

Б. Ф. МАГУСЬКИН

О ЛАПЛАСОВСКОМ ОБОСНОВАНИИ МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ

Как известно, обоснование метода наименьших квадратов исходит из исследований таких известных ученых как Лежандр, Лаплас и Гаусс. В наше время, когда непрестанно увеличиваются требования как к точности непосредственных измерений, так и к их математической обработке, необходимо четко представлять

степень обоснованности метода наименьших квадратов, применение которого в геодезии не имеет какой-либо конкуренции. Задачей настоящей статьи является краткое изложение Лапласовского обоснования метода наименьших квадратов с последующим его анализом.

Основой обоснования по Лапласу является предельная теорема, которую называют сейчас классической предельной теоремой (или центральной предельной теоремой). Суть теоремы состоит в указании условий, при которых распределение суммы независимых случайных величин при неограниченном возрастании числа слагаемых подчиняется нормальному закону [1, 6]. На протяжении последующего столетия доказательство теоремы совершенствовалось, а формулировка ее обобщалась. Основная заслуга в этом принадлежит П. Л. Чебышеву, А. А. Маркову и особенно А. М. Ляпунову, которому принадлежит окончательный вариант формулировки и доказательства. Теорема Ляпунова существенно используется в Лапласовском способе обоснования. Поэтому приведем ее формулировку [3]:

Если для последовательности взаимно независимых случайных величин $\xi_1, \xi_2, \dots, \xi_n$ можно подобрать такое положительное число $\delta > 0$, что при $n \rightarrow \infty$

$$\frac{1}{B_n^{2+\delta}} \sum_{i=1}^n |\xi_i - M\xi_i|^{2+\delta} \rightarrow 0,$$

то при $n \rightarrow \infty$ равномерно по x вероятность

$$P \left\{ \frac{1}{B_n} \sum_{i=1}^n (\xi_i - M\xi_i) < x \right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz.$$

При этом приняты обозначения: $M\xi_i$ — математическое ожидание случайной величины ξ_i ; σ_i^2 — ее дисперсия;

$$B_n^2 = \sum_{i=1}^n \sigma_i^2.$$

Смысл условия сходимости функции распределения суммы случайных величин к нормальному закону в теореме Ляпунова состоит в том, что каждое слагаемое должно оказывать на сумму лишь незначительное и примерно одинаковое влияние.

Перейдем к изложению обоснования метода наименьших квадратов по Лапласу, принимая во внимание приведенную выше формулировку классической предельной теоремы (точнее, той ее части, которая носит название теоремы Ляпунова). Из-за изменения первоначальной формулировки предельной теоремы, а также из-за внесения некоторых других изменений, связанных с достижениями математики, изложение наше будет несколько отличаться от Лапласовского (а также от изложения результатов Лапла-

са в [1] и [5]). Однако способ и идея обоснования метода наименьших квадратов останутся неизменными.

Пусть требуется определить k неизвестных $x_1, x_2, \dots, x_k$ по n результатам измерений $q_1, q_2, \dots, q_n$ равной точности ($k < n$), если искомые неизвестные связаны с результатами измерений линейными уравнениями

$$a_i x_1 + b_i x_2 + c_i x_3 + \dots + g_i x_k + r_i = q_i + \Delta_i \quad (1)$$

($i = 1, 2, \dots, n$), где $a_i, b_i, \dots, g_i, r_i$ — известные коэффициенты; Δ_i — погрешности наблюдений.

Обозначим $r_i - q_i = l_i$. Тогда уравнения (1) примут вид

$$a_i x_1 + b_i x_2 + c_i x_3 + \dots + g_i x_k + l_i = \Delta_i \quad (2)$$

($i = 1, 2, \dots, n$).

В матричной форме это запишется так:

$$Gx + \bar{l} = \bar{\Delta}, \quad (3)$$

где

$$G = \begin{pmatrix} a_1 & b_1 & \dots & g_1 \\ a_2 & b_2 & \dots & g_2 \\ \dots & \dots & \dots & \dots \\ a_n & b_n & \dots & g_n \end{pmatrix} \quad (4)$$

— матрица уравнений погрешностей; $\bar{x} = (x_1, x_2, \dots, x_k)^T$; $\bar{l} = (l_1, l_2, \dots, l_n)^T$; $\bar{\Delta} = (\Delta_1, \Delta_2, \dots, \Delta_n)^T$; T — знак транспонирования.

Поскольку неопределенность нахождения неизвестных из (2) при $k < n$ устраняется требованием найти такие их значения, при которых пределы для ошибок были бы наименее тесными. (Смысл понятия «тесных» границ будет ясен из дальнейшего).

Умножим систему уравнений (3) слева на неопределенную матрицу G^+

$$G^+ G \bar{x} + G^+ \bar{l} = G^+ \bar{\Delta}. \quad (5)$$

Подчиним эту матрицу условию

$$G^+ G = E, \quad (6)$$

где E — единичная матрица, т. е. диагональные элементы ее равны единице, а недиагональные — нулю. Условие (6) показывает, что матрица G^+ — одна из множества обратных матриц к матрице G . Равенство (5) теперь примет вид

$$\bar{x} = G^+ \bar{l} + G^+ \bar{\Delta}. \quad (7)$$

Из (7) вытекает, что неизвестный вектор $\bar{x}$ будет определен в случае нахождения матрицы G^+ , при этом погрешность в $\bar{x}$ окажется равной $G^+ \bar{\Delta}$. Дисперсионную (ковариационную) матрицу погрешности $G^+ \bar{\Delta}$ запишем (см. [4], с. 55)

$$D(G^+ \bar{\Delta}) = G^+ D_{\Delta} G^{+T}, \quad (8)$$

где D_{Δ} — дисперсионная матрица вектора истинных ошибок Δ .
 Так как считается, что измерения независимые и равноточные,
 то

$$D_{\Delta} = \sigma^2 E. \quad (9)$$

Здесь σ^2 — дисперсия одного измерения.

Таким образом,

$$D(G^+ \bar{\Delta}) = \sigma^2 G^+ G^T. \quad (10)$$

Диагональные элементы матрицы $D(G^+ \bar{\Delta})$ суть дисперсии определяемых неизвестных x_i ($i=1, 2, \dots, k$). При использовании теоремы Ляпунова нужны только диагональные элементы, поэтому обозначив диагональную матрицу, элементами которой являются диагональные элементы матрицы $G^+ G^T$, через E_g , получим

$$\bar{B}_n = \sigma^2 E_g, \quad (11)$$

где $\bar{B}_n$ — тоже диагональная матрица, элементы которой равны дисперсиям неизвестных x_i . Согласно теореме Ляпунова при выполнении условий теоремы погрешность $G^+ \bar{\Delta}$ заключена между пределами $-\bar{B}_n y$ и $+\bar{B}_n y$ с вероятностью, стремящейся к

$$\frac{1}{\sqrt{2\pi}} \int_{-y}^{+y} e^{-\frac{z^2}{2}} dz.$$

При неизменности выбранного значения y пределы становятся тем теснее, чем меньше элементы $\bar{B}_n$. Следовательно, наилучшие значения для неизвестных x_i получим, если потребуем, устанавливая конкретное значение для матрицы G^+ , минимума для $\bar{B}_n$, понимая под этим минимизацию каждого элемента этой матрицы

$$\bar{B}_n = E_g^{1/2} \sigma = \min \quad (12)$$

Поскольку $\sigma = \text{const}$, то это приводит к требованию

$$E_g^{1/2} = \min \quad (13)$$

при соблюдении (6).

Составим функцию Лагранжа

$$\begin{aligned} \bar{f} = & E_g - \{2R(G^+ G - E)\}_d = \\ & - (G^+ G^T)_d - \{2R(G^+ G - E)\}_d. \end{aligned} \quad (14)$$

Здесь индекс d внизу означает, что берутся во внимание только диагональные элементы окончательных выражений в скобках. Найдем

$$\frac{d\bar{f}}{dG^+} = 2G^+ - 2RG^T = 0, \quad (15)$$

$$\text{откуда} \quad G^+ = RG^T. \quad (16)$$

Подставим (16) в (6):

$$RG^T G = E. \quad (17)$$

Но произведение $G^T G$ — это матрица нормальных уравнений

$$G^+ G = A, \quad (18)$$

следовательно,

$$RA = E, \quad R = A^{-1}. \quad (19)$$

Теперь матрица G^+ определена:

$$G^+ = RG^T = A^{-1}G^T. \quad (20)$$

Таким образом,

$$\bar{x} = -G^+ l = -A^{-1}G^T l. \quad (21)$$

Это решение совпадает с решением по методу наименьших квадратов. В результате последний считается обоснованным.

К такому же результату приводит требование минимизации средней нормальной погрешности. Средней нормальной положительной погрешностью называется интеграл

$$v_+ = \int_0^{\infty} \Delta \cdot p(\Delta) d\Delta, \quad (22)$$

где $p(\Delta)$ — плотность распределения погрешностей.

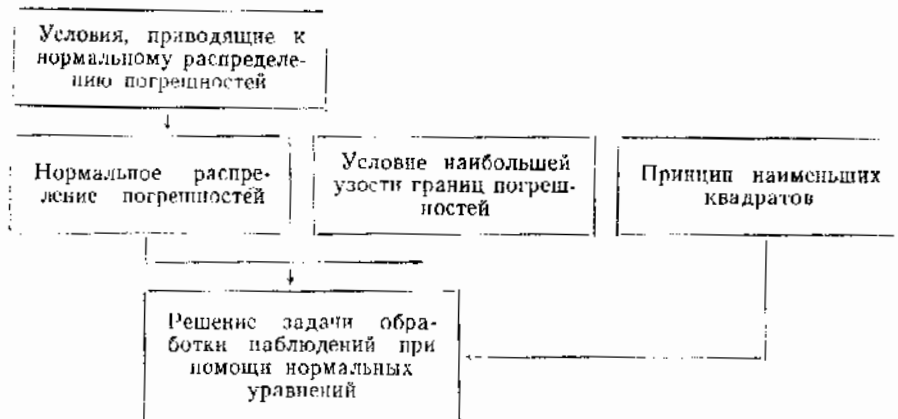
Средней нормальной отрицательной погрешностью называется

$$v_- = \left| \int_{-\infty}^0 \Delta \cdot p(\Delta) d\Delta \right|. \quad (23)$$

Обоснование по Лапласу в чебышевском изложении [5] предусматривает нахождение неизвестных при условии, требующем, чтобы были наиболее тесными вероятные пределы ошибок. Вероятными пределами ошибок называются такие пределы, если вероятность попасть внутрь пределов и вне предела одинакова и равна $\frac{1}{2}$, в остальном вывод остается тем же.

Мы не будем также рассматривать обобщение вывода на неравноточные измерения, так как схема вывода при этом не изменится. Логический ход рассуждений при обосновании по Лапласу представлен на рисунке.

Определенные условия, указанные в формулировке предельной теоремы, приводят при неограниченном увеличении числа слагаемых к нормальному распределению суммы случайных величин. Эта теорема применяется для того случая, когда суммируемыми случайными величинами являются погрешности измерений. Сле-



Логическая схема обоснования по Лапласу

довательно, допускается, что суммарные погрешности подчиняются нормальному распределению. Важнейшая из основных задач математической обработки наблюдений — нахождение по результатам измерений наилучших значений искомых неизвестных решается в условиях наибольшей «тесноты» границ для погрешностей окончательных значений неизвестных. К такому же результату приводят требование минимальности средней нормальной погрешности и требование наибольшей «тесноты» вероятных границ погрешностей неизвестных. К такому же результату приводит и применение принципа наименьших квадратов. Последнее обстоятельство считается доказательством справедливости его применения при математической обработке наблюдений.

Исходными условиями и требованиями для обоснования метода наименьших квадратов по Лапласу являются следующие:

1. Число погрешностей неограниченно большое.
2. Каждая из погрешностей оказывает на их сумму лишь незначительное влияние. Это одно из условий предельной теоремы. Непосредственно у Лапласа вместо этого условия были другие, как считается, более сильные требования: плотность распределения погрешностей измерений симметрична относительно нуля; большие по абсолютной величине ошибки встречаются реже, чем меньшие; погрешности не превосходят определенных границ.
3. Границы погрешностей окончательных значений неизвестных, находимых по результатам измерений, должны быть возможно теснее (это требование может быть заменено требованием наибольшей узости вероятных границ погрешностей или требованием минимальности средней нормальной погрешности). Как видно из предыдущего, смысл условия наибольшей тесноты границ погрешностей заключается в том, что погрешности оценок искомых неизвестных при одной и той же заданной вероятности должны лежать в возможно узких границах, или, другими словами, доверительный интервал в случае одной и той же доверительной ве-

вероятности для указанных погрешностей должен быть минимальным. Условие 3, очевидно, не вызывает никаких сомнений и не требует доказательств.

Предложение «если погрешности оценки неизвестного при любых фиксированных вероятностях лежат в наиболее узких границах, то такая оценка наилучшая» следует считать аксиомой. Это условие лучше, чем требование максимальной вероятности для искомой системы неизвестных в первом обосновании Гаусса для метода наименьших квадратов и требования минимизации дисперсии во втором обосновании Гаусса. Относительно требования максимальной вероятности системы неизвестных сам Гаусс считал, что лучше требовать минимизации возможной ошибки, чем требовать максимизации вероятности системы отдельных значений неизвестных, вероятность которой все равно остается равной нулю [2]. Надо сказать, что это замечание Гаусса касается и метода максимального правдоподобия, поскольку по существу метод максимального правдоподобия является развитием первого обоснования Гаусса, а требование максимизации функции правдоподобия эквивалентно требованию максимальной вероятности для искомой системы неизвестных. Что касается минимальности дисперсии, то Гаусс признавал произвольность этого принципа [2]. Таким образом, наличие условия 3 — это положительная сторона обоснования Лапласа.

Изложение Лапласовского обоснования метода наименьших квадратов показывает, что к недостаткам обоснования следует отнести:

1. Требование неограниченно большого числа измерений. В практике человека число измерений всегда ограничено, следовательно, даже при выполнении следующего второго условия распределение суммарных погрешностей может не быть нормальным.

2. Условие, чтобы каждая из погрешностей оказывала на окончательный результат лишь незначительное влияние. На практике, естественно, часты случаи, когда оно не выполняется.

Оба первых условия можно заменить одним — требованием распределения погрешностей измерений по нормальному закону. Но это требование ограничивает применимость метода наименьших квадратов и поэтому является недостатком обоснования.

3. Сам принцип наименьших квадратов не вытекает непосредственно из логических рассуждений и доказательств, а обосновывается путем сопоставления результатов, вытекающих из выше перечисленных условий с результатами, полученными по методу наименьших квадратов. В результате правомерность применения принципа наименьших квадратов оказывается доказанной только для решения поставленной вначале задачи обработки наблюдений, но не для других случаев применения этого принципа.

Список литературы: 1. Буняковский В. Я. Основания математической теории вероятностей. СПб., 1846. 2. Гаусс К. Ф. Избранные геодезические сочинения. т. 1. М., Изд-во геод. лит., 1957. 3. Гнеденко Б. В. Курс теории вероятностей. М.: Гос. изд-во физ.-мат. лит., 1961. 4. Линник Ю. В. Метод наимень-

ших квадратов и основы теории обработки наблюдений. — М.: Гос. изд-во физ.-мат. лит., 1958. 5. Чебышев П. М. Теория вероятностей. — М.; Л.: Изд-во АН СССР, 1936. 6. *Laplace P. S. Theorie analytique des probabilités. Paris, 1812.*

Статья поступила в редколлегию 10.05.83