Глава VIII

МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

1. Введение. В любой теории движения тела, будь то относительно другого тела или вокруг своей оси вращения, встречаются определенные постоянные, которые должны быть найдены при помощи наблюдений. Если об этих постоянных ничего не известно заранее, как, например, в случае элементов орбиты вновь открытого объекта, то их определение может оказаться затруднительным. Существует обширная литература для достижения этой конкретной цели в указанном случае. Однако если известны приближенные значения постоянных, то их можно ввести в теорию движения, которая затем может быть использована для вычисления теоретического положения объекта. Сравнение этой теории с наблюдениями покажет, что теория не представляет наблюдений точно. Каждое наблюдение дает остаточную разность в смысле «наблюденное место минус вычисленное место» $(O-C)$, которая обусловливается следующими тремя причинами и никакими другими. Во-первых, сама теория может быть неточной; во-вторых, наблюдения отягощены ошибками; в-третьих, в остаточные разности включено влияние ошибок приближенных значений постоянных, использованных при определении вычисленных положений. В этой главе рассматриваются два последних класса ошибок. Мы покажем, каким образом можно улучшить приближенные значения постоянных путем анализа расхождений между наблюдениями и теорией.

Этот метод анализа может быть использован для получения не только более точных значений элементов орбиты, но также улучшенных значений любого другого параметра, от которого зависят наблюдения. В качестве примера можно упомянуть элементы орбиты Земли, массы возмущающих планет, солнечный параллакс, постоянную нутации и другие астрономические постоянные. В каждом отдельном случае необходимы такие наблюдения, чтобы ошибка в принятом значении постоянной оказывала ощутимое влияние.

2. Частотное распределение ошибок наблюдений. Ошибки наблюдений вызываются различными причинами, среди которых следует упомянуть оптические недостатки телескопов, флуктуации плотности атмосферы и дефекты микрометрических винтов, применяемых при измерении фотографических пластинок. Ошибки некоторого ряда наблюдений имеют различную величину. Если мы представим себе величину ошибки, нанесенной на график как абсциссу, а число наблюдений, имеющих эту величину ошибки, как ординату, то построенная кривая называется кривой частотного распределения, или кривой вероятностного распределения ошибок. Такая кривая вообще имеет форму колокола, который более или менее симметричен относительно точки нулевой ошибки, показывая этим, что положительные и отрицательные ошибки встречаются с одинаковой частостью и что малые ошибки более многочисленны, чем большие.

В большинстве случаев, с которыми приходится иметь дело в астрономии, частотное распределение ошибок заранее не известно, а может быть определено лишь после весьма тщательного исследования наблюдательного материала, если это представляется вообще возможным. Однако можно сделать предположение относительно распределения ошибок, которое будет правильным, *если каждая из ошибок является суммой некоторого числа* одновременно возникающих ошибок, имеющих различное происхождение. Такого рода ошибки часто называются случайными ошибками в отличие от систематических ошибок, которые могут иметь общее происхождение. На практике астрономические наблюдения обычно подвержены влиянию ошибок обоих родов, и присутствие систематических ошибок может свести на нет все сделанные выводы. Тем не менее, по крайней мере сначала, мы вынуждены поступить так, как если бы эти ошибки были случайными, так как математической теории систематических ошибок не существует.

Что же касается случайных ошибок, то имеется теорема в теории вероятностей, утверждающая, что частотное распределение суммы случайных переменных величин, каждая из которых имеет собственное произвольное частотное распределение, асимптотически приближается к функции

$$f(x)\,dx = h\pi^{-1/2} \exp\{-h^2x^2\}\,dx, \tag{1}$$

если индивидуальные случайные переменные не зависят друг от друга. В этом выражении $f(x)\,dx$ является относительным числом случаев, в которых случайная переменная величина имеет значение между x и $x+dx$, где π — известная постоянная, выражение $\exp\{-h^2x^2\}$ означает основание натуральных логарифмов, возведенное в степень $-h^2x^2$, а h называется мерой точности. Множитель $h\pi^{-1/2}$ введен для того, чтобы выполнялось условие

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1.$$

Вместо h мы можем ввести среднее из квадратов случайной переменной x. Квадратный корень из этого среднего называется стандартным отклонением σ или средней квадратичной ошибкой (с. к. о.). Соотношение между σ и h имеет следующий вид:

$$\sigma^2 = h\pi^{-1/2} \int_{-\infty}^{+\infty} x^2 \exp\{-h^2x^2\}\,dx = \frac{1}{2h^2},$$

или

$$h^2 = \frac{1}{2\sigma^2}$$

и

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\}.$$

Частотное распределение (1) называется нормальным, или гауссовым, распределением. Если ошибки обладают такого рода распределением, то стандартное отклонение σ называется средней ошибкой, или стандартной ошибкой. Другим параметром, часто используемым в астрономии, является вероятная ошибка, которая определяется как величина, которая превосходит по абсолютной величине одну половину ошибок и меньше второй их половины. Вероятная ошибка (в. о.), определенная таким образом, имеет определенный смысл, каким бы ни было распределение ошибок. Если распределение является гауссовым, то тогда и только тогда можно показать, что вероятная ошибка связана со стандартным отклонением следующим соотношением:

$$\text{в. о.} = 0{,}6745\sigma.$$

Таблица 1 показывает для гауссова распределения 10 000 ошибок число тех ошибок, которые превосходят различные кратности вероятной ошибки.

Таблица 1

Распределение ошибок

Кратное в. о.	Число больших ошибок
1	5000 или 1 в 2 случаях
2	1774 или 1 в 6 »
3	430 или 1 в 23 »
4	70 или 1 в 140 »
5	8 или 1 в 1 200 »
6	1 или 1 в 10 000 »

Любое заметное отклонение от чисел в этой таблице указывает на то, что распределение не является гауссовым.

Если мы имеем линейную комбинацию двух случайных переменных вида

$$x = a_1x_1 + a_2x_2, \tag{2}$$

в которой a_1 и a_2 — постоянные, а переменные x_1 и x_2 имеют нормальные распределения со стандартными отклонениями σ_1 и σ_2, то распределение x можно найти следующим образом.

Одновременная, или совместная, вероятность того, что x_1 принимает значение между x_1 и $x_1 + dx_1$, а x_2 — между x_2 и $x_2 + dx_2$, равна

$$\frac{1}{\sigma_1\sqrt{2\pi}} \cdot \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left\{-\frac{x_1^2}{2\sigma_1^2} - \frac{x_2^2}{2\sigma_2^2}\right\} dx_1\,dx_2. \tag{3}$$

Можно показать, что для

$$x < a_1x_1 + a_2x_2 < x + dx$$

будет иметь место равенство

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\},$$

где

$$\sigma^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2. \tag{4}$$

Следовательно, распределение линейной комбинации двух случайных переменных, каждая из которых имеет нормальное распределение, является нормальным распределением со стандартным отклонением, определяемым формулой (4). Этот закон можно распространить на линейную комбинацию любого числа случайных переменных.

3. Наиболее вероятное значение измеренной величины. Рассмотрим n измерений $x_1, x_2, \ldots, x_n$ некоторой величины и обозначим истинное значение этой величины через a. Тогда ошибка ε_j любого измерения x_j равна $\varepsilon_j = x_j - a$. Если эти ошибки обладают нормальным распределением, то вероятность любого измерения x_j равна

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x_j-a)^2}{2\sigma^2}\right\},$$

а совместная вероятность для этих n измерений равна

$$\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\left[(x_1-a)^2+(x_2-a)^2+\ldots+(x_n-a)^2\right]\right\}. \tag{5}$$

Наиболее вероятное значение величины a определяется как значение, при котором выражение (5) имеет максимум или величина имеет

$$\frac{1}{2\sigma^2}\sum (x_j-a)^2 \tag{6}$$

минимум. Чтобы привести (6) к минимуму, достаточно иметь

$$\frac{\partial\Sigma}{\partial a} = 0 \quad \text{или} \quad 2\Sigma\,(x_j-a) = 0,$$

откуда

$$a = \frac{1}{n}\sum x_j. \tag{7}$$

Таким образом, наивероятнейшим значением величины a является среднее арифметическое из этих наблюдений. Наиболее вероятное значение величины σ можно найти из уравнений

$$\frac{\partial}{\partial\sigma}\left[n\ln\sigma + \frac{\Sigma\,(x_j-a)^2}{2\sigma^2}\right] = 0 \quad \text{или} \quad \frac{n}{\sigma} - \frac{1}{\sigma^3}\sum (x_j-a)^2 = 0,$$

откуда

$$\sigma^2 = \frac{1}{n}\sum (x_j-a)^2. \tag{8}$$

Заметим, что в уравнении (8) a означает истинное (неизвестное) значение измеряемой величины, а не наивероятнейшее значение ее, определяемое формулой (7).

Если частотное распределение ошибок не является нормальным распределением, то все же мы можем найти наиболее вероятные значения для a и σ, если известен вид этого распределения. Пусть вероятность любого единственного измерения есть $f(a, x_1, x_2, \ldots, x_n)$, где f есть некоторая известная функция. Тогда совместная вероятность равна

$$\prod_{j=1}^{n} f(x_j-a),$$

и наиболее вероятные значения можно найти из уравнений

$$\frac{\partial\Pi}{\partial a} = 0 \quad \text{и} \quad \frac{\partial\Pi}{\partial\sigma} = 0.$$

Как видно из выражения (8), средняя ошибка среднего арифметического (7) может быть определена следующим образом. Пусть имеется n наблюдений x_j ($j=1, 2, \ldots, n$) с нормально распределенными истинными ошибками ε_j. Ошибка среднего арифметического равна

$$\varepsilon = \frac{1}{n} \sum \varepsilon_j,$$

что представляет собой линейную комбинацию n случайных переменных. Если σ — средняя ошибка одного наблюдения, то в силу формулы (4) имеем

$$\sigma_\varepsilon^2 = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n},$$

откуда $\sigma_\varepsilon = \sigma n^{-1/2}$, или средняя ошибка среднего арифметического равна средней ошибке одного наблюдения, деленной на квадратный корень из числа наблюдений.

Мы видели, что если бы было известно истинное значение измеряемой величины, то средняя ошибка одного наблюдения давалась бы формулой (8). Однако это истинное значение не известно; мы располагаем лишь оценкой, которая дается формулой (7). После того как эта оценка получена, мы можем вычесть ее из каждого из измеренных значений; разности называются остаточными разностями и обозначаются через v. Обозначим среднее арифметическое из x_j через $\bar{x}$, истинную ошибку наблюдения x_j, как и ранее, через ε_j, а истинное значение измеряемой величины — через a. Тогда

$$v_j = x_j - \bar{x},$$

$$\varepsilon_j = x_j - a.$$

Обозначив через $[x]$ сумму величин x_j, а через $[\varepsilon]$ — сумму истинных ошибок ε_j, мы имеем

$$[\varepsilon] = [x] - na = n\bar{x} - na,$$

откуда

$$\bar{x} = a + \frac{[\varepsilon]}{n}$$

и

$$v_j = \varepsilon_j - \frac{[\varepsilon]}{n}.$$

Теперь, если σ_v является стандартным отклонением остаточных разностей v_j, а σ_ε — стандартным отклонением ошибок ε_j, то мы имеем

$$\sigma_v^2 = \left(\frac{n-1}{n}\right)^2 \sigma_\varepsilon^2 + \frac{n-1}{n^2} \sigma_\varepsilon^2 = \frac{n-1}{n} \sigma_\varepsilon^2.$$

Поскольку $\sigma_v^2 = [v^2]/n$, то отсюда следует, что

$$\sigma_\varepsilon^2 = \frac{[v^2]}{n-1},$$

или

$$\varepsilon_\varepsilon = \sqrt{\frac{[v^2]}{n-1}}.$$

В таком случае, обозначая среднюю ошибку среднего арифметического через μ, имеем

$$\mu = \sqrt{\frac{[v^2]}{n(n-1)}}. \tag{9}$$

4. Веса наблюдений. Обычно мы вынуждены поступать так, как если бы качество различных наблюдений было одинаковым, но вообще это не так. Предположим, что средняя ошибка наблюдения x_1 равна σ_1, а средняя ошибка наблюдения x_2 — σ_2. Тогда совместная вероятность дается выражением

$$\frac{1}{\sigma_1\sqrt{2\pi}} \cdot \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left\{-\frac{(x_1-a)^2}{2\sigma_1^2} - \frac{(x_2-a)^2}{2\sigma_2^2}\right\}.$$

Функция, которую необходимо свести к минимуму, имеет вид

$$\frac{(x_1-a)^2}{\sigma_1^2} + \frac{(x_2-a)^2}{\sigma_2^2},$$

откуда

$$a = \frac{(x_1/\sigma_1^2)+(x_2/\sigma_2^2)}{(1/\sigma_1^2)+(1/\sigma_2^2)}. \tag{10}$$

Если мы положим $p_1 = 1/\sigma_1^2$ и $p_2 = 1/\sigma_2^2$, то

$$a = \frac{p_1x_1+p_2x_2}{p_1+p_2},$$

или для n наблюдений

$$a = \frac{[px]}{[p]}.$$

Величина p_j называется весом наблюдения x_j.

5. Непрямые измерения. Как упоминалось в разд. 1, обычными случаями в астрономии являются такие, когда искомая величина не измеряется непосредственно; чаще всего измеряемая величина является некоторой известной функцией от искомой величины. Эта искомая величина обычно является постоянной, значение которой уже известно приближенно. В самом простом случае ищут лишь одну величину; обозначим эту величину через x, а измеряемую величину — через $y = f(x)$. Допустим, что мы имеем n наблюдений y_j, $j = 1, 2, \ldots, n$, по которым требуется вывести наиболее вероятное значение x. Обозначим приближенно известное значение x через x_c, значение, которое необходимо вывести по наблюдениям, — через x_0, а разность $x_0 - x_c$ — через Δx. Процедура состоит в вычислении y по x_c для каждого наблюдения, что дает y_c, и затем в образовании разностей $\Delta_j y = y_j - y_c$. Поскольку вообще y является функцией как от x, так и от других параметров, то величина y_c меняется при переходе от наблюдения к наблюдению; чтобы показать это изменение, мы можем записать $\Delta_j y = y_j - y_{jc}$.

Теперь Δy можно выразить как функцию от Δx при помощи разложения Тэйлора в виде

$$\Delta y = \frac{\partial f}{\partial x}\Delta x + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\Delta x)^2 + \frac{1}{6}\frac{\partial^3 f}{\partial x^3}(\Delta x)^3 + \ldots.$$

Предполагается, что Δx настолько мало, что можно пренебречь его квадратом и более высокими степенями, откуда с достаточной степенью точности следует

$$\Delta y = \frac{\partial f}{\partial x} \Delta x,$$

так что для каждого наблюдения мы можем найти уравнение следующего вида:

$$\left(\frac{\partial f}{\partial x}\right)_j \Delta x = \Delta_j y, \tag{11}$$

где величина в скобках является известным числом. Уравнение (11) называется условным уравнением, а Δx называется неизвестным.

Запишем для краткости уравнение (11) в виде

$$a\Delta x = \Delta y. \tag{12}$$

Как и в предыдущих разделах, условие для наиболее вероятного значения неизвестного Δx имеет вид

$$\frac{\partial}{\partial \Delta x} \sum (a_j \Delta x - \Delta_j y)^2 = 0,$$

или

$$[a^2 \Delta x] = [a\Delta y], \tag{13}$$

откуда

$$\Delta x = \frac{[a\Delta y]}{[a^2]} \tag{14}$$

и, наконец, x_0 определяется формулой

$$x_0 = x_c + \Delta x.$$

Допустим, что дано n уравнений вида (12), веса которых предполагаются одинаковыми, и что для q из них величина a является постоянной. Тогда эти q уравнений вносят в (13) член вида

$$qa^2 \Delta x = a [\Delta y] = \frac{\sqrt{q} a [\Delta y]}{\sqrt{q}}.$$

Сравнение с уравнением (12) показывает, что мы можем написать условное уравнение вида

$$\sqrt{q} a \Delta x = \frac{[\Delta y]}{\sqrt{q}} = \frac{\sqrt{q} [\Delta y]}{q},$$

или

$$\sqrt{q} a \Delta x = \sqrt{q}\, \overline{\Delta y}, \tag{15}$$

которое показывает, что эти q уравнений можно заменить одним уравнением (15), где $\overline{\Delta y}$ означает среднее из q значений Δy. В этом случае говорят, что уравнение

$$a\Delta x = \overline{\Delta y}$$

имеет вес q, и его необходимо умножить на $\sqrt{q}$, прежде чем использовать с остальными $n - q$ уравнениями для получения уравнений (13) и (14).

6. Условные уравнения. В общем случае требуется определить не одно, а n неизвестных по ряду наблюдений. Обозначим наблюдаемую величину, например сферическую координату тела, через x, и пусть Δx означает расхождение в смысле наблюденное x *минус* вычисленное x $(O-C)$ между каким-нибудь наблюдением и соответствующим вычисленным значением. Пусть c_j, $j=1, 2, \ldots, n$, означает принятые значения постоянных, использованных при вычислении этих расчетных значений, и пусть поправки к принятым значениям постоянных обозначены через ξ_j, так что исправленное значение любой постоянной равно $c_j+\xi_j$. Тогда условные уравнения, выведенные тем же способом, что и уравнения (11), могут быть написаны в следующем виде:

$$\frac{\partial x}{\partial c_1}\xi_1+\frac{\partial x}{\partial c_2}\xi_2+\ldots+\frac{\partial x}{\partial c_n}\xi_n=\Delta x. \tag{16}$$

Допустим пока, что Δx не подвержено влиянию первых двух классов ошибок, упомянутых в разд. 1; тогда оно целиком определяется уравнением (16). Частные производные, входящие в (16) в качестве коэффициентов, могут быть вычислены и выражены числами; при этом почти всегда достаточно трех значащих цифр. Неизвестными в этих уравнениях являются ξ_j, которые необходимо определить. Из теории линейных алгебраических уравнений известно, что n уравнений типа (16) составляют необходимое и достаточное количество уравнений для определения неизвестных ξ_j при условии, что все эти уравнения независимы, т. е. если определитель из коэффициентов левой части не обращается в нуль. Если эти уравнения независимы, то можно найти значения величин ξ_j, которые точно удовлетворят n уравнениям, и если бы наблюдения не были отягощены ошибками, то ничего больше не требовалось бы. Однако присутствие случайных ошибок в правых частях этих уравнений мешает определить истинные значения ξ_j; можно найти лишь приближенные значения, которые будут ближе к истине или дальше от нее, смотря по тому, будут ли ошибки наблюдений меньше или больше. Влияние случайных ошибок можно ослабить при помощи увеличения числа наблюдений и числа уравнений. В астрономических задачах число используемых уравнений редко бывает меньше $2n$ и часто бывает еще больше; когда необходима самая высокая степень точности, для определения нескольких неизвестных могут быть использованы несколько сотен и даже тысяч наблюдений.

Когда число уравнений больше числа неизвестных, то в этом случае из-за ошибок наблюдений точное решение не существует; невозможно найти значения величин ξ_j, которые точно удовлетворяют всем уравнениям. Если любые частные значения величин ξ_j подставлены в левые части этих уравнений, и результаты в отдельности вычтены из Δx, то остающиеся числа называются остаточными разностями «после решения». Как мы видели ранее в этой главе, наиболее вероятными значениями неизвестных ξ_j являются те значения, которые обращают сумму квадратов этих остаточных разностей в минимум; это и есть принцип наименьших квадратов.

Непосредственно не очевидно, как следует решать эти уравнения, чтобы получить требуемую минимальную сумму квадратов остаточных разностей; поэтому остальная часть настоящей главы посвящается изложению метода, быстро ведущего к этой цели, и, что в равной степени важно, методу оценки неточности в значениях неизвестных, когда они окончательно определены.

7. Веса уравнений. Каждое условное уравнение может быть результатом единственного наблюдения; в этом случае, если несколько наблюдений имеют, по предположению, одинаковую точность, говорят, что эти уравнения имеют равный вес. С другой стороны, уравнение может выражать результат нескольких наблюдений, либо может быть известно, что точность различных наблюдений меняется; в любом из этих случаев может оказаться необходимым приписать этим уравнениям различные веса. Если точность различных наблюдений одинакова, то вес уравнения можно выбрать пропорциональным числу наблюдений, от которых оно зависит. Если вероятные ошибки наблюдений являются известным образом меняющимися величинами, то вес одного наблюдения следует взять обратно пропорциональным квадрату этой ошибки. Таким образом, вес p некоторого уравнения, основанного на q наблюдениях, имеющих вероятные ошибки $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_q$, дается следующей формулой:

$$p = k\left(\frac{1}{\varepsilon_1^2} + \frac{1}{\varepsilon_2^2} + \ldots + \frac{1}{\varepsilon_q^2}\right),$$

где k есть постоянная, которую необходимо выбрать так, чтобы веса различных уравнений менялись в удобных пределах, например от 1 до 9 или же от 0,1 до 0,9. Удержание в p более одной значащей цифры редко бывает оправданным, а более двух — почти никогда.

Если значение k фиксированно, а ε_j известны, то вероятная ошибка единицы веса, записываемая часто в виде в. о. 1, равна $k^{1/2}$.

Однако часто случается так, что вероятная ошибка наблюдения заранее неизвестна, но должна быть выведена из самих уравнений. В таких случаях может оказаться удобным приписать каждому наблюдению вес в единицу, и тогда вес уравнения равен числу наблюдений, от которых оно зависит.

Когда условные уравнения имеют неравные веса, то наиболее вероятным результатом не является тот, который обращает сумму квадратов остаточных разностей в минимум; вместо этого необходимо обратить в минимум сумму квадратов чисел, получающихся умножением каждой остаточной разности на корень квадратный из ее веса.

8. Составление нормальных уравнений. Допустим, что имеется m условных уравнений, содержащих n неизвестных величин, $m > n$; тогда критерию наименьших квадратов можно удовлетворить объединением этих m уравнений в n нормальных уравнений, как было указано в разд. 5, и решением этих нормальных уравнений. Вместо умножения каждого уравнения на квадратный корень из его веса может оказаться предпочтительным оперировать самим весом, организуя вычисления так, чтобы получить тот же результат, как если бы использовался этот квадратный корень. В этом случае умножим каждое уравнение на его вес и вычислим сумму всех численных величин, входящих в каждое уравнение. Результат можно записать согласно схеме табл. 2, где для удобства буквенные величины и знак равенства помещены вверху, вместо того чтобы находиться в уравнениях. Все символы ниже этой строки означают числа; частные производные, обозначенные для краткости через a, являются отвлеченными числами, тогда как Δx в уравнениях (16), обозначенные через C, обычно будут выражены в секундах дуги, градусах или радианах. Смешение размерностей в Σ не вызовет никаких трудностей.

Первое нормальное уравнение, называемое нормальным уравнением относительно ξ_1, получается умножением первого из исходных уравнений на $p_1 a_{11}$, второго — на p_2, a_{21}, третьего — на $p_3 a_{31}$ и т. д. и составлением суммы получающихся при этом m уравнений. Нормальное уравнение относительно ξ_2 получается умножением первого из исходных уравнений на $p_1 a_{12}$, второго — на $p_2 a_{22}$ и т. д. и суммированием получающихся m уравнений. Повторение этого процесса n раз доставляет n нормальных

Таблица 2

Схема для условных уравнений

ξ_1	ξ_2	ξ_3	...	ξ_n	=	C	Σ
a_{11}	a_{12}	a_{13}	...	a_{1n}		C_1	Σ_1
$p_1 a_{11}$	$p_1 a_{12}$	$p_1 a_{13}$	...	$p_1 a_{1n}$		$p_1 C_1$	$p_1 \Sigma_1$
a_{21}	a_{22}	a_{23}	...	a_{2n}		C_2	Σ_2
$p_2 a_{21}$	$p_2 a_{22}$	$p_2 a_{23}$	...	$p_2 a_{2n}$		$p_2 C_2$	$p_2 \Sigma_2$
...	...	...	...	...		...	...
...	...	...	...	...		...	...
a_{m1}	a_{m2}	a_{m3}	...	a_{mn}		C_m	Σ_m
$p_m a_{m1}$	$p_m a_{m2}$	$p_m a_{m3}$	...	$p_m a_{mn}$		$p_m C_m$	$p_m \Sigma_m$

уравнений. Коэффициенты этих нормальных уравнений следует фактически вычислять по одному, накапливая отдельные произведения. Так, например, коэффициент при ξ_2 в нормальном уравнении относительно ξ_1 равен

$$p_1 a_{11} a_{12} + p_2 a_{21} a_{22} + \ldots + p_m a_{m1} a_{m2}$$

и свободный член в нормальном уравнении относительно ξ_3 равен

$$p_1 a_{13} C_1 + p_2 a_{23} C_2 + \ldots + p_m a_{m3} C_m.$$

Легко заметить, что величины слева от знака равенства в нормальных уравнениях образуют квадратную таблицу, симметричную относительно диагонали, проведенной из верхнего левого угла в нижний правый угол. Таким образом, числа, расположенные ниже и слева от главной диагонали, вычислять не нужно.

Та же операция, произведенная над суммами, представляет ценный контроль составления нормальных уравнений. Каждый результат должен быть суммой чисел, образуемой при продвижении влево вплоть до главной диагонали, а затем вверх по столбцу. Если число условных уравнений более десяти, то обычно имеет смысл удерживать в нормальных уравнениях на один десятичный знак больше, чем в условных уравнениях; если же это число порядка сотни, то можно сохранить два лишних знака.

9. Нормальные уравнения. Нормальные уравнения могут быть записаны, как показано в табл. 3, в которой для новых чисел применяются те же обозначения, что и использованные в условных уравнениях.

Мы получили n уравнений с n неизвестными, которые необходимо решить. Симметрия, обусловленная методом, примененным для получения этих уравнений, упрощает их решение.

Из теории алгебраических уравнений известно, что система из n независимых линейных уравнений относительно n неизвестных допуска-

Таблица 3

Нормальные уравнения

ξ_1	ξ_2	ξ_3	$\ldots$	ξ_n	$=$	C	Σ
a_{11}	a_{12}	a_{13}	$\ldots$	a_{1n}		C_1	Σ_1
	a_{22}	a_{23}	$\ldots$	a_{2n}		C_2	Σ_2
		a_{33}	$\ldots$	a_{3n}		C_3	Σ_3
			$\ldots$	$\ldots$		$\ldots$	$\ldots$
				$\ldots$		$\ldots$	$\ldots$
				a_{nn}		C_n	Σ_n

ет единственное решение тогда и только тогда, если определитель из коэффициентов отличен от нуля. В этой системе уравнений определитель из коэффициентов равен

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & a_{n3} & & a_{nn} \end{vmatrix}$$

и $a_{ij} = a_{ji}$. Значение неизвестного ξ_i равно отношению двух определителей. В этом отношении знаменателем является определитель из коэффициентов, написанный выше, а числителем — определитель, который можно получить из первого подстановкой вместо i-го столбца значений C. Задача состоит в том, чтобы найти самый эффективный метод определения численных значений отношений этих различных определителей.

Хотя в строго математическом смысле решение может быть получено всегда, если определитель в знаменателе отличен от нуля, на практике часто бывает так, что этот определитель, не будучи равен нулю, является очень малой величиной, т. е. малой величиной по сравнению с произведением величин, расположенных по главной диагонали этого определителя. Например, если главная диагональ состоит из чисел порядка единицы, то 0,01 есть малая величина. Если это имеет место и если числители не являются также малыми, то физическая интерпретация заключается просто в том, что исходные наблюдения не вполне пригодны для определения искомых величин, как, например, в случае, когда пытаются определить малый угол по его косинусу. В результатах будет существовать потеря значащих цифр, и первым побуждением вычислителя будет преодолеть эту трудность путем удержания лишних десятичных знаков в уравнениях. Таким путем можно получить формальное решение с любым желаемым числом значащих цифр, однако это оказывается бесполезной затратой труда, коль скоро речь идет о каком-либо физическом смысле такого решения. На это укажут большие вероятные ошибки неизвестных. Для вычислителя имеет смысл оценивать эти ошибки, прежде чем увеличивать число значащих цифр в уравнениях, и благодаря такой оценке он во многих случаях избавит себя от ненужной работы.

Еще одна трудность, часто возникающая на практике, заключается в том, что как определитель в знаменателе, так и некоторые или все определители в числителе являются малыми величинами. В этом случае говорят, что решение является почти неопределенным, и это происходит тогда, когда два или более неизвестных находятся в точной корреляции, т. е. существует определенное множество решений. В этом случае уравнения почти точно удовлетворяются для широкого интервала значений коррелирующих неизвестных при условии, что они связаны друг с другом определенным линейным соотношением. Вообще целесообразно в таких случаях не пытаться решать уравнения относительно каждого из коррелирующих неизвестных в отдельности, а вместо этого решить их относительно линейной комбинации этих неизвестных и относительно только одного из каждой пары самих коррелирующих неизвестных. При такой методике не будет потеряна реальная информация, а вычислитель будет избавлен от большого количества работы. Кроме того, при этом гораздо легче усмотреть физический смысл полученных результатов.

Эти тонкости работы рассматриваются ниже. Сначала же дается формальное решение нормальных уравнений.

10. Формальное решение. Решению нормальных уравнений было уделено внимание многих авторов, что представляется естественным для вопроса, имеющего столь широкое применение. Теория проста и понятна, ее можно найти почти в любой книге по теории уравнений. Проблема заключается в расположении фактических вычислений в форме, требующей записи немногих чисел, чтобы сделать доступными все относящиеся к задаче сведения, содержащиеся в неявном виде в уравнениях, и обеспечить достаточный контроль численных операций. Метод, излагаемый здесь, представляет собой результат совместных

Таблица 4

Решение нормальных уравнений

	ξ_1	ξ_2	ξ_3	ξ_4	C_1	C_2	C_3	C_4	Σ
N_1	a_{11}	a_{12}	a_{13}	a_{14}	1	0	0	0	Σ_1
N_2		a_{22}	a_{23}	a_{24}	0	1	0	0	Σ_2
N_3			a_{33}	a_{34}	0	0	1	0	Σ_3
N_4				a_{44}	0	0	0	1	Σ_4
N_1	a_{11}	a_{12}	a_{13}	a_{14}	1	0	0	0	Σ_1
E_1	1	b_{12}	b_{13}	b_{14}	$1/a_{11}$	0	0	0	Σ_1/a_{11}
		d_{22}	d_{23}	d_{24}	e_1	1	0	0	Σ_{21}
E_2		1	b_{23}	b_{24}	e_1/d_{22}	$1/d_{22}$	0	0	Σ_{21}/d_{22}
			d_{33}	d_{34}	f_1	f_2	1	0	Σ_{31}
E_3			1	b_{34}	f_1/d_{33}	f_2/d_{33}	$1/d_{33}$	0	Σ_{31}/d_{33}
				d_{44}	g_1	g_2	g_3	1	Σ_{41}
E_4				1	g_1/d_{44}	g_2/d_{44}	g_3/d_{44}	$1/d_{44}$	Σ_{41}/d_{44}
S_4				1	s_{41}	s_{42}	s_{43}	s_{44}	Σ_{44}
S_3			1		s_{31}	s_{32}	s_{33}	s_{34}	Σ_{34}
S_2		1			s_{21}	s_{22}	s_{23}	s_{24}	Σ_{24}
S_1	1				s_{11}	s_{12}	s_{13}	s_{14}	Σ_{14}

усилий, приложенных в различное время многими опытными вычислителями. По-видимому, он впервые в принципе был задуман в 1878 г. М. Г. Дулиттлем, однако его детали с тех пор были сильно изменены, и приписывать этот метод в том виде, который он имеет в настоящее время, какому-нибудь одному лицу не представляется возможным. Согласно принципу, гласящему, что ни одно решение не представляет большой ценности без оценки связанных с этим решением погрешностей, любому методу, который не удовлетворяет такому требованию, не уделяется никакого внимания. Для того чтобы избежать громоздких обозначений, решение в табл. 4 дано в буквенном виде для случая четырех неизвестных. Оно сопровождается численным примером. Однажды усвоенный принцип решения легко можно распространить на любое число неизвестных.

Сначала напишем по порядку нормальные уравнения, но свободные члены запишем вверху в строку вместо того, чтобы писать их в столбец. Под каждым свободным членом напишем 1 в строке, содержащей уравнение, которому он принадлежит, и 0 — в остальных строках. Запишем контрольные суммы в правом столбце. Они отличаются от сумм, полученных при составлении нормальных уравнений, и равны

$$\begin{aligned}
\textstyle\sum_1 &= a_{11}+a_{12}+a_{13}+a_{14}+1,\\
\textstyle\sum_2 &= a_{12}+a_{22}+a_{23}+a_{24}+1,\\
\textstyle\sum_3 &= a_{13}+a_{23}+a_{33}+a_{34}+1,\\
\textstyle\sum_4 &= a_{14}+a_{24}+a_{34}+a_{44}+1.
\end{aligned}$$

1. Перепишем первое нормальное уравнение в следующую строку.

2. Умножим каждое число в строке N_1 на $1/a_{11}$, чтобы получить строку E_1, и проверим контрольную сумму, ставя «птичку» сбоку. Это есть уравнение, при помощи которого определится ξ_1 после того, как станут известны ξ_2, ξ_3, ξ_4; его часто называют элиминационным уравнением относительно ξ_1.

3. Закроем верхнюю строку N_1 полоской бумаги и вычислим

$$\begin{aligned}
d_{22} &= a_{22}-b_{12}a_{12}, & e_1 &= -b_{12},\\
d_{23} &= a_{23}-b_{12}a_{13}, & \textstyle\sum_{21} &= \textstyle\sum_2 - b_{12}\sum_1,\\
d_{24} &= a_{24}-b_{12}a_{14}, & &
\end{aligned}$$

проверяя контрольную сумму.

4. Умножим каждое число в только что написанной строке на $1/d_{22}$, чтобы получить строку E_2, и проверим контрольную сумму.

5. Закроем строки N_1 и N_2 и вычислим

$$\begin{aligned}
d_{33} &= a_{33}-b_{13}a_{13}-b_{23}d_{23}, & f_1 &= -b_{13}-b_{23}e_1,\\
d_{34} &= a_{34}-b_{13}a_{14}-b_{23}d_{24}, & f_2 &= -b_{23},\\
& & \textstyle\sum_{31} &= \textstyle\sum_3 - b_{13}\sum_1 - b_{23}\sum_{21},
\end{aligned}$$

проверяя контрольную сумму.

6. Умножим каждое число только что написанной строки на $1/d_{33}$, чтобы получить строку E_3, и проверим контрольную сумму.

7. Закроем строки N_1, N_2 и N_3 и вычислим

$$d_{44} = a_{44} - b_{14}a_{14} - b_{24}d_{24} - b_{34}d_{34},$$
$$g_1 = -b_{14} - b_{24}e_1 - b_{34}f_1,$$
$$g_2 = -b_{24} - b_{34}f_2,$$
$$g_3 = -b_{34},$$
$$\sum_{41} = \sum_4 - b_{14}\sum_1 - b_{24}\sum_{21} - b_{34}\sum_{31}.$$

8. Умножим каждое число в только что написанной строке на $1/d_{44}$, чтобы получить строку E_4, и проверим контрольную сумму.

Шаги с 1 по 8 составляют то, что известно под названием «восходящее решение». Легко видеть, что строка E_4 дает значение ξ_4, выраженное через C_1, C_2, C_3 и C_4. В нисходящем решении получаются аналогичные выражения для остальных неизвестных. В приведенной схеме строка E_4 переписывается в строку ниже и обозначается через S_4, однако здесь это сделано для удобства в обозначениях, а на практике от записи этой строки дважды никакого выигрыша не получится.

9. Строка S_3 получается при помощи формул

$$s_{31} = f_1/d_{33} - s_{41}b_{34},$$
$$s_{32} = f_2/d_{33} - s_{42}b_{34},$$
$$s_{33} = 1/d_{33} - s_{43}b_{34},$$
$$s_{34} = \qquad - s_{44}b_{34}.$$
$$\sum_{34} = \sum_{31}/d_{33} - \sum_{44}b_{34}.$$

10. Строка S_2 получается по формулам

$$s_{21} = e_1/d_{22} - s_{31}b_{23} - s_{41}b_{24},$$
$$s_{22} = 1/d_{22} - s_{32}b_{23} - s_{42}b_{24},$$
$$s_{23} = \qquad - s_{33}b_{23} - s_{43}b_{24},$$
$$s_{24} = \qquad - s_{34}b_{23} - s_{44}b_{24},$$
$$\sum_{24} = \sum_{21}/d_{22} - \sum_{34}b_{23} - \sum_{44}b_{24}.$$

11. Строка S_1 получается посредством формул

$$s_{11} = 1/a_{11} - s_{21}b_{12} - s_{31}b_{13} - s_{41}b_{14},$$
$$s_{12} = \qquad - s_{22}b_{12} - s_{32}b_{13} - s_{42}b_{14},$$
$$s_{13} = \qquad - s_{23}b_{12} - s_{33}b_{13} - s_{43}b_{14},$$
$$s_{14} = \qquad - s_{24}b_{12} - s_{34}b_{13} - s_{44}b_{14},$$
$$\sum_{14} = \sum_1/a_{11} - \sum_{24}b_{12} - \sum_{34}b_{13} - \sum_{44}b_{14}.$$

Легко видеть, что величины s_{34}, s_{23}, s_{24}, s_{12}, s_{13}, s_{14} соответственно совпадают с величинами s_{43}, s_{32}, s_{42}, s_{21}, s_{31}, s_{41}. Следовательно, их не нужно вычислять, а можно вписать по усмотрению на соответствующие места. На каждом этапе вычислений необходимо проверять контрольные суммы.

12. Теперь остается вычислить значения неизвестных по следующим формулам:

$$\xi_1 = s_{11}C_1 + s_{12}C_2 + s_{13}C_3 + s_{14}C_4,$$

$$\xi_2 = s_{21}C_1 + s_{22}C_2 + s_{23}C_3 + s_{24}C_4,$$

$$\xi_3 = s_{31}C_1 + s_{32}C_2 + s_{33}C_3 + s_{34}C_4,$$

$$\xi_4 = s_{41}C_1 + s_{42}C_2 + s_{43}C_3 + s_{44}S_4$$

и табулировать их вместе с множителями, которые, будучи умножены на вероятную ошибку единицы веса, дадут вероятные ошибки этих неизвестных. Эти множители суть соответственно $s_{11}^{1/2}$, $s_{22}^{1/2}$, $s_{33}^{1/2}$, $s_{44}^{1/2}$.

13. Далее необходимо подставить значения неизвестных в нормальные уравнения, которые должны удовлетвориться с точностью до ошибок округления.

14. Затем подставить значения неизвестных в первоначальные уравнения и вычислить остаточные разности v_j. Если величины v_j обладают гауссовым распределением, то вероятная ошибка единицы веса дается следующей формулой:

$$\text{в. о.}1 = 0{,}6745\left(\frac{\sum_{1}^{m} p_j v_j^2}{m-n}\right)^{1/2}$$

и вероятная ошибка любого неизвестного ξ_j определяется умножением на множитель, вычисленный ранее.

Если закон распределения остаточных разностей не является гауссовым, то это служит хорошим доказательством того, что наблюдения отягощены некоторыми систематическими ошибками, кроме тех, которые обозначаются через $\xi_1, \xi_2, \ldots, \xi_n$. В этом случае вероятные ошибки, вычисленные только что указанным способом, могут быть совершенно иллюзорными, а физическая интерпретация самих неизвестных — сомнительной. Эта сомнительность будет тем более серьезной, чем больше отклонение остаточных разностей от гауссова распределения.

Заметим, что если число условных уравнений равно числу неизвестных $(m = n)$, то вероятную ошибку вычислить невозможно; эти уравнения будут точно удовлетворяться, и в.о.1 будет вида 0/0. Если m превосходит n менее чем в 2 раза, то вероятная ошибка определяется неточно и заслуживает малого доверия.

11. **Численный пример.** В качестве иллюстрации содержания этой главы приводится следующий пример. Его необходимо проработать как упражнение каждому вычислителю, применяющему эту схему впервые.

Система условных уравнений вместе с их весами дана в табл. 5. Величины v являются остаточными разностями, определенными после решения.

Сначала необходимо отметить, что коэффициенты имеют различные порядки величины; например, коэффициенты при ξ_4 примерно в десять раз больше остальных. Предстоящая вычислительная работа упростится, если их привести примерно к одной и той же величине. Поэтому отдельные столбцы умножаются на следующие множители:

$$0{,}34;\ 0{,}53;\ 0{,}21;\ 0{,}068;\ 0{,}080,$$

Таблица 5

Условные уравнения

p	ξ_1	ξ_2	ξ_3	ξ_4	=	C	v
1	+1,155	+1,507	+2,622	−12,29		+9″,76	+0″,79
2	+2,417	+0,833	+4,702	+14,62		+12,43	−0 ,44
1	+0,931	+1,904	+1,554	+6,10		+5 ,48	+0 ,41
1	+2,955	+0,725	+3,324	−10,48		+6 ,98	+0 ,44
1	+1,164	+1,820	+0,949	+11,75		+1 ,60	−0 ,24
3	+1,188	+0,828	−0,016	+11,48		−2 ,37	+0 ,44
3	+1,199	+1,735	−0,647	−12,11		−4 ,49	−0 ,40
1	+0,947	+1,854	−1,318	−6,38		−6 ,30	−0 ,10

причем каждое число приблизительно является обратной величиной наибольшего числа в соответствующем столбце из приведенных выше; это равносильно введению новых неизвестных η_1, η_2, η_3, η_4, таких, что

$$\xi_1 = 4{,}25\eta_1, \qquad \xi_3 = 2{,}62\eta_3,$$
$$\xi_2 = 6{,}62\eta_2, \qquad \xi_4 = 0{,}85\eta_4.$$

Новые условные уравнения даются в табл. 6. Произведение каждого уравнения на его вес, всякий раз когда этот вес отличен от единицы, приводится непосредственно под уравнением, и добавлены контрольные суммы.

Таблица 6

Преобразованные условные уравнения

η_1	η_2	η_3	η_4	=	C	Σ
+0,393	+0,799	+0,551	−0,836		+0,781	+1,688
+0,822	+0,441	+0,987	+0,994		+0,994	+4,238
+1,644	+0,882	+1,974	+1,988		+1,988	+8,476
+0,317	+1,009	+0,326	+0,415		+0,438	+2,505
+1,005	+0,384	+0,698	−0,713		+0,558	+1,932
+0,396	+0,965	+0,199	+0,799		+0,128	+2,487
+0,404	+0,439	−0,003	+0,781		−0,190	+1,431
+1,212	+1,317	−0,009	+2,343		−0,570	+4,293
+0,408	+0,920	−0,136	−0,823		−0,359	+0,010
+1,224	+2,760	−0,408	−2,469		−1,077	+0,030
+0,322	+0,983	−0,277	−0,434		−0,504	+0,090

Нормальные уравнения и их решение представлены в табл. 7 в соответствии с описанной схемой. Контрольные суммы, использованные для проверки составления нормальных уравнений, не напечатаны.

$$\eta_1 = -0{,}604 \pm 1{,}35 \text{ в. о. } 1_\eta, \qquad \xi_1 = -2'',57 \pm 0{,}46 \text{ в. о. } 1_\xi,$$
$$\eta_2 = +0{,}111 \pm 0{,}73 \text{ в. о. } 1_\eta, \qquad \xi_2 = +0'',73 \pm 0{,}39 \text{ в. о. } 1_\xi,$$
$$\eta_3 = +1{,}532 \pm 1{,}04 \text{ в. о. } 1_\eta, \qquad \xi_3 = +4'',01 \pm 0{,}22 \text{ в. о. } 1_\xi,$$
$$\eta_4 = -0{,}030 \pm 0{,}38 \text{ в. о. } 1_\eta, \qquad \xi_4 = -0'',026 \pm 0{,}026 \text{ в. о. } 1_\xi.$$

Таблица 7

Нормальные уравнения и решение

	η_1	η_2	η_3	η_4	C_1 +1,859	C_2 +0,544	C_3 +3,238	C_4 +1,869	Σ
N_1	+3,866	+4,102	+2,464	+0,836	1	0	0	0	+12,268
N_2		+7,208	+1,448	−0,545	0	1	0	0	+13,213
N_3			+3,017	+1,747	0	0	1	0	+9,676
N_4				+8,044	0	0	0	1	+11,082
N_1	+3,866	+4,102	+2,464	+0,836	1	0	0	0	+12,268
E_1	1	+1,061	+0,637	+0,216	+0,259	0	0	0	+3,173
		+2,856	−1,166	−1,432	−1,061	1	0	0	+0,197
E_2		1	−0,408	−0,501	−0,371	+0,350	0	0	+0,069
			+0,972	+0,630	−1,070	+0,408	1	0	+1,942
E_3			1	+0,648	−1,101	+0,420	+1,029	0	+1,998
				+6,738	−0,054	+0,237	−0,648	1	+7,272
$E_4=S_4$				1	−0,008	+0,035	−0,096	0,148	+1,079
S_3			1		−1,096	+0,397	+1,091	−0,096	+1,299
S_2		1			−0,822	+0,530	+0,397	+0,035	+1,140
S_1	1				+1,831	−0,823	−1,095	−0,008	+0,903

Эти значения величин η подставляются в нормальные уравнения, и мы находим, что уравнения удовлетворяются с точностью до ошибок округления. Затем в исходные условные уравнения подставляются значения величин ξ, что дает остаточные разности v. Сумма квадратов величин v, если каждый квадрат умножен на его вес, равна 2,50, откуда

$$\text{в. о. } 1_\xi = 0{,}6745(2{,}50/4)^{1/2} = \pm 0'',53.$$

Легко видеть, что из этих восьми остаточных разностей семь меньше вероятной ошибки единицы веса, что не является существенным, так как не все остаточные разности имеют единичный вес. Вероятная величина остаточной разности веса p равна $\pm\, 0'',53/p^{1/2}$. Когда остаточные разности сравниваются со своими вероятными величинами, то видно, что четыре из них больше, а четыре — меньше последних по абсолютной величине. Кроме того, половина остаточных разностей положительна, а половина отрицательна, и ни одна остаточная разность не достигает своей удвоенной вероятной величины. Все эти условия более благоприятны, чем те, которые часто встречаются, и они указывают, что результаты имеют право на большое доверие.

Окончательные результаты таковы:

$$\xi_1 = -2'',57 \pm 0'',24, \qquad \xi_3 = +4'',01 \pm 0'',12,$$
$$\xi_2 = +0'',73 \pm 0'',21, \qquad \xi_4 = -0'',026 \pm 0'',014.$$

12. Комбинации неизвестных. После решения нередко требуется объединить два или более неизвестных. Например, при исправлении почти круговой орбиты может оказаться выгодным использовать в качестве неизвестных $\Delta(e \sin \tilde{\omega})$ и $\Delta(e \cos \tilde{\omega})$, где e и $\tilde{\omega}$ — эксцентриситет и долгота перицентра; а после этого требуется найти Δe и $\Delta\tilde{\omega}$ вместе с их вероятными ошибками. Комбинация самих неизвестных обычно не встречает никаких затруднений, однако метод определения вероятных ошибок непосредственно не очевиден.

Пусть требуется определить вероятные ошибки двух величин u и v, которые являются известными функциями от ξ_1 и ξ_2. Вероятная ошибка величины u дается формулой вида

$$\text{в. о. } 1_\xi \left[s_{11} \left(\frac{\partial u}{\partial \xi_1} \right)^2 + s_{22} \left(\frac{\partial u}{\partial \xi_2} \right)^2 + 2s_{12} \left(\frac{\partial u}{\partial \xi_1} \cdot \frac{\partial u}{\partial \xi_2} \right) \right]^{1/2},$$

а вероятная ошибка величины v — аналогичным выражением.

Часто необходима вероятная ошибка линейной комбинации неизвестных. Допустим, например, что

$$u = k_1\xi_1 + k_2\xi_2 + k_3\xi_3,$$

где величины k — постоянные. Тогда вероятная ошибка u определяется следующим выражением:

$$\text{в. о. } 1_\xi (s_{11}k_1^2 + s_{22}k_2^2 + s_{33}k_3^2 + 2s_{12}k_1k_2 + 2s_{13}k_1k_3 + 2s_{23}k_2k_3)^{1/2}.$$

Иногда может потребоваться выяснить, какая линейная комбинация из двух неизвестных, например ξ_1 и ξ_2, определяется с наименьшей вероятной ошибкой, т. е. необходимо найти значения постоянных k_1 и k_2, которые обращают выражение $s_{11}k_1^2 + s_{22}k_2^2 + 2s_{12}k_1k_2$ в минимум. Эта задача в такой формулировке является неопределенной, так как k_1 и k_2 могут быть всегда выбраны так, чтобы сделать вероятную ошибку выражения $k_1\xi_1 + k_2\xi_2$ как угодно малой. Однако если какое-либо из этих k фиксированно, то задача становится определенной. Допустим, что $k_1 = 1$. Тогда выражение, которое необходимо свести к минимуму, имеет вид $s_{11} + s_{22}k_2^2 + 2s_{12}k_2$, и оно достигает минимума, если $k_2 = -s_{12}/s_{22}$.

Найдем в предыдущем численном примере, какая комбинация из ξ_1 и ξ_2 определяется с наименьшей вероятной ошибкой. Принимая $k_2 = 1$ и оперируя сначала с величинами η, мы имеем $k_1 = 0{,}823/1{,}831 = 0{,}45$, и искомая комбинация есть $0{,}45\eta_1 + \eta_2 = -0{,}161$. Вероятная ошибка такой комбинации равна

$$\text{в. о. } 1_\eta\,[1{,}831\,(0{,}45)^2 + 0{,}530\,(1)^2 - 1{,}645\,(0{,}45)]^{1/2} = \text{в. о. } 1_\eta\,(0{,}40).$$

Теперь мы имеем комбинацию

$$0{,}45\,\eta_1 + \eta_2 = 0{,}106\,\xi_1 + 0{,}151\,\xi_2,$$

которая определяется с вероятной ошибкой, равной

$$\text{в. о. } 1_\eta\,(0{,}40) = \text{в. о. } 1_\xi\,(0{,}40)\,(0{,}080) = \text{в. о. } 1_\xi\,(0{,}0320),$$

откуда $0{,}70\,\xi_1 + \xi_2 = -1{,}07$ определяется с вероятной ошибкой, равной в. о. $1_\xi\,(0{,}21)$, или $0'',11$, которая значительно меньше вероятной ошибки одного ξ_1 либо одного ξ_2.

13. Корреляции. Всякий раз, когда линейная комбинация двух неизвестных может быть определена с вероятной ошибкой, меньшей, чем вероятная ошибка любого одного из них в отдельности, эти неизвестные называются коррелирующими. Два неизвестных ξ_1 и ξ_2 являются коррелирующими всякий раз, когда смешанный член s_{12} отличен от нуля, или же, равным образом, в большинстве случаев, когда смешанный коэффициент a_{12} нормальных уравнений отличен от нуля. Корреляции являются правилом, а не исключением при решении по методу наименьших квадратов; в численном примере этой главы корреляции существуют между каждой парой неизвестных.

По величине корреляции бывают различной степени. Чем сильнее корреляция, тем больше смешанный коэффициент a_{ij} по сравнению с $(a_{ii}a_{jj})^{1/2}$. В силу способа составления нормальных уравнений a_{ij} не может превосходить $(a_{ii}a_{jj})^{1/2}$, но может быть очень близким к этой величине. В таких случаях возникает затруднение с решением, которое является неопределенным; в ходе вычислительной работы оказывается, что одно из чисел d_{22}, d_{33}, d_{44} становится очень малым, что вызывает потерю значащих цифр в элиминационных уравнениях. Первым побуждением вычислителя может явиться повторение всех вычислений с дополнительными десятичными знаками, и формальное решение всегда можно получить этим путем, часто при огромной затрате времени и энергии. Однако оба коррелирующих неизвестных будут иметь большие вероятные ошибки и поэтому мало физического смысла. Во многих случаях можно избежать этой трудности без введения лишних десятичных знаков непосредственным решением уравнений относительно только одного из коррелирующих неизвестных, подставляя вместо другого неизвестного их линейную комбинацию. Наилучшую комбинацию, т. е. определяемую с наименьшей вероятной ошибкой, нельзя определить заранее до решения. Поэтому обычно рекомендуется сначала, если это возможно, решить уравнения, затем установить, какой комбинацией неизвестных лучше всего воспользоваться в данном случае, и повторить решение с новыми неизвестными. Хорошее приближение к наилучшей комбинации часто может быть получено таким выбором, при котором смешанный член в нормальных уравнениях обращается в нуль. Это можно сделать до того, как будет получено какое-нибудь решение уравнений следующим путем. Если рассматриваемыми неизвестными являются ξ_1 и ξ_2, заменим при решении ξ_1 на комбинацию $\xi_1 + a_{12}/a_{11}\,\xi_2$. Эта замена требует, чтобы все коэффициенты в нормальном уравнении относительно ξ_2 были изменены соответствующим образом при помощи следующих подстановок:

вместо a_{12} напишем 0,

вместо a_{22} напишем $a_{22} - (a_{12})^2/a_{11}$,

вместо a_{23} напишем $a_{23} - (a_{12}/a_{11})\, a_{13}$,

вместо a_{24} напишем $a_{24} - (a_{12}/a_{11})\, a_{14}$,

. . .

вместо C_2 напишем $C_2 - (a_{12}/a_{11})\, C_1$,

вместо $\sum_2$ напишем $\sum_2 - (a_{12}/a_{11}) \sum_1$.

Нормальное уравнение относительно ξ_1 остается без изменения, за исключением коэффициента a_{12}, и становится уравнением относительно комбинации $\xi_1 + a_{12}/a_{11}\xi_2$.

Применяя этот принцип к η_1 и η_2 из численного примера, мы получим следующую комбинацию:

$$\eta_1 + \eta_2 + (4{,}102/3{,}866) = \eta_1 + 1{,}06\,\eta_2,$$

которую необходимо подставить вместо неизвестного η_1.

Или же мы можем вместо η_2 подставить комбинацию

$$\eta_1\,(4{,}102/7{,}208) + \eta_2 = 0{,}57\,\eta_1 + \eta_2.$$

Следует отметить, что ни одна из этих комбинаций не является комбинацией, определяемой с наименьшей вероятной ошибкой, которая может быть найдена только после решения уравнений; в настоящем примере мы ничего не выиграли бы при помощи такой подстановки, так как корреляции не очень сильны.

14. Нормальные места. Если взято несколько наблюдений, распределенных на достаточно коротком интервале времени, то условные уравнения будут мало меняться на этом интервале. В таких случаях является выгодным осреднить отдельные остаточные разности $O-C$ и составить одно условное уравнение для средней эпохи наблюдений, приписывая ему вес, равный числу этих наблюдений. Остаточная разность $O-C$, полученная из нескольких наблюдений, объединенных таким образом, называется нормальным местом. Для составления нормальных мест нельзя дать определенного правила; здесь важную роль играют опыт и рассудительность вычислителя. Однако такое объединение наблюдений всегда допустимо в тех случаях, когда рассматриваемые остаточные разности можно считать равномерно меняющимися со временем. Кроме того, если наблюдения очень многочисленны, то имеет смысл представить остаточные разности параболической кривой и использовать постоянный член уравнения этой параболы в качестве свободного члена условного уравнения.

Замечания. Литература

Метод наименьших квадратов принадлежит к числу самых общих вопросов математической обработки наблюдений. Хорошей начальной книгой является руководство Брента (D. Brunt, The Combination of Observations, Cambridge Univ. Press, London and New York, 1931). Стандартным пособием является книга Уиттекера и Робинсона (E. T. Whittaker, G. Robinson, The Calculus of Observations, Blackie, London, 1924; русский перевод 2-го изд.: Э. Т. Уиттекер, Г. Робинсон Математическая обработка результатов наблюдений, М., ОНТИ, 1935).

Метод решения нормальных уравнений, данный в этой главе, часто называется элиминационным методом, или методом исключения, так как при восходящем решении неизвестные исключаются последовательно до тех пор, пока не останется одно уравнение с одним неизвестным. Другими методами, пользующимися популярностью в некоторых приложениях, являются методы квадратного корня и релаксации. Они, повидимому, не обладают преимуществами для задач, подобных рассмотренным здесь, а методы релаксации, в частности, могут привести к непригодным результатам в решениях с сильной корреляцией некоторых неизвестных. Рассмотрим для примера случай, когда два неизвестных, например x и y, сильно коррелируют. Уравнения будут удовлетворяться большим интервалом значений x и y, но этого не будет, если x и y меняются в своих пределах значений независимо друг от друга; любому допустимому значению x соответствует одно и только одно значение y, удовлетворяющее уравнениям. Метод исключения благодаря последовательному определению неизвестных делает значение x, например, зависимым от заданного заранее значения y и сохраняет необходимую связь между ними, тогда как эта связь не сохраняется при использовании релаксационных методов. Применяя эти методы, можно обнаружить в конце вычислений, что полученные результаты не удовлетворяют уравнениям.

Против предыдущих замечаний иногда возражают на том основании, что согласующиеся значения x и y имеют не больше «физической реальности», чем несогласующиеся значения, полученные при помощи релаксации. Такое возражение имеет некоторый смысл. При этом, однако, обычно не замечают того, что согласующиеся значения x и y подходят лучше других для использования при предвычислении будущих наблюдений, а также для проверки правильности выкладок при решении.