необходимые для организации вычислительного процесса величины $\hat{x}(t_0)$, $K(t_0)$ и $\tilde{P}(t_0)$ определены в (24), (25), (25'). Легко видеть, что (34), (34') имеют ту же форму, что и (3) из гл. 1: уточненная оценка получается из оценки предшествующего шага добавлением поправки, линейной по невязке.

4. Можно несколько видоизменить соотношения, определяющие фильтр Калмана, и уменьшить их число. Подставляя (34) в (27), получим

$$x^*(t+1) = A(t) x^*(t) + A(t) K(t) \{y(t) - C(t) x^*(t)\} = \\ = A(t) x^*(t) + \tilde{K}(t) \{y(t) - C(t) x^*(t)\}, \quad (39)$$

где матрица усиления $\tilde{K}(t) \triangleq A(t) K(t)$ определяется с учетом (37) соотношением

$$\tilde{K}(t+1) = A(t+1) P^*(t+1) C^T(t+1) [C(t+1) \times \\ \times P^*(t+1) C^T(t+1) + R(t+1)]^{-1}. \quad (40)$$

Наконец, для вычисления ковариационной матрицы оценки $x^*(t+1)$ запишем $x^*(t+1) - x(t+1)$, используя (1), (3), в виде

$$x^*(t+1) - x(t+1) = A(t) (x^*(t) - x(t)) - v(t) + \\ + \tilde{K}(t) \{\omega(t) - C(t) (x^*(t) - x(t))\}.$$

Вычисляя ковариационную матрицу этого выражения, получим

$$P^*(t+1) = [A(t) - \tilde{K}(t) C(t)] P^*(t) [A(t) - \\ - \tilde{K}(t) C(t)]^T + Q(t) + \tilde{K}(t) R(t) \tilde{K}(t). \quad (41)$$

Совокупность формул (39)–(41) представляет собой другую форму записи фильтра Калмана.

Соотношение (41) можно записать также еще в двух эквивалентных формах, используя (40) и известное матричное тождество [62; 68; 91]

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1},$$

справедливое для произвольных матриц A, B, C, D , размерности которых согласованы так, чтобы соответствующие матричные операции имели смысл. Модификации (41) суть следующие:

$$P^*(t+1) = A(t)P^*(t)A^T(t) - \\ - R(t)C(t)P^*(t)A^T(t) + Q(t), \quad (41')$$

$$P^*(t+1) = A(t)\{P^*(t) - P^*(t)C^T(t)[R(t) + \\ + C(t)P^*(t)C^T(t)]^{-1}C(t)P^*(t)\}A^T(t) + Q(t). \quad (41'')$$

Заметим, что матрицы $P^*(t)$, $\hat{P}(t)$ и $K(t)$ в (28), (37), (38) (соответственно $K(t)$ и $P^*(t)$ в (40), (41)) не зависят от $\{y(t)\}$ и могут быть вычислены заранее, до начала наблюдений. Это целесообразно, если основные требования в ходе обработки предъявляются к быстрдействию, а не к памяти, поскольку обращение матрицы, необходимое для реализации (37), (40), — самая трудоемкая матричная операция. Если же хранение указанных матриц нецелесообразно, то, учитывая симметричность обращаемой в (37), (40) матрицы, обычно используют обращение на основе факторизации Холецкого, обеспечивающее большую быстроту по сравнению с другими методами при хороших вычислительных свойствах [91].

Модель (1), (3) допускает обобщение [14] на случай ненулевых м.о. вектора состояний и наличия известных входных воздействий. Обобщенная модель, описывающая эволюцию вектора состояний, имеет вид

$$x(t+1) = A(t)x(t) + B(t)u(t) + v(t), \quad Ex(t_0) = \xi \neq 0, \quad (42)$$

$u(t)$ — вектор входных воздействий, значения которого $u(t_0), \dots, u(t)$ в моменты наблюдений заданы. Уравнение (3) по-прежнему описывает вектор наблю-

дений. Условия на $v(t)$, $w(t)$ и их авто- и взаимные ковариационные матрицы, а также их ковариацию $x(t_0)$ с $v(t)$ и $w(t)$ совпадают с (2), (4), (8). Тогда вместо (27) используется соотношение

$$x^*(t+1) = A(t)\hat{x}(t) + B(t)u(t), \quad x^*(t_0) = \xi, \quad (43)$$

соотношения (28), (34), (37), (38) сохраняют силу.

Прежде всего очевидна линейность $\hat{x}(t)$ по y . Рассмотрим ошибку экстраполированной и текущей оценки. С учетом (42), (43) имеем

$$x(t+1) - x^*(t+1) = A(t)\bar{x}(t) + v(t), \quad (44)$$

причем в начальный момент t_0

$$x(t_0) - x^*(t_0) = x(t_0) - \xi. \quad (45)$$

Воспользовавшись (34) и (3), получим для ошибки предсказания

$$\begin{aligned} \bar{x}(t) = x(t) - \hat{x}(t) = x(t) - x^*(t) - K(t)\{C(t)x(t) + \\ + w(t) - C(t)x^*(t)\} = \{I_n - K(t)C(t)\}(x(t) - \\ - x^*(t)) - K(t)w(t). \end{aligned} \quad (46)$$

Вычисляя м. о. от левой и правой части (45), получим $E[x(t_0) - x^*(t_0)] = 0$. Обратившись теперь к вычислению $E\bar{x}(t_0+1)$, в силу свойств $w(t)$ и только что полученной формулы найдем $E\bar{x}(t_0+1) = 0$. Теперь, вычисляя м. о. от обеих частей равенства (44) при $t = t_0$, получим с учетом свойств $v(t)$

$$E[x(t_0+1) - x^*(t_0+1)] = 0.$$

После этого находим из (46) $E\bar{x}(t_0+2) = 0$ и т. д., отсюда $E\bar{x}(t) = 0$ и $E x(t) = E x^*(t)$ в моменты наблюдений $t_0, t_0+1, \dots, t-1$ независимо от значений ξ и u , т. е. оценка $\hat{x}(t)$ несмещенная. Ковариационная матрица разности $x(t+1) - x^*(t+1)$ с учетом (44) записывается в виде

$$P^*(t+1) = A(t)P(t)A^T(t) + Q(t),$$

что совпадает с (28). Ее начальное значение согласно (45) есть $P(t_0)$. Для матрицы ковариаций с. в. $\bar{x}(t)$ из (46) имеем

$$P(t) = [I_n - K(t)C(t)]P^*(t)[I_n - C^T(t)K^T(t)] + \\ + K(t)R(t)K^T(t)$$

или, раскрывая скобки,

$$P(t) = P^*(t) - K(t)C(t)P^*(t) - P^*(t)C^T(t)K^T(t) + \\ + K(t)\{C(t)P^*(t)C^T(t) + R(t)\}K^T(t). \quad (47)$$

Если $K(t)$ вычислена по (37), обе последние суммы исчезают, а остаток, совпадающий с (38), дает ковариационную матрицу оценки, имеющей минимальную дисперсию ошибки оценивания.

С результатом обобщения модели (1), (3) на случай небелого шума ошибок $w(t)$ и $v(t)$, их возможной коррелированности, а также с постановкой задачи и результатами для нелинейной связи между x и y читатель может познакомиться в [62].

Покажем, как связана весьма общая модель (1)–(4), (6)–(8) с задачами обработки экспериментов, которые описываются линейной моделью типа (5) из гл. 6 с неслучайным вектором состояний и случайным вектором наблюдений. Положим

$$A(t) = I_n, \quad v(t) = 0, \quad C \doteq C(t) = \text{const}, \quad R \doteq R(t) = \text{const},$$

тогда (42) перепишем в виде

$$x(t+1) = x(t), \quad Ex(t_0) = a.$$

Используя соотношения для фильтра Калмана в форме (39), (40), (41''), получим, что для этого случая (39) переходит в соотношение

$$x^*(t+1) = x^*(t) + K(t)\{y(t) - Cx^*(t)\}, \quad x^*(t_0) = a, \quad (48)$$

а вместо (40) и (41'') выполнены соотношения

$$K(t+1) = P^*(t+1)C^T(CP^*(t+1)C^T + R)^{-1}, \quad (49)$$

$$P^*(t+1) = P^*(t) - P^*(t)C^T[R + CP^*(t)C^T]^{-1}CP^*(t), \quad (50)$$

$$P^*(t_0) = P(t_0).$$

Формулы (39), (48)–(50) определяют одну из возможных реализаций рекуррентного метода наименьших квадратов [14; 40; 62]. С другими возможными

подходами к построению процедур рекуррентного МНК можно познакомиться в [68; 123].

В частности, рассматривая задачу решения линейного интегрального уравнения первого рода, после замены интеграла квадратурами приходим к линейной модели описанного вида.

Имея измерения экспериментального выхода y_1, \dots, \dots, y_N в моменты времени $t=1, \dots, N$, можно построить решение через условные байесовские оценки [40], фактически идентичное уравнениям (48)–(50) рекуррентного МНК. При этом суммирование рекуррентных соотношений дает [40] выражение для оценок через начальные значения $x_0 \triangleq x(t_0)$, $P_0 \triangleq P(t_0)$, имеющее вид

$$x^*(N) = \left[\frac{1}{N} I_n + P_0 C^T R^{-1} C \right]^{-1} \left[\frac{x_0}{N} + P_0 C^T R^{-1} \frac{1}{N} \sum_{s=1}^N y_s \right]. \quad (51)$$

Для их ковариационной матрицы справедливо соотношение

$$P^*(N) = \left[\frac{1}{N} I_n + P_0 C^T R^{-1} C \right]^{-1} \frac{1}{N} P_0. \quad (52)$$

В частном случае R , коммутирующего с C , имеем

$$x^*(N) = \left[\frac{1}{N} R P_0^{-1} + C^T C \right]^{-1} \times \\ \times \left[\frac{1}{N} R P_0^{-1} x_0 + C^T \frac{1}{N} \sum_{s=1}^N y_s \right], \quad (51')$$

$$P^*(N) = \left[\frac{1}{N} R P_0^{-1} + C^T C \right]^{-1} \frac{1}{N} R. \quad (52')$$

Легко показать, что оценка (51') минимизирует функционал

$$J = (Cx - \bar{y})^T R^{-1} (Cx - \bar{y}) + \\ + \frac{1}{N} (x - x_0)^T P_0^{-1} (x - x_0), \quad (53)$$

т. е. представляет собой регуляризованную оценку (ридж-оценка) задачи МНК. Нетрудно видеть, что (53) является квадратичной формой, стоящей в показателе экспоненты ФП для задачи оценки вектора x по результатам косвенных измерений y_1, \dots, y_N , если x распределен по нормальному закону с априорной оценкой м. о. x_0 и априорной ковариационной матрицей P_0 .

ЛИТЕРАТУРА

1. Айвазян С. А. Статистическое исследование зависимостей. М.: Металлургия, 1968.
2. Айвазян С. А., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: Основы моделирования и первичная обработка данных. М.: Финансы и статистика, 1983.
3. Айвазян С. А., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: Исследование зависимостей. М.: Финансы и статистика, 1985.
4. Альберт А. Регрессия, псевдоинверсия и рекуррентное оценивание. М.: Наука, 1977.
5. Андерсон Т. Введение в многомерный статистический анализ. М.: Физматгиз, 1963.
6. Арсенин В. Я., Крянев А. В. Применение статистических методов решения некорректных задач для обработки результатов физических экспериментов//Автоматизация научных исследований в экспериментальной физике. М.: Энергоатомиздат, 1986. С. 19—30.
7. Бард Й. Нелинейное оценивание параметров. М.: Статистика, 1979.
8. Бслов А. Г., Галкин В. Я., Уфимцев М. В. Вероятностно-статистические задачи при экспериментальном разделении множественных процессов. М.: Изд-во Моск. ун-та, 1985.
9. Библиотека программ по обработке и интерпретации физических экспериментов/Под ред. А. Н. Тихонова, П. Н. Заикина. М.: Изд-во Моск. ун-та, 1983.
10. Бикматов Р. Р., Уфимцев М. В. Исследование свойств некоторых робастных оценок параметра размещения при асимметричных загрязнениях//Математические задачи обработки эксперимента. М.: Изд-во Моск. ун-та, 1984. С. 65—86.
11. Большев Л. Н. Об исключении грубых наблюдений//Теория всроятностей и ее применения. 1961. № 6. С. 482—484.

12. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. М.: Наука, 1965.

13. Большев Л. Н. О критериях исключения резко выделяющихся наблюдений//Труды института прикладной математики Тбилисского государственного университета. Т. II. Тбилиси: Изд-во ТГУ, 1969. С. 159—177.

14. Браммер К., Зиффлинг Г. Фильтр Калмана—Бьюси: Детерминированное наблюдение и стохастическая фильтрация. М.: Наука, 1982.

15. Васильев Ф. П. Численные методы решения экстремальных задач. М.: Наука, 1980.

16. Воеводин В. В. Вычислительные основы линейной алгебры. М.: Наука, 1977.

17. Гаек Я., Шидак З. Теория ранговых критериев. М.: Наука, 1971.

18. Об оптимальной обработке и организации эксперимента по статистическому разделению выходов ядерных реакций различной множественности/В. Я. Галкин, Б. И. Горячев, В. Н. Орлин, М. Ф. Уфимцев//Вычислительные методы и программирование. Вып. 18. М.: Изд-во Моск. ун-та, 1972. С. 161—172.

19. Галкин В. Я., Орлин В. Н. Первичная обработка экспериментальных данных, полученных в автоматическом режиме измерений//Вычислительные методы и программирование. Вып. 21. М.: Изд-во Моск. ун-та, 1973. С. 74—107.

20. Галкин В. Я., Орлин В. Н., Литвакова И. Э. Учет систематических искажений при обработке данных фотоядерных измерений//Вычислительные методы и программирование. Вып. 21. М.: Изд-во Моск. ун-та, 1973. С. 108—117.

21. Галкин В. Я., Уфимцев М. В. Свойства некоторых оценок параметра сдвига, устойчивых к отклонениям от нормальности//Обработка и интерпретация физических экспериментов. Вып. 4. М.: Изд-во Моск. ун-та, 1976. С. 25—41.

22. Алгоритм статистических испытаний при расчете аппаратурной функции детектора гамма-квантов/В. Я. Галкин, Б. С. Ишханов, В. И. Макеев, М. В. Уфимцев//Обработка и интерпретация физических экспериментов. М.: Изд-во Моск. ун-та, 1979. С. 46—56.

23. Галкин В. Я., Уфимцев М. В. Монте-Карловское исследование многомерных робастных оценок параметра сдвига//Математическая статистика и ее применения. 1981. № 7. С. 20—23.

24. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. М.: Мир, 1985.

25. Гнеденко Б. В. Курс теории вероятностей. М.: Наука, 1969.

26. Голенко Д. И. Моделирование и статистический анализ псевдослучайных чисел на ЭВМ. М.: Наука, 1965.

27. Гольданский В. И., Куценко А. В., Подгорецкий М. И. Статистика отсчетов при регистрации ядерных частиц. М.: Физматгиз, 1959.

28. Гончарский А. В., Леонов А. С., Ягола А. Г. Об одном регуляризирующем алгоритме для некорректно поставленных задач с приближенно заданным оператором//ЖВМ и МФ. 1972. Т. 12, № 6. С. 1592—1594.

29. Гончарский А. В., Леонов А. С., Ягола А. Г. Обобщенный принцип невязки//ЖВМ и МФ. 1973. Т. 13, № 2. С. 294—302.

30. Де Гроот М. Оптимальные статистические решения. М.: Мир, 1974.

31. Демиденко Е. З. Линейная и нелинейная регрессия. Фортран-IV. М.: Изд-во Института мировой экономики и международных отношений, 1979.

32. Демиденко Е. З. Линейная и нелинейная регрессия. М.: Финансы и статистика, 1981.

33. Джонсон Н., Лион Ф. Статистика и планирование эксперимента в технике и науке: Методы обработки данных. М.: Мир, 1980.

34. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. М.: Статистика, 1973.

35. Дэйвид Г. Порядковые статистики. М.: Наука, 1979.

36. Ермаков С. М. Метод Монте-Карло и смежные вопросы. М.: Наука, 1975.

37. Ермаков С. М., Михайлов Г. А. Курс статистического моделирования. М.: Наука, 1976.

38. Ершов А. А. Стабильные методы оценки параметров (обзор)//Автоматика и телемеханика. 1978. № 8. С. 66—100.

39. Жуковский Е. Л. О применении статистических регуляризирующих алгоритмов к решению некоторых типов некорректно поставленных задач//Некоторые вопросы автоматизированной обработки и интерпретации физических экспериментов. М.: Изд-во Моск. ун-та, 1973. С. 133—169.

40. Жуковский Е. Л. Статистическая регуляризация систем алгебраических уравнений//ЖВМ и МФ. 1972. Т. 12, № 1. С. 185—191.

41. Иберла К. Факторный анализ. М.: Статистика, 1980.

42. Иванов В. К. О некорректно поставленных задачах//Математический сборник. 1963. Т. 61, № 2. С. 211—223.

43. Калиткин Н. Н. Численные методы. М.: Наука, 1978.

44. Кендалл М. Дж., Стьюарт А. Теория распределений. М.: Наука, 1966.

45. Кендалл М. Дж., Стьюарт А. Статистические выводы и связи. М.: Наука, 1973.

46. Клейнен Дж. Статистические методы в имитационном моделировании. Вып. 1, 2. М.: Статистика, 1978.

47. Кнут Д. Искусство программирования для ЭВМ. Получисленные алгоритмы. Т. 2. М.: Мир, 1977; Сортировка и поиск. Т. 3. М.: Мир, 1978.

48. Конюховский В. В. Критерии согласия, однородности и независимости. М.: Изд-во Моск. ун-та, 1970.

49. Крамер Г. Математические методы статистики. М.: Мир, 1975.

50. Крянев А. В. Статистическая форма регуляризованного метода наименьших квадратов (РМНК) А. Н. Тихонова//ДАН СССР. 1986. Т. 291, № 4. С. 780—785.

51. Кукс Я., Ольман В. Минимаксная линейная оценка коэффициентов регрессии//Известия АН Эстонской ССР. Сер. физико-математическая. 1972. Т. 21, № 1. С. 66—72.

52. Лаврентьев М. М., Федотов А. М. О постановке некоторых задач математической физики со случайными исходными данными//ЖВМ и МФ. 1982. Т. 22, № 1. С. 133—143.

53. Линник Ю. В. Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений. М.: Физматгиз, 1962.
54. Лоули Д. Н., Максвелл А. Э. Факторный анализ как статистический метод. М.: Мир, 1967.
55. Лозв М. Теория вероятностей. М.: ИЛ, 1962.
56. Математическая теория планирования эксперимента/Под ред. С. М. Ермакова. М.: Наука, 1983.
57. Математическое обеспечение ЕС ЭВМ. Вып. 2, 10, 16, 24, 25, 44. Минск: Изд-во Института математики АН БССР, 1973; 1976; 1978; 1980; 1983.
58. Мелешко В. И. Решение некорректных стохастических алгебраических уравнений//ЖВМ и МФ. 1980. Т. 20, № 1. С. 11—26.
59. Мешалкин Л. Д. Параметризация многомерных распределений//Прикладной многомерный статистический анализ. М.: Наука, 1978. С. 11—18.
60. Морозов В. А. О принципе невязки при решении операторных уравнений методом регуляризации//ЖВМ и МФ. 1968. Т. 8, № 2. С. 295—309.
61. Мостеллер Ф., Тьюки Дж. Анализ данных и регрессия. Вып. 1, 2. М.: Финансы и статистика, 1982.
62. Мудров В. И., Кушко В. Л. Методы обработки измерений: Квазиправдоподобные оценки. М.: Радио и связь, 1983.
63. Налимов В. В., Чернова Н. А. Статистические методы планирования экстремальных экспериментов. М.: Наука, 1965.
64. Немыцкий В. В., Слудская М. И., Черкасов А. Н. Курс математического анализа. Т. 1. М.: ГТТИ, 1957.
65. Петров А. П. Оценка линейных функционалов для решения некоторых обратных задач//ЖВМ и МФ. 1967. Т. 7, № 3. С. 648—654.
66. Поллард Дж. Справочник по вычислительным методам статистики. М.: Финансы и статистика, 1982.
67. Сборник научных программ на Фортране. Руководство для программиста. Вып. 1. М.: Статистика, 1974.

68. Себер Дж. Линейный регрессионный анализ. М.: Мир, 1980.

69. Смоляк С. А., Титаренко Б. П. Устойчивые методы оценивания. М.: Статистика, 1980.

70. Соболев И. М. Численные методы Монте-Карло. М.: Наука, 1973.

71. Справочник по специальным функциям/Под ред. М. Абрамовица и И. Стиган. М.: Наука, 1979.

72. Тихонов А. Н. О решении некорректно поставленных задач и методе регуляризации//ДАН СССР. 1963. Т. 151, № 3. С. 501—504.

73. Тихонов А. Н. О регуляризации некорректно поставленных задач//ДАН СССР. 1963. Т. 153, № 1. С. 49—52.

74. Тихонов А. Н. О некорректных задачах линейной алгебры и устойчивых методах их решения//ДАН СССР. 1965. Т. 163, № 3. С. 591—594.

75. Тихонов А. Н. Об устойчивости алгоритмов для решения вырожденных систем линейных алгебраических уравнений//ЖВМ и МФ. 1965. Т. 5, № 4. С. 718—722.

76. Система сплошной автоматической обработки результатов эксперимента по исследованию сечений фотоядерных реакций/А. Н. Тихонов, В. Г. Шевченко, В. Я. Галкин и др.//Вычислительные методы и программирование. Вып. 14. М.: Изд-во Моск. ун-та, 1970. С. 3—26.

77. Применение факторного анализа для учета аппаратного дрейфа при обработке результатов автоматических измерений/А. Н. Тихонов, Л. Н. Большев, В. Я. Галкин и др.//Вычислительные методы и программирование. Вып. 21. М.: Изд-во Моск. ун-та, 1973. С. 118—123.

78. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. М.: Наука, 1974.

79. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. М.: Наука, 1979.

80. Тихонов А. Н. Об одном принципе взаимности//ДАН СССР. 1980. Т. 253, № 2. С. 305—308.

81. Тихонов А. Н. О нормальных решениях приближенных систем линейных алгебраических уравнений//ДАН СССР. 1980. Т. 254, № 3. С. 549—554.

82. Тихонов А. Н. О приближенных системах линейных алгебраических уравнений//ЖВМ и МФ. 1980. Т. 20, № 6. С. 1373—1383.
83. Регуляризирующие алгоритмы и априорная информация/А. Н. Тихонов, А. В. Гончарский, В. В. Степанов, А. Г. Ягола. М.: Наука, 1983.
84. Тихонов А. Н. Об обратных задачах//Некорректные задачи математической физики и анализа. Семинар. Новосибирск, 19—24 июля 1982 г. Новосибирск, 1984. С. 161—171.
85. Тихонов А. Н. О задачах с неточно заданной исходной информацией//ДАН СССР. 1985. Т. 280, № 3. С. 559—562.
86. Уилкинсон Дж., Райнш К. Справочник алгоритмов на языке Алгол. Линейная алгебра. М.: Машиностроение, 1976.
87. Уилкс С. Математическая статистика. М.: Наука, 1967.
88. Устойчивые статистические методы оценки данных/Под ред. Р. Лонера, Г. Уилкинсона. М.: Машиностроение, 1984.
89. Уфимцев М. В. Применение факторного анализа для учета влияния аппаратного дрейфа в некоторых классах физических экспериментов//Некоторые вопросы автоматизированной обработки и интерпретации физических экспериментов. Вып. 1. М.: Изд-во Моск. ун-та, 1973. С. 123—132.
90. Уфимцев М. В. Некоторые свойства многомерных «устойчивых» оценок параметра сдвига//Обработка и интерпретация физических экспериментов. Вып. 4. М.: Изд-во Моск. ун-та, 1976. С. 42—68.
91. Фаддеев Д. К., Фаддеева В. Н. Вычислительные методы линейной алгебры. М.: Физматгиз, 1963.
92. Федотов А. М. Линейные некорректные задачи со случайными ошибками в данных. Новосибирск: Наука, 1982.
93. Феллер В. Введение в теорию вероятностей и ее применение. Т. 1, 2. М.: Мир, 1967.
94. Федоров В. В. Теория оптимального эксперимента (планирование регрессионных экспериментов). М.: Наука, 1971.

95. Форсайт Дж., Молер К. Численное решение систем линейных алгебраических уравнений. М.: Мир, 1969.
96. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений. М.: Мир, 1980.
97. Харман Г. Современный факторный анализ. М.: Статистика, 1972.
98. Химмельблау Д. Анализ процессов статистическими методами. М.: Мир, 1973.
99. Шеффе Г. Дисперсионный анализ. М.: Физматгиз, 1963.
100. Шурьгин А. М. Устойчивые оценки параметров распределения//Заводская лаборатория. 1982. № 2. С. 88—93.
101. Anderson T. W., Rubin H. Statistical inference in factor analysis//Proc. 3-rd Berkeley Symposium. Berkeley. 1956. V. 5. P. 111—150.
102. Robust estimates of location. Survey and advances/Andrews et al. Princeton: Princeton Univ. Press, 1972.
103. Bessel F. W. Fundamenta Astronomiae. Königsberg, 1818.
104. Bickel P. J. On some alternative estimates for shift in the p-variate one sample problem//Ann. of Math. Statist. 1964. V. 35. P. 1079—1090.
105. Businger P., Golub G. H. Linear least squares solutions by Householder transformations//Numer. Math. 1965. V. 7. P. 269.
106. Cochran W. G. The χ^2 test of goodness of fit//Ann. of Math. Statist. 1952. V. 23. P. 315.
107. Cochran W. G. Some methods for strengthening the common χ^2 tests//Biometrics. 1954. V. 10. P. 417.
108. Ferguson T. S. On the rejection of outliers//Proc. 4-th Berkeley Sympos. Math. Statist. and Probability. Berkeley. 1961. V. 1. P. 177.
109. Fisher R. A. On the mathematical foundations of theoretical statistics//Phil. Trans. Ser. A. 1921. V. 222. P. 309.
110. Gentleman W. Algorithm AS75: basic procedures for large sparse or weighted linear least squares problems//Appl. Statist. 1974. V. 23. P. 448.

111. Gnanadesikan R., Kettenring J. R. Robust estimates, residuals and outlier detection with multiresponse data//Biometrics. 1972. V. 28. P. 81.
112. Golub G. H., Reinsch C. Singular value decomposition and least squares solutions//Numer. Math. 1970. V. 14. P. 403.
113. Hodges J. L., Lehmann E. L. Estimates of location based on rank tests//An. of Math. Statist. 1962. V. 34. P. 598.
114. Hoerl A. E., Kennard R. W. Ridge regression. Biased estimation for non-orthogonal problems//Technometrics. 1970. V. 12. P. 55.
115. Hoerl A. E., Kennard R. W. Ridge regression. Applications to non-orthogonal problems//Technometrics. 1970. V. 12. P. 69.
116. Hogg R. V. Some observations on robust estimation//J. Amer. Statist. Assoc. 1967. V. 62. P. 1179.
117. Hogg R. V. Adaptive robust procedures: a partial review and some suggestions for future applications and theory//J. Amer. Statist. Assoc. 1974. V. 69. P. 909.
118. Hogg R. V., Randles R. H. Adaptive distribution-free regression methods and their applications//Technometrics. 1975. V. 17. P. 399.
119. Huber P. J. Robust estimation of location parameter//Ann. of Math. Statist. 1964. V. 35, N 1. P. 73.
120. Huber P. J. Robust statistics: a review//Ann. of Math. Statist. 1973. V. 43, N 4. P. 1041.
121. Jackel L. A. Some flexible estimates of location//Ann. of Math. Statist. 1971. V. 42, N 5. P. 1540.
122. Kolmogorov A. N. Sulla determinazione empirica di una legge di distribuzione//Giorn. del Inst. Ital. Attuari. 1933. V. 4. P. 83.
123. Lawson C. L., Hanson R. J. Solving least squares problems. New Jersey, 1974. (русский перевод: Лоусон Ч., Хенсон Р. Численное решение задач метода наименьших квадратов. М.: Наука. 1986).
124. Lehmer P. H. Mathematical methods in large-scale computing units//Ann. Comput. Lab. Harvard University. 1951. V. 26. P. 141.

125. Marquardt D. W. An algorithm for least squares estimation of nonlinear parameters//J. Soc. Industr. and Appl. Math. 1963. V. 11, N 2. P. 431.
126. Meshalkin L. D. Approximation of multidimensional densities by normal distributions. Reports of the 7-th Intern. Biometrical Conference. Hannover, Germany, 1970.
127. Neyman J., Pearson E. S. On the use and interpretation of certain test criteria for the purposes of statistical inference//Biometrika. 1928. V. 20A. P. 175 and 263.
128. Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables in such that it can be reasonably supposed to have arisen in random sampling//Phil. Mag. 1900. V. 50. P. 157.
129. Phillips D. L. A technique for numerical solution of certain integral equations of the first kind//J. Assoc. Comput. Mach. 1962. V. 9. P. 84.
130. Plackett R. L. A historical note on the method of least squares//Biometrika. 1949. V. 36. P. 458.
131. Ralston M. L., Jennrich R. I. Duda derivative-free algorithm for nonlinear least squares//Technometrics. 1978. V. 20, N 1. P. 7.
132. Rao C. R. Information and accuracy attainable in the estimation of statistical parameters//Bull. Calcutta Math. Soc. 1945. V. 37. P. 81.
133. Rao C. R. Estimation of parameters in the linear model//Ann. of Math. Statist. 1976. V. 4, N 6. P. 1023.
134. Slacter M. J. Comparative validity of the chi-square and two modified chi-square goodness-of-fit tests for small but equal expected frequencies//Biometrika. 1966. V. 53. P. 619.
135. Swed F. S., Eisenhart G. Tables for testing randomness of grouping in sequence of alternatives//Ann. of Math. Statist. 1943. V. 14. P. 66.
136. On overall automatization of data processing for determining the photonuclear reaction cross-sections/Tichonov A. N. et al.//IFIP Congress. V. 12. Edinburg, 1968. P. 1549.

137. Tukey J. W. The future of Data Analysis//
//Ann. of Math. Statist. 1962. V. 33, N 1. P. 1.
138. Wilks S. S. Multivariate statistical outliers//
//Sankhya. Ser. A. 1962. V. 25, part 4. P. 407.

Учебное пособие

**Тихонов Андрей Николаевич,
Уфимцев Михаил Валентинович**

**СТАТИСТИЧЕСКАЯ ОБРАБОТКА
РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТОВ**

Зав. редакцией С. И. Зеленский
Редактор Е. Д. Егорушкина
Художественный редактор Ю. М. Добрянская
Технический редактор М. Б. Терентьева
Корректор Л. А. Айдарбекова

ИБ № 2952

Сдано в набор 24.09.87. Подписано в печать 22.02.88.
Л-36560. Формат 84×108/32. Бумага тип. № 1.
Гарнитура литературная. Высокая печать. Усл. печ. л. 9,24.
Уч.-изд. л. 8,6. Тираж 10 500 экз. Заказ 220.
Изд. № 4854. Цена 30 коп.

Ордена «Знак Почета» издательство
Московского университета.
103009, Москва, ул. Герцена, 5/7.
Типография ордена «Знак Почета» изд-ва МГУ.
119899, Москва, Ленинские горы

**МАГАЗИНЫ — ОПОРНЫЕ ПУНКТЫ
ПО ИЗУЧЕНИЮ СПРОСА
И РАСПРОСТРАНЕНИЮ
КНИГ ИЗДАТЕЛЬСТВА МГУ**

117296, Москва,
Ломоносовский проспект, 18
магазин № 110
«Университетская книжная лавка»
(для москвичей)

252001, Киев, ул. Крещатик, 44
магазин № 12 «Книги»

630090, Новосибирск, ул. Ильича, 6
магазин № 2

Магазин № 93 «Книга — почтой» принимает предварительные заказы на книги Издательства МГУ и высылает литературу иногородним покупателям.

Адрес магазина: 117168, Москва,
ул. Кржижановского, 14 магазин № 93
«Книга — почтой».

Цена 30 коп