

## МОДЕЛИ, СОДЕРЖАЩИЕ РАЗЛИЧНЫЕ ФУНКЦИИ ПРЕДИКТОРОВ. ПОЛИНОМИАЛЬНЫЕ МОДЕЛИ

Общие методы построения регрессии для линейных моделей, описанные в главах 6–12, остаются в силе и для модельной функции произвольного вида (линейной относительно параметров). Теоретически модель может быть любой. Для начала мы обсудим общий случай, а затем рассмотрим конкретное множество данных, для которого подберем модель второго порядка. Этот пример поможет понять, на чем основаны предлагаемые правила исключения членов из полиномиальных моделей.

### 13.1. БОЛЕЕ СЛОЖНЫЕ МОДЕЛЬНЫЕ ФУНКЦИИ

На практике большинство подбираемых моделей содержит только сами наблюдаемые предикторные переменные  $X_1, X_2, \dots, X_k$ , т.е. модель имеет вид

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \epsilon, \quad (13.1.1)$$

однако возможны и более общие формы. Наиболее общая линейная модель, содержащая переменные  $X_1, X_2, \dots, X_k$ , имеет вид

$$Y = \beta_0 Z_0 + \beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_{p-1} Z_{p-1} + \epsilon. \quad (13.1.2)$$

*Примечание.*  $Z_0 = 1$  — фиктивная переменная, которая всегда равна единице и обычно опускается. Однако для удобства вычислений иногда имеет смысл ввести в модель такую переменную. Например, если

$$(Z_{1i}; Z_{2i}; \dots; Z_{(p-1)i}), \quad i = 1, 2, \dots, n,$$

—  $n$  наборов значений переменных  $Z_j$ ,  $j = 1, 2, \dots, p - 1$ , соответствующих наблюдениям  $Y_i$ ,  $i = 1, 2, \dots, n$ , то при  $j \neq 0$  и  $Z_{0i} = 1$

$$\sum_{i=1}^n Z_{ji} = \sum_{i=1}^n Z_{ji} Z_{0i},$$

и при записи нормальных уравнений данная сумма может быть представлена как перекрестное произведение общего вида:

$$\sum_{i=1}^n Z_{ji} Z_{li}.$$

Отметим, что  $\sum_{i=1}^n Z_{0i}^2 = n$ .

В уравнении (13.1.2) каждая переменная  $Z_j$ ,  $j = 1, 2, \dots, p-1$ , является произвольной функцией переменных  $X_1, X_2, \dots, X_k$ . В некоторых примерах каждая переменная  $Z_j$  может содержать только одну переменную  $X_i$ .

Любую модель, которую в результате преобразований можно записать в виде (13.1.2), можно анализировать с помощью общих методов, описанных в главах 6–12. Рассмотрим несколько конкретных примеров моделей, к которым применимы данные методы, и свяжем их с общей формой (13.1.2).

### Полиномиальные модели различного порядка относительно $X_j$

#### Модели первого порядка

1. Если в уравнении (13.1.2)  $p = 2$ , а  $Z_1 = X$ , то имеем простую модель первого порядка с одной предикторной переменной:

$$Y = \beta_0 + \beta_1 X + \epsilon. \quad (13.1.3)$$

2. Если  $p = k + 1$ , а  $Z_j = X_j$ , получаем модель первого порядка с  $k$  предикторными переменными:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \epsilon. \quad (13.1.4)$$

#### Модели второго порядка

1. При  $p = 3$ ,  $Z_1 = X$ ,  $Z_2 = X^2$ ,  $\beta_2 = \beta_{11}$  получается модель второго порядка (квадратичная) с одной предикторной переменной:

$$Y = \beta_0 + \beta_1 X + \beta_{11} X^2 + \epsilon. \quad (13.1.5)$$

2. При  $p = 6$ ,  $Z_1 = X_1$ ,  $Z_2 = X_2$ ,  $Z_3 = X_1^2$ ,  $Z_4 = X_2^2$ ,  $Z_5 = X_1 X_2$ ,  $\beta_3 = \beta_{11}$ ,  $\beta_4 = \beta_{22}$ ,  $\beta_5 = \beta_{12}$  мы получаем модель второго порядка с двумя предикторными переменными:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{12} X_1 X_2 + \epsilon. \quad (13.1.6)$$

Полная модель второго порядка для  $k$  переменных получается аналогичным образом при  $p = 1 + k + k + \frac{1}{2}k(k-1) = \frac{1}{2}(k+1)(k+2)$ . Модели второго порядка используются преимущественно при исследовании поверхностей

отклика, когда характеристики неизвестной поверхности отклика желательно аппроксимировать с помощью полинома невысокого порядка. Заметим, что модель включает все возможные члены второго порядка. Это важно, поскольку исключение членов подразумевает знание того, что определенные типы поверхностей (которые невозможно представить без исключенных членов) возникнуть не могут. Такие знания имеются не часто. При наличии подобной информации можно проводить более обоснованные исследования.

Пример анализа поверхности откликов второго порядка приводится в разделе 13.2.

### Модели третьего порядка

1. При  $p = 4$ ,  $Z_1 = X$ ,  $Z_2 = X^2$ ,  $Z_3 = X^3$ ,  $\beta_2 = \beta_{11}$ ,  $\beta_3 = \beta_{111}$  получается модель третьего порядка с одной предикторной переменной:

$$Y = \beta_0 + \beta_1 X + \beta_{11} X^2 + \beta_{111} X^3 + \epsilon. \quad (13.1.7)$$

2. При  $p = 10$  и введении соответствующих обозначений для  $\beta_i$  и  $Z_i$  уравнение (13.1.2) может представлять модель третьего порядка с двумя предикторными переменными:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{12} X_1 X_2 + \beta_{22} X_2^2 + \beta_{111} X_1^3 + \beta_{112} X_1^2 X_2 + \beta_{122} X_1 X_2^2 + \beta_{222} X_2^3 + \epsilon. \quad (13.1.8)$$

Модель третьего порядка с  $k$  предикторными переменными  $X_1, X_2, \dots, X_k$  получается аналогичным образом. Модели третьего порядка также используются при работе с поверхностями отклика, хотя значительно реже, чем модели второго порядка. Обратите внимание на принцип индексирования параметров  $\beta$ . Он может показаться громоздким, но позволяет легко связать коэффициенты с соответствующими степенями переменных  $X$ . Например, члену с  $X_1 X_2^2 = X_1 X_2 X_2$  соответствует коэффициент  $\beta_{122}$ , и т.д. Аналогичная система обозначений используется в моделях второго порядка; это стандарт при работе с поверхностью отклика. (Заметьте, что мы пишем  $\beta_{122}$ , а не  $\beta_{212}$  или  $\beta_{221}$ , т.е. меньшие индексы предшествуют бóльшим.)

Продолжая описанный процесс, из уравнения (13.1.2) можно получить модель любого порядка.

### Преобразования

Если модель второго порядка неадекватна, то, возможно, подойдет модель третьего порядка. Однако не следует бездумно включать в модель члены более высокого порядка. Обычно более продуктивным оказывается исследование влияния других преобразований предикторов, откликов или тех, и других

одновременно. То же замечание справедливо и в случае перехода от модели первого порядка к модели второго порядка. Например, подбор прямолинейной зависимости  $\log Y$  от  $X$  обычно предпочтительнее, чем подбор квадратичной зависимости  $Y$  от  $X$  при условии, что анализ остатков указывает на возможность обоих вариантов.

### Модели, включающие преобразования, отличные от возведения в целые степени

Представленные выше полиномиальные модели содержат степени и перекрестные произведения степеней предикторных переменных  $X_1, X_2, \dots, X_k$ . Рассмотрим несколько примеров других преобразований, которые часто используются при формировании модельной функции.

*«Обратное» преобразование.* Если в уравнении (13.1.2) мы положим  $p = 2$ ,  $Z_1 = 1/X_1$ ,  $Z_2 = 1/X_2$ , то получим модель:

$$Y = \beta_0 + \beta_1 \left( \frac{1}{X_1} \right) + \beta_2 \left( \frac{1}{X_2} \right) + \epsilon. \quad (13.1.9)$$

*Логарифмическое преобразование.* Если взять  $p = 2$ ,  $Z_1 = \ln X_1$ ,  $Z_2 = \ln X_2$ , то уравнение (13.1.2) можно записать как

$$Y = \beta_0 + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \epsilon. \quad (13.1.10)$$

*Преобразование типа квадратного корня.* Например,

$$Y = \beta_0 + \beta_1 X_1^{1/2} + \beta_2 X_2^{1/2} + \epsilon. \quad (13.1.11)$$

Очевидно, что существует множество всевозможных преобразований и можно постулировать модели, содержащие несколько таких членов. *Одна модель может содержать несколько различных преобразований.* (В примерах (13.1.9)–(13.1.11) в одной модели используются только одинаковые преобразования.) Часто выбор преобразования не является очевидным и осуществляется на основе предыдущих знаний о факторах<sup>1</sup> данной задачи. Цель преобразований такого рода состоит в том, чтобы получить для преобразованных переменных более простую регрессионную модель, чем для исходных переменных.

Преобразования могут также включать функции нескольких переменных одновременно, например,  $Z_1 = X_1^{1/2} \ln X_2$ . Преобразования такого вида иногда вытекают из вида подобранного уравнения в исходных переменных. Когда не известно, какую степень  $X$  лучше использовать, можно сделать ее параметром. В таких случаях обычно помогают методы нелинейного оценивания.

<sup>1</sup>Здесь и далее термин «фактор» означает то же, что «предиктор». В разделе 2.1 приводится целый ряд синонимов, обозначающих то же понятие. — *Прим. ред.*

### О пользе графиков

Выбор подходящих преобразований предикторов иногда можно сделать на основе анализа различных диаграмм данных; см., например, работы Хоэрла [172], Долби [92] и Тьюки [329].

Все вышесказанное касается выбора модельной функции. (Отклик  $Y$  остается неизменным.) Если считать, что предикторы не подвержены ошибкам, преобразовать их не составит труда.

Преобразования отклика рассматриваются в главе 14. В таких случаях следует особенно внимательно проверять, не нарушаются ли при выполнении данного преобразования предположения, используемые в методе наименьших квадратов, а именно: что ошибки независимы и подчиняются нормальному распределению  $N(0; \sigma^2)$ . Такая проверка обычно осуществляется путем анализа остатков в *преобразованной метрике*, после того как было выполнено преобразование и подобрана соответствующая модель.

### Использование отношений в качестве откликов и/или предикторов

Когда в регрессионном анализе используются отношения с общей компонентой, существует опасность, что эта компонента внесет ложную сильную регрессионную зависимость. Первые предостережения на этот счет появились в литературе почти сто лет назад, но проблемы продолжают возникать. Обсуждение возможных проблем содержится в работе Кронмела [207]. Раздел заключений и рекомендаций в его книге (сс. 390–391) начинается такими словами: «Идея данной статьи в том, что отношения переменных следует использовать только в полной линейной модели, в которую входят составляющие отношение переменные и свободный член. Распространенная в регрессионном анализе практика использования отношений только для отклика или предикторов может приводить к неверным выводам и редко что-то дает». Далее Кронмел предлагает несколько советов тем исследователям, которые не желают «отказаться от своего самого ценного отношения или индекса».

## 13.2. ПРИМЕРЫ ПОДБОРА ПОВЕРХНОСТЕЙ ВТОРОГО ПОРЯДКА ДЛЯ ДВУХ И ТРЕХ ПРЕДИКТОРОВ

В статье Эйа, Голдсмита и Муни [3] «Predicting Stoichiometric  $\text{CaHPO}_4 \cdot 2\text{H}_2\text{O}$ » описывается исследование, проведенные в рамках производственного эксперимента. В этом разделе мы используем материалы данной статьи с разрешения Американского химического общества. Мы решили не останавливаться на химических аспектах задачи и внесли некоторые изменения в исходный анализ.

В рассматриваемом эксперименте имелось семь потенциальных предикторных переменных, однако значения четырех из них оставались постоянными. Три переменные, выбранные для изучения поверхности отклика, а также диапазоны их изменения описаны ниже.

Переменная	Обозначение	Диапазон
Молярное отношение $\text{NH}_3/\text{CaCl}_2$ в растворе хлорида кальция	$r$	0,70–1,00
Дополнительное время (в минутах) превращения аммиачного $\text{CaCl}_2$ в $\text{NH}_4\text{H}_2\text{PO}_4$	$t$	10–90
Начальное рН раствора $\text{NH}_4\text{H}_2\text{PO}_4$	рН	2–5

Представляющих интерес переменных отклика было семь. Для каждой из них была выдвинута гипотеза, что ее можно представить как функцию второго порядка от  $r$ ,  $t$  и рН. Мы рассмотрим только первую переменную отклика (которую обозначим через  $Y$ ); приведем также значения четвертого отклика ( $Y_4$ ), которые будут использоваться в упражнениях.

В качестве плана эксперимента была выбрана схема «куб плюс звезда плюс шесть центральных точек» со звездным плечом  $\alpha = \frac{5}{3} = 1,667$ . (Чтобы план обладал свойством «ротабельности»<sup>2</sup>, плечо  $\alpha$  должно иметь значение  $2^{3/4} = 1,6818$ , так что выбранный план близок к ротабельному, но не точно ротабельный. Авторы статьи пренебрегли этим обстоятельством и фактически использовали в своих вычислениях  $\alpha = 1,6818$ , поэтому наши вычисления несколько отличаются от авторских.)

В плане эксперимента предусмотрены пять уровней для каждой переменной. Переменные кодировались при помощи преобразований:

$$\begin{aligned} X_1 &= (r - 0,85)/0,09; \\ X_2 &= (t - 59)/24; \\ X_3 &= (\text{рН} - 3,5)/0,9. \end{aligned} \tag{13.2.1}$$

Таким образом, соответствие фактических и кодированных уровней можно охарактеризовать с помощью следующей таблицы:

<sup>2</sup>«Ротабельность» означает инвариантность относительно вращения системы координат; планы, обладающие этим свойством, называют *ротабельными*. В случае использования ротабельных планов дисперсия предсказанного значения отклика не зависит от направления от центральной точки плана, в котором ведется предсказание, а зависит только от расстояния до центра. — Прим. ред.

Кодированные уровни $X_1, X_2, X_3$	Фактические уровни		
	$r$	$t$	pH
$\frac{5}{3}$	1,00	90	5,0
1	0,94	74	4,4
0	0,85	50	3,5
-1	0,76	26	2,6
$-\frac{5}{3}$	0,70	10	2,0

Данные реального плана эксперимента в кодированных единицах выделены в табл. 13.1 как матрица  $\mathbf{X}$ . Модель второго порядка для отклика  $Y$  в этих единицах можно записать следующим образом:

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{23} X_2 X_3. \quad (13.2.2)$$

Модель для отклика  $Y_4$  имеет такой же вид.

**Таблица 13.1.** Рабочий пример: матрица  $X$  и два отклика

$u$	Матрица $\mathbf{X}$									$Y$	$Y_4$	
	1	$X_1$	$X_2$	$X_3$	$X_1^2$	$X_2^2$	$X_3^2$	$X_1 X_2$	$X_1 X_3$			$X_2 X_3$
1	1	-1	-1	-1	1	1	1	1	1	1	52,8	6,95
2	1	1	-1	-1	1	1	1	-1	-1	1	67,9	5,90
3	1	-1	1	-1	1	1	1	-1	1	-1	55,4	7,10
4	1	1	1	-1	1	1	1	1	-1	-1	64,2	7,08
5	1	-1	-1	1	1	1	1	1	-1	-1	75,1	5,64
6	1	1	-1	1	1	1	1	-1	1	-1	81,6	5,18
7	1	-1	1	1	1	1	1	-1	-1	1	73,8	6,84
8	1	1	1	1	1	1	1	1	1	1	79,5	5,67
9	1	$-\frac{5}{3}$	0	0	$\frac{25}{9}$	0	0	0	0	0	68,1	6,00
10	1	$\frac{5}{3}$	0	0	$\frac{25}{9}$	0	0	0	0	0	91,2	5,67
11	1	0	$-\frac{5}{3}$	0	0	$\frac{25}{9}$	0	0	0	0	80,6	5,52
12	1	0	$\frac{5}{3}$	0	0	$\frac{25}{9}$	0	0	0	0	77,5	6,47
13	1	0	0	$-\frac{5}{3}$	0	0	$\frac{25}{9}$	0	0	0	36,8	7,17
14	1	0	0	$\frac{5}{3}$	0	0	$\frac{25}{9}$	0	0	0	78,0	5,36
15	1	0	0	0	0	0	0	0	0	0	74,6	6,48
16	1	0	0	0	0	0	0	0	0	0	75,9	5,91
17	1	0	0	0	0	0	0	0	0	0	76,9	6,39
18	1	0	0	0	0	0	0	0	0	0	72,3	5,99
19	1	0	0	0	0	0	0	0	0	0	75,9	5,86
20	1	0	0	0	0	0	0	0	0	0	79,8	5,96

Опыты выполнялись в произвольном порядке; наблюдаемые значения семи откликов записывались. Значения двух переменных отклика:  $Y$  — выход как процент теоретического выхода и  $Y_4$  — насыпная плотность (грамм на кубический дюйм), приведены в табл. 13.1. Матрица  $\mathbf{X}'\mathbf{X}$  не зависит от подбираемого отклика и имеет вид

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} N & 0 & 0 & 0 & B & B & B & 0 & 0 & 0 \\ 0 & B & 0 & 0 & & & & & & \\ 0 & 0 & B & 0 & & \mathbf{0} & & & & \mathbf{0} \\ 0 & B & 0 & 0 & & & & & & \\ B & & & & C & D & D & & & \\ B & & \mathbf{0} & & D & C & D & & & \mathbf{0} \\ B & & & & D & D & C & & & \\ 0 & & & & & & & D & 0 & 0 \\ 0 & & \mathbf{0} & & & \mathbf{0} & & 0 & D & 0 \\ 0 & & & & & & & 0 & 0 & D \end{bmatrix}. \quad (13.2.3)$$

В нашем случае

$$\begin{aligned} N &= 20, & C &= 8 + 2\alpha^4 = \frac{1898}{81}, \\ B &= 8 + 2\alpha^2 = \frac{122}{9}, & D &= 8. \end{aligned} \quad (13.2.4)$$

Матрицы такого вида часто встречаются при исследовании поверхности отклика второго порядка с помощью спланированных экспериментов; для них достаточно легко найти обратную. В общем случае, когда имеются  $k$  предикторов, матрицу  $(\mathbf{X}'\mathbf{X})^{-1}$ , используя те же обозначения, можно записать как

	0	1	2	...	$k$	11	22	...	$kk$	12	13	...	$(k-1)k$	
$P$	0	0	0	...	0	$Q$	$Q$	...	$Q$	0	0	...	0	0
0	$1/B$	0	...	0										1
0	0	$1/B$	...	0										2
⋮	⋮								$\mathbf{0}$					⋮
0	0	0	...	$1/B$										$k$
$Q$						$R$	$S$	...	$S$					11
$Q$						$S$	$R$	...	$S$					12
⋮						⋮								⋮
$Q$						$S$	$S$	...	$R$					$kk$
0										$1/D$	0	...	0	12
0										0	$1/D$	...	0	13
⋮										⋮				⋮
0										0	0	...	$1/D$	$(k-1)k$

(13.2.5)

Значения  $P$ ,  $Q$ ,  $R$  и  $S$  представлены во втором столбце, обозначенном  $C \neq 3D$ , табл. 13.2. (Значения в третьем столбце отвечают упрощенной

форме, когда  $C = 3D$ , что происходит в случае ротабельного плана, т.е. когда контуры<sup>3</sup> функции  $V\{\hat{Y}(X)\}$  являются сферическими. В таких случаях поворот плана в  $X$ -пространстве не влияет на точность получаемой информации.)

В нашем случае  $k = 3$ ,  $3D = \frac{1944}{81}$ ; таким образом, величина  $C - 3D$  мала, но не равна нулю, как для ротабельного плана. Следовательно, мы получаем

$$\begin{aligned} P &= 1597/9614, & Q &= -549/9614, \\ R &= 685,3248/9614, & S &= 62,3376/9614. \end{aligned} \quad (13.2.6)$$

При этом  $1/B = \frac{9}{122}$ ,  $1/D = \frac{1}{8}$ . (Это точные значения; чтобы в дальнейшем избежать ошибок округления, деление на 9614 откладывается до момента, когда будет выполнено умножение матриц.) Теперь нужно найти  $\mathbf{X}'\mathbf{Y}$ . Эти вычисления выполнены ниже для откликов  $Y$  и  $Y_4$ .

$$\mathbf{X}'\mathbf{Y} = \frac{1}{9} \begin{bmatrix} 12941,1 \\ 671,4 \\ -87,0 \\ 1245,3 \\ 8935,2 \\ 8905,2 \\ 7822,7 \\ -63,9 \\ -105,3 \\ -20,7 \end{bmatrix}, \quad \mathbf{X}'\mathbf{Y}_4 = \frac{1}{9} \begin{bmatrix} 1108,26 \\ -29,25 \\ 41,43 \\ -60,45 \\ 744,99 \\ 752,99 \\ 766,49 \\ 2,88 \\ -5,04 \\ 3,24 \end{bmatrix}. \quad (13.2.7)$$

Применив стандартную формулу  $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$  для оценивания параметров регрессии первого отклика  $Y$ , получаем следующее подобранное уравнение:

$$\begin{aligned} \hat{Y} &= 76,022 + 5,503X_1 - 0,713X_2 + 10,207X_3 + 0,712X_1^2 + \\ &+ 0,496X_2^2 - 7,298X_3^2 - 0,888X_1X_2 - 1,463X_1X_3 - 0,288X_2X_3. \end{aligned} \quad (13.2.8)$$

**Таблица 13.2.** Формулы для получения элементов матрицы  $(\mathbf{X}'\mathbf{X})^{-1}$

Символ	Значение при $C \neq 3D$	Значение при $C = 3D$ (ротабельный план)
$P$	$(C - D)(C + (k - 1)D)/A$	$2(k + 2)D^2/A$
$Q$	$-(C - D)B/A$	$-2DB/A$
$R$	$\{N(C + (k - 2)D) - (k - 1)B^2\}/A$	$\{N(k + 1)D - (k - 1)B^2\}/A$
$S$	$(B^2 - ND)/A$	$(B^2 - ND)/A$
$A$	$(C - D)\{N(C + (k - 1)D) - kB^2\}$	$2D\{N(k + 2)D - kB^2\}$

*Примечание.* Обратите внимание, что  $A$  присутствует в формулах для  $P$ ,  $Q$ ,  $R$ ,  $S$ .

<sup>3</sup>Под «контуром» функции многих переменных здесь и далее подразумевается ее линия (или поверхность) уровня. — Прим. ред.

Таблица 13.3. ТДА для планов второго порядка определенного вида

Источник	СС	СК
$b_0$ (среднее)	1	$(\sum_{u=1}^N Y_u)^2 / N$
$b_i$ (члены первого порядка)	$k$	$\sum i = 1^k b_i(iY)$
$b_{ii} b_0$ («чистые» члены второго порядка при данном $b_0$ )	$k$	$b_0(0Y) + \sum_{i=1}^k b_{ii}(iiY) - (\sum_{u=1}^N Y_u)^2 / N$
$b_{ij}$ («смешанные» члены второго порядка)	$\frac{1}{2}k(k-1)$	$\sum_{i=1}^k \sum_{j=1}^k b_{ij}(ijY)$
Неадекватность	$N - n_e - \frac{1}{2}(k+1)(k+2)$	результат вычитания
Чистая ошибка	$n_e$	результат обычных вычислений
Общий	$N$	$\sum_{u=1}^N Y_u^2$

Таблица дисперсионного анализа для случая, когда план второго порядка приводит к такой матрице  $\mathbf{X}'\mathbf{X}$ , что обратная ей матрица имеет вид (13.2.5), представлена в табл. 13.3. В данном случае

$$\begin{aligned} (0y) &= \sum_{u=1}^N Y_u, & (i iy) &= \sum_{u=1}^N X_{iu}^2 Y_u, \\ (iy) &= \sum_{u=1}^N X_{iu} Y_u, & (ij y) &= \sum_{u=1}^N X_{iu} X_{ju} Y_u. \end{aligned} \quad (13.2.9)$$

Все эти выражения являются перекрестными произведениями столбцов матрицы  $\mathbf{X}$  и столбца наблюдений  $\mathbf{Y}$ , т.е. элементами вектора  $\mathbf{X}'\mathbf{Y}$ . Обычно, комбинируя СК( $b_{ii}|b_0$ ) и СК( $b_{ij}$ ), мы получаем

$$\text{СК(члены второго порядка}|b_0) \text{ с } \frac{1}{2}k(k+1) \text{ степенями свободы,}$$

однако мы указали их в таблице отдельно, чтобы подчеркнуть, что единственная *дополнительная сумма квадратов* — СК( $b_{ii}|b_0$ ), что объясняется ортогональностью многих пар столбцов матрицы  $\mathbf{X}$ . Подобным свойством обладают лишь определенные планы экспериментов. Теперь мы можем продолжать работу обычным образом, проверяя неадекватность и полезность членов первого и второго порядков.

Следует отметить, что многие специальные приемы МНК-оценивания и дисперсионного анализа применимы только для планов, в которых матрица  $\mathbf{X}'\mathbf{X}$  имеет указанный выше вид, так что обратная матрица,  $(\mathbf{X}'\mathbf{X})^{-1}$ , может быть вычислена по формуле (13.2.5). Планы, не обладающие этим свойством, должны подвергаться стандартному МНК-анализу, в котором не используются специальные формулы. Тем не менее, столбец «Источник» в приведенной таблице дисперсионного анализа может служить образцом: сумма квадратов чистой ошибки получается обычным способом, а все последующие элементы сумм квадратов для оценок параметров вычисляются как дополнительные суммы квадратов (см. главу 7).

Результаты дисперсионного анализа для нашего примера представлены в табл. 13.4. Поскольку  $F(5; 5; 0,95) = 5,05 > 3,04$ , неадекватность не выявлена. Можно объединить суммы квадратов, обусловленные неадекватностью и чистой ошибкой, и оценить  $V(Y_i) = \sigma^2$  величиной

$$s^2 = (93,91 + 30,86)/(5 + 5) = 12,477.$$

Разделив на эту величину средний квадрат, обусловленный членами первого порядка, получим значение  $609,93/12,477 = 48,88$ , которое превышает  $F(3; 10; 0,999) = 12,55$ . В то же время для среднего квадрата, обусловленного членами второго порядка, получается значение  $135,59/12,477 = 10,88$ , превышающее  $F(6; 10; 0,999) = 9,93$ . Таким образом, в подобранной модели необходимы члены как первого, так и второго порядков.

### Нужна ли нам переменная $X_2$ ?

В статье-первоисточнике авторы обращают внимание на малую величину коэффициентов с индексом 2 по сравнению с их стандартными ошибками и пришли к выводу, что модель не должна содержать переменную  $X_2$ . Когда ситуация настолько определенная — все коэффициенты малы по сравнению с их стандартными ошибками — подобное заключение вряд ли будет ошибочным. Однако в общем случае необходимо применять принцип дополнительной суммы квадратов; мы так и поступим, чтобы продемонстрировать, как это делается.

Предположим, мы хотим проверить нуль-гипотезу  $H_0: \beta_2 = \beta_{22} = \beta_{12} = \beta_{23} = 0$ ; альтернативная гипотеза  $H_1$  состоит в том, что хотя бы один из этих коэффициентов  $\beta$  не равен нулю. Из таблицы дисперсионного анализа сумма квадратов регрессии для полной модели второго порядка, содержащей  $X_1$ ,  $X_2$  и  $X_3$ , при заданном  $b_0$  составляет

$$S_1 = \text{СК(члены первого порядка)} + \text{СК(члены второго порядка)} | b_0 = \\ = 1829,80 + 813,54 = 2643,34 \quad (\text{с } 3 + 6 = 9 \text{ степенями свободы}).$$

**Таблица 13.4.** ТДА для подобранной модели

Источник	СС	СК	СрК	$F$
Среднее ( $b_0$ )	1	103377,82		
Члены первого порядка	3	1829,80	609,93	
Члены второго порядка   $b_0$	6	813,54	135,59	
Неадекватность	5	93,91	18,78	3,04
Чистая ошибка	5	30,86	6,17	
Общий	20	106145,93		

Принятие гипотезы  $H_0$  в отношении исходной модели означает переход к использованию сокращенной модели:

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \beta_{11} X_1^2 + \beta_{33} X_3^2 + \beta_{13} X_1 X_3.$$

Соответствующая матрица  $\mathbf{X}$  может быть получена из табл. 13.1 удалением столбцов  $X_2$ ,  $X_2^2$ ,  $X_1 X_2$  и  $X_2 X_3$ . Матрица  $\mathbf{X}'\mathbf{X}$  получается из ранее представленной матрицы  $\mathbf{X}'\mathbf{X}$  удалением строк и столбцов, соответствующих  $X_2$ ,  $X_2^2$ ,  $X_1 X_2$  и  $X_2 X_3$ . Точно также, вектор  $\mathbf{X}'\mathbf{Y}$  получается из приведенного ранее удалением соответствующих строк. Подобранный уравнение принимает вид:

$$\hat{Y} = 76,420 + 5,503X_1 + 10,207X_3 + 0,667X_1^2 - 7,343X_3^2 - 1,463X_1X_3.$$

Теперь нужно вычислить сумму квадратов регрессии при заданном  $b_0$  для данной сокращенной модели. Находим, что она равна  $S_2 = 2626,025$  (с 5 СС). Следовательно, дополнительная сумма квадратов, связанная с  $b_2$ ,  $b_{22}$ ,  $b_{12}$ ,  $b_{23}$ , равна

$$S_1 - S_2 = 2643,34 - 2626,03 = 17,31 \quad (\text{с } 9 - 5 = 4 \text{ степенями свободы}).$$

Таким образом, средний квадрат составляет  $17,31/4 = 4,33$ , и его можно сравнить со среднеквадратичной оценкой  $\sigma^2$  исходной регрессии с тремя предикторами. Нуль-гипотеза, состоящая в предположении, что  $\beta_2 = \beta_{22} = \beta_{12} = \beta_{23} = 0$ , не может быть отвергнута. Поэтому представляется разумным принять сокращенную модель, не содержащую  $X_2$ . Табл. 13.5 представляет собой таблицу дисперсионного анализа, соответствующую этой сокращенной подобранной модели. Дисперсионный анализ не выявляет неадекватности, и регрессия высоко значима как для членов первого порядка, так и для членов второго.

Чтобы исследовать подобранную поверхность отклика второго порядка, мы обычно производим «канонический анализ»: поверхность описывается в координатах, расположенных вдоль ее основных осей. Такой анализ чрезвычайно полезен и позволяет оценить ситуацию в целом, даже при большом

**Таблица 13.5.** ТДА для модели второго порядка с переменными  $X_1$  и  $X_3$

Источник	СС	СК	СрК	$F$
Первый порядок	2	1822,91	911,46	89,80
Второй порядок   $b_0$	3	803,12	267,71	26,38
Неадекватность	9	111,22	12,36	10,15
Чистая ошибка	5	30,86	6,17	
Общий скорректированный	19	2768,11		

количестве предикторных переменных. Когда же предикторов всего два, как в данном случае, можно непосредственно построить на графике линии уровней  $\hat{Y}$ , записав подобранное уравнение в виде:

$$-7,343X_3^2 + (10,207 - 1,463X_1)X_3 + (0,667X_1^2 + 5,503X_1 + 76,420 - \hat{Y}) = 0.$$

Выбрав некое значение  $\hat{Y}$ , можно построить соответствующий контур, подставляя значения  $X_1$  и решая данное уравнение относительно  $X_3$ . Полученные таким способом контуры представлены на рис. 13.1. Экспериментальные точки показаны на графике маленькими кружками. Кратные точки не выделены; чтобы выявить их, нужно обратиться к табл. 13.6. Горизонталы образуют картину поднимающегося гребня. Исследование этой системы контуров позволило авторам выдвинуть гипотезу о том, какие химические реакции могут лежать в основе такого поведения переменных. (Достаточно часто исследования поверхности отклика являются первым шагом в детальном теоретическом исследовании изучаемой системы.)

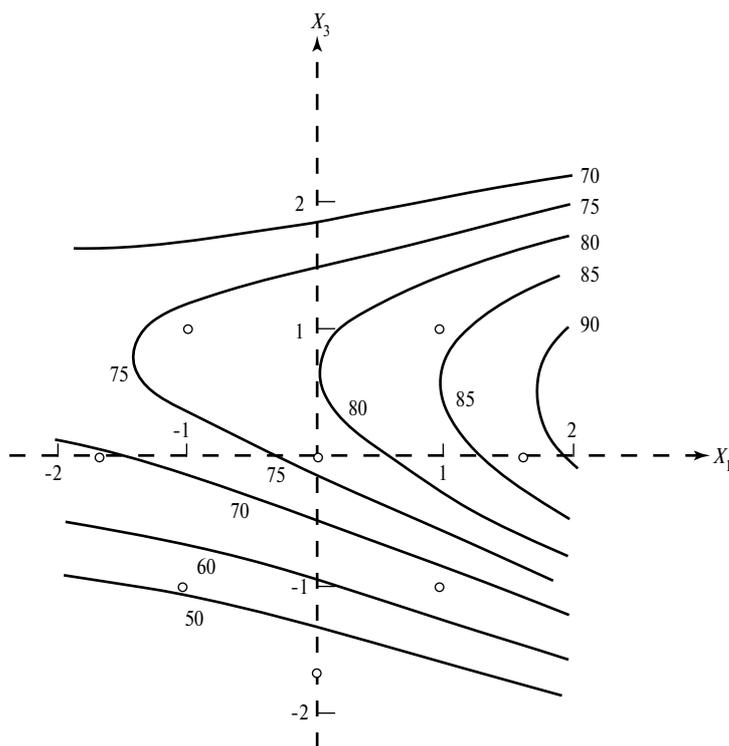
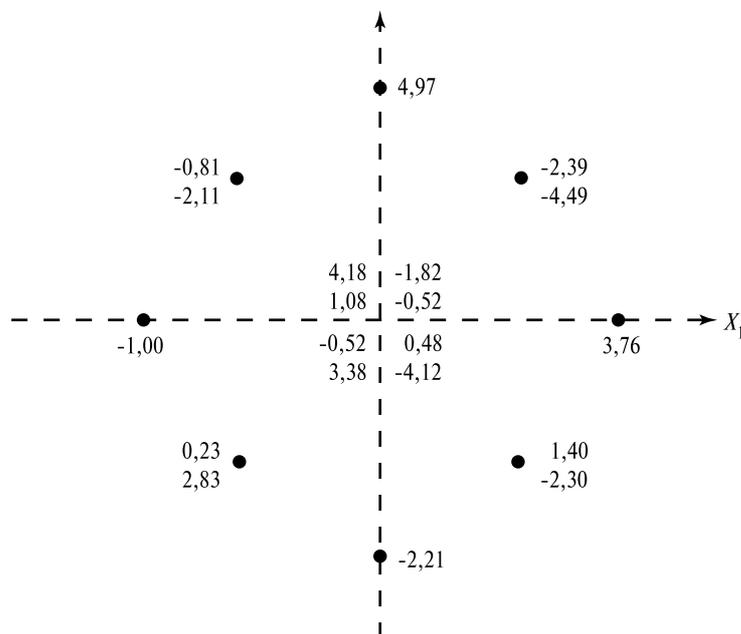


Рис. 13.1. Горизонталы подобранного уравнения второго порядка, связывающего отклик  $\hat{Y}$  с переменными  $X_1$  и  $X_3$

**Таблица 13.6.** Подобранные значения и остатки, полученные в результате подбора поверхности второго порядка  $\hat{Y} = f(X_1, X_3)$ 

$u$	$X_1$	$X_3$	$Y$	$\hat{Y}$	$e = Y - \hat{Y}$
1	-1	-1	52,8	52,57	0,23
2	1	-1	67,9	66,50	1,40
3	-1	-1	55,4	52,57	2,83
4	1	-1	64,2	66,50	-2,30
5	-1	1	75,1	75,91	-0,81
6	1	1	81,6	83,99	-2,39
7	-1	1	73,8	75,91	-2,11
8	1	1	79,5	83,99	-4,49
9	$-\frac{5}{3}$	0	68,1	69,10	-1,00
10	$\frac{5}{3}$	0	91,2	87,44	3,76
11	0	0	80,6	76,42	4,18
12	0	0	77,5	76,42	1,08
13	0	$-\frac{5}{3}$	36,8	39,01	-2,21
14	0	$\frac{5}{3}$	78,0	73,04	4,97
15	0	0	74,6	76,42	-1,82
16	0	0	75,9	76,42	-0,52
17	0	0	76,9	76,42	0,48
18	0	0	72,3	76,42	-4,12
19	0	0	75,9	76,42	-0,52
20	0	0	79,8	76,42	3,38

Эти контуры можно также рассматривать совместно с остатками, приведенными в табл. 13.6. График, на котором каждый остаток помещен рядом с соответствующей точкой плана, приведен на рис. 13.2. Среди двадцати остатков шесть с наибольшими абсолютными значениями соответствуют точкам  $(X_1; X_3) = (0; 0)$  (три остатка);  $(\frac{5}{2}; 0)$ ,  $(0; \frac{5}{2})$ ,  $(1; 1)$  (по одному остатку). Получается, что модель хуже всего соответствует данным в первом квадранте плоскости  $(X_1; X_3)$ , и все заключения, основанные на правильности подобранной поверхности, в этой области могут подвергаться сомнению. (Как это может отразиться на заключениях авторов статьи-первоисточника — судить химику-технологу, а не специалисту по статистике, поэтому мы не будем здесь обсуждать этот вопрос.) Иногда сомнения такого рода можно разрешить, проведя дополнительные опыты в области, где форма подобранной поверхности сомнительна, и подобрав для нее подходящую функцию отклика.

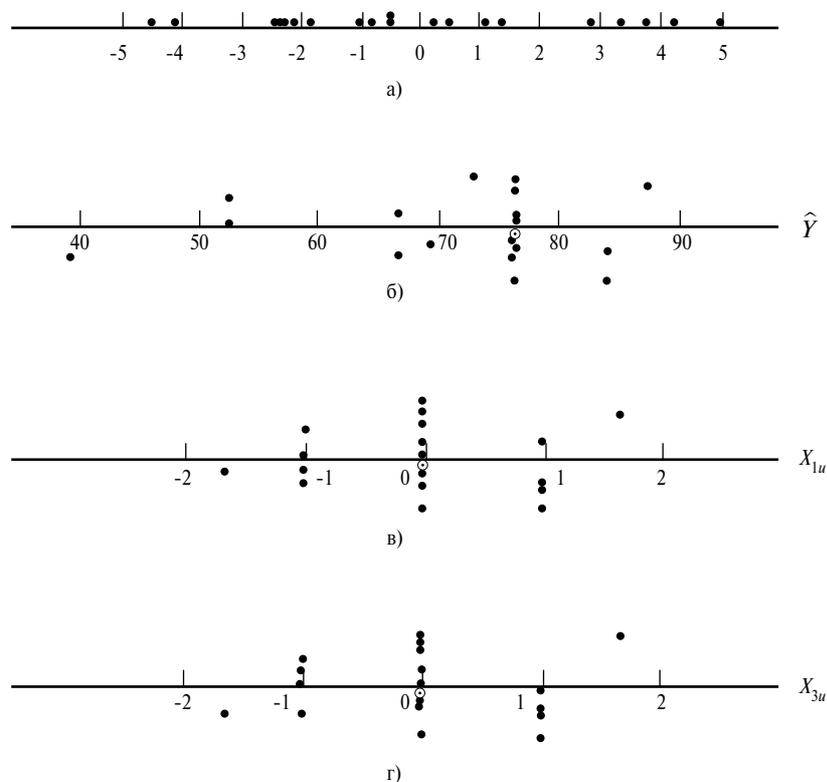


**Рис. 13.2.** «Структура» графика остатков для модели, связывающей отклик  $Y$  с факторами  $X_1$  и  $X_3$

Теперь мы можем исследовать остатки другими способами, чтобы выявить любые отклонения от нормальности, если они существуют. На рис. 13.3 приведены стандартные графики остатков.

Общий вид графика не противоречит предположению о нормальности, которое неявно используется в ходе дисперсионного анализа. График зависимости остатков от  $\hat{Y}_u$  обнаруживает тенденцию к «расширению» в первой точке, но это обманчиво, поскольку остатки в основном велики и диапазон значений остатков при малых значениях  $\hat{Y}_u$  определен не достаточно четко. Аналогичная картина наблюдается на графиках зависимости от  $X_{1u}$  и  $X_{3u}$ , поскольку диапазон значений остатков в крайних точках определен нечетко. Таким образом, ни один график не позволяет утверждать о существовании отклонений от нормальности. Следовательно, базовые предположения, судя по всему, не нарушаются. (Поскольку нам не известен порядок проведения наблюдений, мы не можем проверить, влияет ли на отклик временной тренд.)

Если бы мы хотели продолжить исследование, следовало бы попытаться учесть большие остатки в первом квадранте, повторно изучить исходные данные и их взаимосвязи с другими предикторами, изменения которых, возможно, не были учтены. Таким образом можно добиться улучшения модели. Кроме этого (или вместо этого), можно, как предлагалось ранее, более тщательно изучить область, в которой подбор модели вызывает сомнения.



**Рис. 13.3.** Стандартные графики остатков, полученных в результате подбора уравнения второго порядка, связывающего отклик  $Y$  с переменными  $X_1$  и  $X_3$ : а) общий вид, б) зависимость от  $\hat{Y}_u$ , в) зависимость от  $X_{1u}$ , г) зависимость от  $X_{3u}$

### Интерпретация чистой ошибки при исключении факторов

Теперь мы попытаемся ответить на вопрос, который избегали задавать ранее: «Следует ли пересматривать интерпретацию истинной ошибки, когда из модели исключается фактор (например,  $X_2$ )?» В табл. 13.1 повторами являются только опыты 15–20, но после исключения  $X_2$  данные выглядят так, как в табл. 13.6, и повторами по переменным  $X_1$  и  $X_3$  являются еще четыре пары опытов: (1; 3), (2; 4), (5; 7) и (6; 8). Кроме того, опыты 11 и 12, очевидно, являются центральными точками. Таким образом, план теперь следует рассматривать как факторный эксперимент  $2^2$  с повторами (восемь точек 1–8) плюс двухфакторная звезда (четыре точки: 9, 10, 13, 14) плюс восемь центральных точек (11, 12, 15, ..., 20).

В этом случае некоторые элементы табл. 13.5 также изменятся:

СК, обусловленная неадекватностью = 78,26 (3 СС),  $CpK = 26,08$ ;

СК, обусловленная чистой ошибкой = 62,82 (11 СС),  $CpK = 5,80$ .

Соответствующее  $F$ -отношение равно  $4,50 > F(3; 11; 0,95) = 3,59$ ; это приводит нас к несколько неожиданному заключению о наличии неадекватности. Казалось бы, присутствие в модели переменной  $X_2$  не является необходимым. Но после ее исключения проявляется неадекватность. Однако ясно, что  $X_2$  мало помогает в объяснении дисперсии наблюдений. Фактически, когда в модели *используется*  $X_2$ , величина среднего квадрата, обусловленного чистой ошибкой, увеличивается. В то же время уменьшение числа степеней свободы чистой ошибки влечет уменьшение чувствительности  $F$ -критерия при проверке модели на адекватность.

Какой же анализ дает правильные результаты? Можно привести аргументы в пользу обоих подходов. Однако мы предпочитаем использовать чистую ошибку, вычисленную до исключения факторов. Предполагается, что повторы в исходных данных являются истинными, чего нельзя сказать об экспериментах, которые выглядят как повторы после исключения одного или нескольких факторов. Кроме того, во многих наборах данных может возникать обратная проблема: реально существующая неадекватность может не обнаруживаться после исключения факторов, поскольку «новые» повторы будут демонстрировать более сильную изменчивость в сравнении с истинными.

Надежнее всего провести анализ обоими способами и проверить, согласуются ли результаты. Для многих наборов данных они будут согласованными. В противном случае данные можно подвергнуть дальнейшей проверке.

В нашем примере мы предпримем следующие действия. Прежде всего необходимо подвергнуть сомнению вид модели. Однако вошедшие в нее члены первого и второго порядка отвечают за  $R^2 = (1822,91 + 803,12) / 2768,11 = 0,949$  полного (скорректированного на среднее) отклонения данных при потере всего пяти степеней свободы. (При включении членов, содержащих  $X_2$ , этот показатель возрастает всего лишь до 0,995.) Иными словами, данная модель объясняет 95% отклонения от среднего значения, хотя технически неадекватность возможна. Исследуя, как и выше, совместно подобранные контуры и остатки, можно обнаружить, где может иметь место неадекватность. Если сглаживание истинной поверхности выглядит удовлетворительно в большой области  $X$ -пространства, полученные на основании подобранной модели заключения для данной области вполне могут оставаться верными. Дальнейшее исследование остатков может также показать, какие из базовых предположений (нормальность, постоянство дисперсии, независимость наблюдений) нарушаются, или подсказать альтернативную модель.

Иногда чистая ошибка оказывается «достаточно малой» просто потому, что при проведении эксперимента опыты, на основании которых определяется чистая ошибка, не были рандомизированы (или, как минимум, распределены). Если несколько повторяющихся опытов по определению чистой ошибки прово-

дятся последовательно через короткие промежутки времени, то отклики могут оказаться более похожими, чем при других обстоятельствах. Иными словами, чистая ошибка в таком случае не представляет диапазон ошибок, характерный для всего эксперимента. Иногда это приводит к ложному заключению о наличии неадекватности и необходимости более тщательного исследования. Повторный анализ тех же опытов иногда также дает «правильные повторы».

*Комментарий.* Приведенный пример несколько необычен. После того как по веским соображениям из модели исключаются некоторые члены, в сокращенной модели редко обнаруживается неадекватность, за исключением случаев, когда в данных существуют особенности. Как мы могли убедиться, эти особенности, судя по всему, не связаны с нарушением базовых предположений метода наименьших квадратов. Их источник остается предметом обсуждения.

В более широком смысле этот пример не является необычным. Эксперимент позволил ответить на ряд вопросов, но некоторые все же остались не выясненными. Они становятся предметом дальнейшего исследования. Такая ситуация вполне типична в практике экспериментальных работ.

### Интерпретация чистой ошибки в планах, разбитых на блоки

В плане эксперимента, приведенном в табл. 13.6, разбиение данных на блоки отсутствует. Однако часто для построения поверхностей отклика используются планы с разбиением на блоки, ортогональные модели. Опыты, которые были бы повторами в плане без разбиения на блоки, обычно распределяются по блокам. В таком случае, эти опыты не являются более повторами, *если не принадлежат одному блоку*, и чистая ошибка должна вычисляться на этой основе. Кроме того, дисперсионный анализ должен содержать сумму квадратов для блоков. Поскольку блоки ортогональны модели, соответствующая сумма квадратов для блоков в таблице дисперсионного анализа обычно находится как

$$\text{СК(блоков)} = \sum_{w=1}^m \frac{B_w^2}{n_w} - \frac{G^2}{N}, \quad \text{с } (m-1) \text{ степенями свободы,}$$

где  $m$  — общее число блоков,  $B_w$  — сумма  $n_w$  наблюдений в блоке  $w$ , а  $G$  — сумма всех наблюдений по всем  $m$  блокам. Если блоки не ортогональны модели, применяется принцип дополнительной суммы квадратов. (Естественно, он применим и в ортогональном случае; при этом результат будет тот же, что приведен выше.)

В некоторых случаях даже опыты, относящиеся к *разным* блокам, могут использоваться для измерения «почти чистой» ошибки при условии, что известны оценки разностей между блоками. См., например, работу Бокса и Дрейпера [41] (с. 375).

### Исключение членов

В нашем примере с полиномиальной моделью второго порядка мы решили исключить все члены, содержащие  $X_2$ . Однако возникает более общий вопрос: «Какие члены можно исключить, не нарушая связи между членами модели?» (Может быть, разумно было бы исключить, например, только член  $X_2$ ? Ответ на этот вопрос — определенно отрицательный.) В разделе 13.3 мы продолжим обсуждение поставленного вопроса и предложим два критерия, которыми следует руководствоваться при исключении членов.

## 13.3. СОХРАНЕНИЕ ЧЛЕНОВ В ПОЛИНОМИАЛЬНОЙ МОДЕЛИ

В данном разделе мы покажем, что в общем случае отдельные члены не нужно исключать из полиномиальной модели второго и более высокого порядка без тщательного анализа ситуации. Мы предложим два критерия и два вытекающих из них правила, которые рекомендуем использовать. Из всего сказанного следует, что использовать уравнение, полученное с помощью некой механической процедуры для отбора членов полиномиальной модели, достаточно рискованно.

### Пример 1. Квадратное уравнение относительно $X$

Рассмотрим квадратичную модель с одним предиктором:

$$\hat{Y} = b_0 + b_1X + b_{11}X^2. \quad (13.3.1)$$

Максимум или минимум этого квадратичного выражения достигается в точке, где первая производная  $\hat{Y}$  по  $X$  равна нулю. Это имеет место там, где выполняется условие  $b_1 + 2b_{11}X = 0$ , т.е. при  $X = -b_1/(2b_{11})$ . Если значение  $b_{11}$  положительно, мы получим минимум, если отрицательно — максимум.

Рассмотрим последствия исключения одного члена. (Естественно, в этом случае нужно заново подобрать уравнение и получить новые коэффициенты.)

- а) Исключение  $b_0$ . Ранее свободный член при  $X = 0$  был равен  $\hat{Y} = b_0$ . В новом уравнении он должен быть равен нулю.
- б) Исключение  $b_1$ . Максимум или минимум функции находился в точке  $X = -b_1/(2b_{11})$ . Теперь он должен находиться в нуле.
- в) Исключение  $b_{11}$ . Первоначально мы имели квадратичную функцию со свободным членом  $b_0$ . Теперь мы получим прямую линию.

Даже если оценки  $b_0$  или  $b_1$  не являются статистически значимыми, действия (а) и (б) обычно приводят к искажению повторного подбора. Однако если оценка  $b_{11}$  оказывается статистически незначимой, т.е. искривление линии

регрессии незначительно, полученная в результате исключения соответствующего члена прямая в большинстве случаев будет вполне удовлетворительной моделью.

Таким образом, мы не рекомендуем исключать ни  $b_0$ , ни  $b_1$ , если  $b_{11}$  остается в уравнении, поскольку это приведет к тому, что подобранная квадратичная функция будет ограниченной и смещенной. В то же время, если искривление выражено слабо, исключение  $b_{11}$  вполне обоснованно, поскольку полученная в результате прямая будет хорошим приближением (в противном случае  $b_{11}$  было бы большим). Те же соображения справедливы и для моделей, содержащих несколько предикторных переменных.

Чтобы понять, какими будут последствия трех разных исключений, можно рассмотреть, что происходит с сокращенными уравнениями при переносе начала координат. Пусть  $Z = X - a$ , что соответствует переносу начала координат из точки  $X = 0$  в точку  $X = a$ . Подставив  $X = Z + a$  в выражения (а)–(в), получим:

- а)  $b_1X + b_{11}X^2 = a(b_1 + ab_{11}) + a(b_1 + 2b_{11})Z + b_{11}Z^2$ ;
- б)  $b'_0 + b'_{11}X^2 = (b'_0 + a^2b'_{11}) + 2ab'_{11}Z + b'_{11}Z^2$ ;
- в)  $b''_0 + b''_1X = (b''_0 + ab''_1) + b''_1Z$ .

Как видим, только третье уравнение сохранило первоначальный вид. В двух других появились ранее исключенные члены! Этот пример служит основой первого критерия и первого правила. Ниже мы рассмотрим второй критерий и второе правило. В дальнейшем мы будем использовать модельную функцию с параметрами  $\beta$  вместо подобранного уравнения с оценками  $b$ , но это никак не влияет на справедливость наших рассуждений. Главный вопрос: «Какие сокращенные модели следует принимать, а какие — не следует?»

### Критерий 1. Критерий сдвига начала координат

Будем считать, что сокращенная модель является осмысленной (или «хорошо сформулированной»), если сдвиг начала координат в  $X$ -пространстве приводит к модели *неизменного вида* в новых переменных  $Z_1, Z_2, \dots, Z_k$ .

### Пример 2. Полином второго порядка с двумя переменными $X$

Рассмотрим квадратичную модель с двумя переменными:

$$\hat{Y} = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_{11}X_1^2 + \beta_{22}X_2^2 + \beta_{12}X_1X_2 + \epsilon. \quad (13.3.2)$$

Если в полную модель второго порядка (13.3.2) подставить  $X_1 = Z_1 + a_1$ ,  $X_2 = Z_2 + a_2$ , что соответствует переносу начала координат в точку  $(X_1; X_2) = (a_1; a_2)$ , получится уравнение

$$\hat{Y} = \beta_0 + \beta_1(Z_1 + a_1) + \beta_2(Z_2 + a_2) + \beta_{11}(Z_1 + a_1)^2 + \beta_{22}(Z_2 + a_2)^2 + \beta_{12}(Z_1 + a_1)(Z_2 + a_2) + \epsilon = \quad (13.3.3)$$

$$= \alpha_0 + \alpha_1 Z_1 + \alpha_2 Z_2 + \alpha_{11} Z_1^2 + \alpha_{22} Z_2^2 + \alpha_{12} Z_1 Z_2 + \epsilon, \quad (13.3.4)$$

где

$$\begin{aligned} \alpha_0 &= \beta_0 + \beta_1 a_1 + \beta_2 a_2 + \beta_{11} a_1^2 + \beta_{22} a_2^2 + \beta_{12} a_1 a_2 \\ &\quad - \text{квадратичная форма относительно } (a_1, a_2), \\ \alpha_1 &= \beta_1 + 2\beta_{11} a_1 + \beta_{12} a_2, \\ \alpha_2 &= \beta_2 + \beta_{12} a_1 + 2\beta_{22} a_2, \\ \alpha_{11} &= \beta_{11}, \\ \alpha_{22} &= \beta_{22}, \\ \alpha_{12} &= \beta_{12}. \end{aligned} \quad (13.3.5)$$

Какие члены можно исключить из (13.3.2), чтобы сокращенная модель (13.3.4) имела тот же вид? Иными словами, какие коэффициенты  $\beta$  в (13.3.2) можно приравнять к нулю, чтобы получить модель (13.3.4) того же вида? Очевидно, можно исключить только квадратичные члены. Например, положив  $\beta_{11} = 0$ , мы получим  $\alpha_{11} = 0$ , но остальные коэффициенты в обеих моделях останутся. (Мы игнорируем относительно маловероятный случай, когда значения  $a$  и  $\beta$  таковы, что один из коэффициентов  $\alpha$  также исчезает. Это происходит не при всех сдвигах начала координат и на практике встречается редко.)

Общая закономерность, которая следует из (13.3.5), состоит в том, что при переносе начала координат коэффициенты высшего порядка остаются неизменными, в то время как коэффициенты низших порядков в результате сдвига превращаются в комбинации коэффициентов как низших, так и высших порядков. На основании вышеизложенного можно сформулировать следующее правило.

### Правило 1

Если модель должна оставаться согласованной при сдвиге начала координат, сначала можно исключать только члены наивысшего порядка; при всех последующих исключениях модель должна оставаться хорошо сформулированной. Более того, если какие-либо члены высшего порядка оставлены в модели, все члены низших порядков, на которые они оказывают влияние после переноса начала координат, также должны быть сохранены, независимо от того, являются ли их оценки значимыми в подобранном регрессионном уравнении.  
*Примечание.* Модель без свободного члена ( $\beta_0$ ) ни при каких обстоятельствах не может быть хорошо сформулированной при переносе начала координат.

**Пример 2. Продолжение**

Предположим,  $\beta_{11} \neq 0$ ,  $\beta_{22} = 0$ ,  $\beta_{12} = 0$ . Тогда в силу (13.3.5)  $\beta_1 X_1$  необходимо оставить в модели, поскольку  $\alpha_1$  зависит от  $\beta_{11}$ . В то же время, член  $\beta_2 X_2$  можно исключить, поскольку  $\alpha_2 = \beta_2$ , когда  $\beta_{22} = \beta_{12} = 0$ .

**Пример 3. Полином третьего порядка с тремя факторами**

Чтобы рассмотреть данное правило в более широком контексте, мы приведем эквивалент формулы (13.3.5) для моделей второго и третьего порядков с тремя переменными  $(X_1; X_2; X_3)$  при переносе начала координат в точку  $(a_1; a_2; a_3)$ . Для кубической модели вида

$$\begin{aligned} \hat{Y} = & \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \\ & + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{33} X_3^2 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{23} X_2 X_3 + \\ & + \beta_{111} X_1^3 + \beta_{222} X_2^3 + \beta_{333} X_3^3 + \\ & + \beta_{122} X_1 X_2^2 + \beta_{133} X_1 X_3^2 + \beta_{112} X_1^2 X_2 + \beta_{233} X_2 X_3^2 + \beta_{113} X_1^2 X_3 + \\ & + \beta_{223} X_2^2 X_3 + \beta_{123} X_1 X_2 X_3 + \epsilon \end{aligned} \quad (13.3.6)$$

мы находим, что

$$\begin{aligned} \alpha_0 &= \text{полная кубическая форма относительно } (a_1, a_2, a_3), \\ \alpha_1 &= \beta_1 + 2\beta_{11}a_1 + \beta_{12}a_2 + 3\beta_{111}a_1^2 + \beta_{122}a_2^2 + \beta_{133}a_3^2 + \\ & \quad + 2\beta_{112}a_1a_2 + 2\beta_{113}a_1a_3 + \beta_{123}a_2a_3, \\ \alpha_2 &= \beta_2 + \beta_{12}a_1 + 2\beta_{22}a_2 + \beta_{23}a_3 + 3\beta_{222}a_2^2 + \beta_{233}a_3^2 + \\ & \quad + 2\beta_{122}a_1a_2 + \beta_{223}a_2a_3 + \beta_{123}a_1a_3, \\ \alpha_3 &= \beta_3 + \beta_{13}a_2 + 2\beta_{33}a_3 + \beta_{113}a_1^2 + \beta_{223}a_2^2 + 3\beta_{333}a_3^2 + \\ & \quad + 2\beta_{133}a_1a_3 + 2\beta_{233}a_2a_3 + \beta_{123}a_1a_2, \\ \alpha_{11} &= \beta_{11} + 2\beta_{111}a_1 + \beta_{112}a_2 + \beta_{113}a_3, \\ \alpha_{22} &= \beta_{22} + \beta_{122}a_1 + 3\beta_{222}a_2 + \beta_{223}a_3, \\ \alpha_{33} &= \beta_{33} + \beta_{133}a_1 + \beta_{233}a_2 + 3\beta_{333}a_3, \\ \alpha_{12} &= \beta_{12} + 2\beta_{112}a_1 + 2\beta_{122}a_2 + \beta_{123}a_3, \\ \alpha_{13} &= \beta_{13} + 2\beta_{113}a_1 + \beta_{123}a_2 + 2\beta_{133}a_3, \\ \alpha_{23} &= \beta_{23} + \beta_{123}a_1 + 2\beta_{223}a_2 + 2\beta_{233}a_3, \\ \alpha_{111} &= \beta_{111}, & \alpha_{222} &= \beta_{222}, & \alpha_{333} &= \beta_{333}, \\ \alpha_{122} &= \beta_{122}, & \alpha_{133} &= \beta_{133}, & \alpha_{112} &= \beta_{112}, \\ \alpha_{233} &= \beta_{233}, & \alpha_{113} &= \beta_{113}, & \alpha_{223} &= \beta_{223}, \\ \alpha_{123} &= \beta_{123}. \end{aligned} \quad (13.3.7)$$

Выражения для параметров квадратичной модели с тремя переменными после переноса начала координат можно найти, приравняв к нулю все параметры  $\beta$  с тремя индексами в (13.3.6) и (13.3.7). Более полные выводы на основании предложенных в примере 2 идей читатель может сделать самостоятельно.

#### Пример 4

В работе Дж. Л. Пеиксото [262] приводится более общий пример хорошо сформулированной полиномиальной модели, удовлетворяющей критерию 1:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \beta_{22} X_2^2 + \beta_{122} X_1 X_2^2 + \beta_{222} X_2^3 + \beta_{1222} X_1 X_2^3 + \epsilon. \quad (13.3.8)$$

*Задание:* подставьте  $X_1 = Z_1 + a_1$ ,  $X_2 = Z_2 + a_2$  в (13.3.8) и покажите, что в такой расширенной модели не появятся полиномиальные члены нового типа.

#### Критерий 2. Критерий поворота осей

Мы будем считать сокращенную модель осмысленной, если поворот осей  $X$ -пространства не меняет *вида* модели относительно переменных  $X$ .

#### Пример 5. Полином второго порядка относительно двух предикторов

Вернемся к модели второго порядка (13.3.2). Рассмотрим поворот, при котором осуществляется переход от переменных  $X$ , скажем, к переменным  $W$ , таким, что

$$\begin{aligned} W_1 &= c_1 X_1 + c_2 X_2, \\ W_2 &= d_1 X_1 + d_2 X_2, \end{aligned} \quad (13.3.9)$$

или в матричном виде —  $\mathbf{W} = \mathbf{M}\mathbf{X}$ , где  $\mathbf{M}$  — ортогональная матрица, такая, что  $\mathbf{M}'\mathbf{M} = \mathbf{M}\mathbf{M}' = \mathbf{M}\mathbf{M}^{-1} = \mathbf{I}$ . Эти условия, которые сохраняют как длину, так и взаимную ортогональность осей при повороте, означают, что матрица  $\mathbf{M}$  должна иметь вид

$$\mathbf{M} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \quad (13.3.10)$$

или подобный, с возможной заменой знаков в строках и/или столбцах.

Чтобы подставить  $X$  в (13.3.2), выполним «обратное» преобразование  $\mathbf{X} = \mathbf{M}^{-1}\mathbf{W} = \mathbf{M}'\mathbf{W}$ . Получим:

$$\begin{aligned} X_1 &= cW_1 - sW_2, \\ X_2 &= sW_1 + cW_2. \end{aligned} \quad (13.3.11)$$

Уравнение поверхности после выполнения поворота осей примет вид

$$Y = \gamma_0 + \gamma_1 W_1 + \gamma_2 W_2 + \gamma_{11} W_1^2 + \gamma_{22} W_2^2 + \gamma_{12} W_1 W_2 + \epsilon,$$

где

$$\begin{aligned} \gamma_0 &= \beta_0, \\ \gamma_1 &= c\beta_1 + s\beta_2, \\ \gamma_2 &= -s\beta_1 + c\beta_2, \\ \gamma_{11} &= c^2\beta_{11} + s^2\beta_{22} + cs\beta_{12}, \\ \gamma_{22} &= s^2\beta_{11} + c^2\beta_{22} - cs\beta_{12}, \\ \gamma_{12} &= -2cs\beta_{11} + 2cs\beta_{22} + (c^2 - s^2)\beta_{12}. \end{aligned} \quad (13.3.12)$$

(Случай  $\theta = 45^\circ$ ,  $c = s = 2^{-1/2}$  использовался в работе Бокса и Дрейпера [41], сс. 447–448.) Мы видим, что порядок полностью сохраняется, а именно: коэффициенты  $\gamma$  нулевого, первого и второго порядков являются линейными комбинациями коэффициентов  $\beta$  исключительно нулевого, первого и второго порядков соответственно. Данный пример, распространенный на большее количество переменных  $X$ , иллюстрирует более общее правило 2.

## Правило 2

Если модель должна оставаться согласованной при повороте осей, все члены одного порядка должны рассматриваться как единое целое. Нельзя исключать подмножества членов определенного порядка. (Например, бессмысленно рассматривать исключение только коэффициента  $\beta_{12}$ , поскольку это не приведет к исключению коэффициента  $\gamma_{12}$ , который останется в виде комбинации  $\beta_{11}$  и  $\beta_{22}$ .)

## Совместное применение правил 1 и 2

Когда модель должна оставаться согласованной как при сдвиге начала координат, так и при повороте осей, то члены могут исключаться из нее только группами членов одного порядка (например, всех членов второго порядка или всех членов первого). Более того, нельзя исключать члены определенного порядка, если члены более высокого порядка остаются в модели.

Когда направления осей можно считать фиксированными, так что возможность поворота осей можно не учитывать (скажем, из практических соображений), применяется только правило 1.

### Потребность в той или иной предикторной переменной

Допустим, мы не хотим рассматривать повороты поверхности, т.е. правило 2 можно не учитывать. Тогда можно ставить вопрос о целесообразности присутствия в модели тех или иных предикторных переменных. Обратимся вновь к формулам (13.3.5) и зададим вопрос «Вносит ли переменная  $X_2$  вклад в кривизну второго порядка?». Это эквивалентно проверке гипотезы  $H_0: \beta_{22} = \beta_{12} = 0$  против альтернативы  $H_1$ , что это не так. Мы видим, что в этом случае  $\alpha_{22} = \alpha_{12} = 0$ , так что правило 1 выполняется. Затем можно проверить нуль-гипотезу обычным способом с помощью  $F$ -критерия для дополнительной суммы квадратов. Можно также проверить, *нужна ли вообще* переменная  $X_2$ , сформулировав гипотезу  $H_0: \beta_2 = \beta_{22} = \beta_{12} = 0$ , поскольку тогда  $\alpha_2 = \alpha_{22} = \alpha_{12} = 0$ . Из этого следует, что если повороты осей невозможны в силу природы предикторов, то подмножества членов более высокого порядка, зависящих от одной или нескольких указанных предикторных переменных, или все такие члены могут быть исключены, но только в таких комбинациях, чтобы модель осталась хорошо сформулированной, т.е. чтобы выполнялось правило 1. К примеру, в (13.3.7) возможны представленные ниже варианты исключений.

- а)  $\beta_{111} = \beta_{122} = \beta_{133} = \beta_{113} = \beta_{123} = 0$ . (Вносит ли переменная  $X_1$  вклад в кривизну третьего порядка?)
- б)  $\beta_{11} = \beta_{12} = \beta_{13} = 0$  в дополнение к (а). (Вносит ли переменная  $X_1$  вклад в кривизну второго и третьего порядков?)
- в)  $\beta_1 = 0$  в дополнение к (б). (Нужна ли переменная  $X_1$  вообще?)

### Заключительный совет

При принятии решения относительно того, какие члены можно исключить из полиномиальной модели, мы советуем:

1. всегда применять правило 1;
2. в зависимости от природы исходных предикторов поступать так:
  - а) если исходные переменные не являются фиксированными, так что возможен поворот осей с целью получения поверхности в новых переменных, являющихся линейными комбинациями исходных, следует в дополнение к правилу 1 применять правило 2;
  - б) если исходные переменные являются фиксированными (т.е. поверхность отклика всегда описывается в данных переменных), применяется только правило 1, при этом можно рассматривать гипотезы относительно отдельных предикторных переменных, как в приведенном выше примере.

### Процедура отбора для полиномиальных уравнений

Процедуры отбора (см. главу 16) не содержат описанных здесь правил, поэтому к их использованию для получения полиномиальных моделей нужно подходить с осторожностью. На основании процедуры можно предложить вариант, но полученное уравнение затем следует уточнить с учетом приведенных выше критериев.

### Литература

Бокс и Дрейпер [41]; Дрисколл и Андерсон [103]; Пеиксото [261, 262].

### УПРАЖНЕНИЯ К ГЛАВЕ 13

1. Получены восемнадцать наблюдений четырех предикторов и одной переменной отклика. Данные представлены в следующей таблице.

$X_1$	$X_2$	$X_3$	$X_4$	$Y$
20	50	75	15	27
27	55	60	20	23
22	62	68	16	18
27	55	60	20	26
24	75	72	8	23
30	62	73	18	27
32	79	71	11	30
24	75	72	8	23
22	62	68	16	22
27	55	60	20	24
40	90	78	32	16
32	79	71	11	28
50	84	72	12	31
40	90	78	32	22
20	50	75	15	24
50	84	72	12	31
30	62	73	18	29
27	55	60	20	22

Предполагается, что корректна модель

$$Y = \beta_0 X_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{11} X_1^2 + \beta_{12} X_1 X_2 + \beta_{22} X_2^2 + \beta_{13} X_1 X_3 + \beta_{14} X_1 X_4 + \epsilon.$$

- а) Исследуйте данные и модель. Можно ли подобрать предложенную модель по имеющимся данным? Почему?
- б) Оцените  $V(Y_i) = \sigma^2$ .

2. По представленным ниже данным подберите модель

$$Y = (\alpha_0 + \alpha_1 Z + \alpha_2 Z^2) + (\beta_0 + \beta_1 Z + \beta_{11} Z^2) X_1 + (\gamma_0 + \gamma_1 Z + \gamma_{11} Z^2) X_2 + \epsilon,$$

где  $Z = \ln(X_3 + 1)$ , и проведите полный анализ. Обратите внимание, что в плане эксперимента имеются шесть повторов.

$X_1$	$X_2$	$X_3$	$Y$	$X_1$	$X_2$	$X_3$	$Y$
47,1	33,9	7,5	11,97	47,1	33,9	750,0	8,46
72,9	33,9	750,0	8,63	60,0	21,0	75,0	10,65
47,1	8,1	750,0	8,80	60,0	21,0	3000,0	7,60
60,0	21,0	75,0	10,73	60,0	21,0	3,0	13,06
60,0	21,0	75,0	10,69	39,0	21,0	75,0	10,51
72,9	8,1	7,5	13,12	60,0	0,0	75,0	11,22
47,1	8,1	7,5	12,58	60,0	21,0	75,0	10,67
72,9	33,9	7,5	12,24	60,0	42,0	75,0	10,24
60,0	21,0	75,0	10,64	81,0	21,0	75,0	10,74
72,9	8,1	750,0	0,09	60,0	21,0	75,0	10,69

Источник: G. C. Derringer, «Variable shear rate viscosity of SBR-filter-plasticizer systems», Rubber Chemistry and Technology, 47, September 1974, pp. 825–836.

3. В приведенной ниже таблице представлены результаты испытаний 24 одинаковых плоских обогревателей.  $T$  — температура в градусах по Фаренгейту, а  $Y$  — срок службы (в часах) отдельного обогревателя при данной температуре. При каждом значении температуры проверялись шесть обогревателей. Постройте точечную диаграмму данных. Подберите модель вида  $\log_{10} Y = \beta_0 + \beta_1 \{1000/(T + 460)\} + \epsilon$  и проведите необходимый анализ.

$T$	$Y$
1520	1953, 2135, 2471, 4727, 6143, 6314
1620	1190, 1286, 1550, 2125, 2557, 2845
1660	651, 837, 848, 1038, 1361, 1543
1708	511, 651, 651, 652, 688, 729

Примечание.  $T + 460$  — абсолютная температура в °F.

Источник: Wayne Nelson, «A short life test for comparing a sample with previous accelerated test results», Technometrics, 14, 1972, pp. 175–185.

4. В приведенной ниже таблице собраны данные производственного эксперимента, проводившегося для исследования влияния изменений процентного содержания определенной составляющей смеси ( $X_1$ ), температуры смеси ( $X_2$ ) и скорости протекания ( $X_3$ ) на величины  $Y_1$ ,  $Y_2$ , и  $Y_3$

(отклики). Предикторные переменные закодированы, а переменные отклика приводятся в исходных единицах. План эксперимента представляет собой центральную составную схему, состоящую из восьми вершин куба  $(X_1; X_2; X_3) = (\pm 1; \pm 1; \pm 1)$ , шести осевых точек  $(\pm\alpha; 0; 0)$ ,  $(0; \pm\alpha; 0)$ ,  $(0; 0; \pm\alpha)$ , где  $\alpha = 1,2154$ , и одной центральной  $(0; 0; 0)$ . Порядок следования результатов в таблице отражает случайный порядок, в котором проводились опыты.

$X_1$	$X_2$	$X_3$	$Y_1$	$Y_2$	$Y_3$
-1	-1	1	85,3	72,7	97,1
1	1	-1	72,3	57,6	96,9
0	1,2154	0	71,4	56,5	96,4
0	-1,2154	0	72,0	64,6	96,8
-1	-1	-1	87,0	79,2	97,0
1	1	1	55,6	32,6	96,2
0	0	-1,2154	85,0	75,9	97,2
1,2154	0	0	70,9	53,4	97,9
0	0	0	75,9	59,3	97,4
1	-1	1	76,1	63,2	97,4
-1	1	-1	85,0	75,3	97,2
0	0	1,2154	68,0	57,2	95,5
-1,2154	0	0	89,6	83,6	97,2
-1	1	1	75,0	61,5	96,5
1	-1	-1	74,2	61,0	98,2

Является ли данный план ротабельным? (См. раздел 13.2.) Используя методы множественной регрессии, сформулируйте и подберите подходящие модели первого или второго порядка для  $Y_1$ ,  $Y_2$ ,  $Y_3$  по отдельности. Выполните полный анализ и дайте практические рекомендации. Если предпочтительны более высокие значения откликов, в какой части  $X$ -пространства лучше действовать?

5. В лаборатории крупной корпорации, производящей хлебобулочные изделия, исследуется новый продукт. Наибольший интерес представляет высота теста в форме непосредственно перед помещением в печь. Предполагается, что на высоту подъема теста могут влиять четыре основных ингредиента: процентное содержание жира, процентное содержание воды, количество муки в опаре и скорость вращения миксера (оборотов в минуту). Данные о проведенных экспериментах приведены в табл. 13.7. Для всех опытов фиксировались значения максимальной высоты подъема теста; они также приведены в таблице. Обратите внимание на четыре повтора при условиях (12; 50; 20; 130), значения отклика для которых составили 492, 523, 530 и 590.

- а) Выбрав подходящие центральные уровни и масштабные множители, закодируйте все четыре предикторные переменные так, чтобы их уровнями были значения  $-1$ ,  $0$  и  $1$ . Запишите план эксперимента в кодированных переменных и убедитесь, что это схема «куб плюс звезда плюс четыре центральные точки». Является ли данный план ротатбельным?
- б) Используя методы множественной регрессии, постройте подходящую модель первого или второго порядка для предсказания максимальной высоты подъема теста. В заключение охарактеризуйте относительную важность предикторных переменных и сделайте другие необходимые замечания.
6. Подберите полную модель второго порядка и выполните полный анализ, используя данные про  $Y_4$  из табл. 13.1.
7. Является ли модель

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \beta_{22} X_2^2 + \beta_{122} X_1 X_2^2 + \epsilon$$

хорошо сформулированной в смысле критерия сдвига начала координат?

**Таблица 13.7.** Экспериментальные данные для исследования максимальной высоты подъема теста

Содержание жира, %	Муки в опаре	Содержание воды, %								
		46			50			54		
		об/мин			об/мин			об/мин		
		90	130	170	90	130	170	90	130	170
8	10	833	540		537			673	493	
	20	577		547		660			512	
	30	547		547			547		547	
12	10	653		650			553		487	
	20	653		650			553		487	
	30	653		650			553		487	
16	10	802	477		575			710	520	
	20	568		401		575			483	
	30	568		401		575			483	

8. Основанная на теоретических соображениях модель имеет вид

$$Y = \alpha X_1^\beta X_2^\gamma X_3^\delta \epsilon.$$

Преобразовав модель выполните подбор методом наименьших квадратов. Определите, какая предикторная переменная представляется наиболее важной, и проверьте статистическую значимость всех коэффициентов (примите  $\alpha = 0,05$ ). Является ли модель удовлетворительной?

Ниже представлены данные, относящиеся к изучению количественного содержания витамина В<sub>2</sub> в свежей репе, где переменные обозначают:  $X_1$  — солнечная радиация, выраженная в грамм-калориях в минуту, на протяжении предшествующей половины дня (кодированная путем деления на 100),  $X_2$  — среднее давление почвенной влаги (кодированное путем деления на 100),  $X_3$  — температура воздуха в градусах по Фаренгейту (кодированная путем деления на 10),  $Y$  — содержание витамина В<sub>2</sub> в зеленой репе (миллиграммов в грамме). Данные взяты из статьи Дж. Т. Уэкли *Annual progress report on soils-weather project, 1948*, University of North Carolina (Raleigh) Institute of Statistics Mimeo Series 19 (1949).

Эти данные Р. Л. Андерсон и Т. А. Бенкрофт использовали в своей книге *Statistical Theory in Research*, McGraw-Hill, New York, 1959 (с. 192) для подбора модели

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{12} X_1 X_2 + \epsilon.$$

Выберите для этих данных подходящую модель, подберите регрессионное уравнение и сравните его с уравнением, подобранным Андерсоном и Бенкрофтом.

$X_1$	$X_2$	$X_3$	$Y$	$X_1$	$X_2$	$X_3$	$Y$
1,76	0,070	7,8	110,4	1,80	0,020	7,3	75,3
1,55	0,070	8,9	102,8	1,80	0,020	6,5	91,0
2,73	0,070	8,9	101,0	1,77	0,020	7,6	82,4
2,73	0,070	7,2	108,4	2,30	0,020	8,2	77,1
2,56	0,070	8,4	100,7	2,03	0,474	7,6	74,0
2,80	0,070	8,7	100,3	1,91	0,474	8,3	65,7
2,80	0,070	7,4	102,0	1,91	0,474	8,2	56,8
1,84	0,070	8,7	93,7	1,91	0,474	6,9	62,1
2,16	0,070	8,8	98,9	0,76	0,474	7,4	61,0
1,98	0,020	7,6	96,6	2,13	0,474	7,6	53,2
0,59	0,020	6,5	99,4	2,13	0,474	6,9	59,4
0,80	0,020	6,7	96,2	1,51	0,474	7,5	58,7
0,80	0,020	6,2	99,0	2,05	0,474	7,6	58,0
1,05	0,020	7,0	88,4				