ISSN 0233-7665. Кинематика и физика небес. тел. 2011. Т. 27, № 6

519.281.2:528.1:521.1

И. В. Джунь

Международный экономико-гуманитарный университет
33018 г. Ровно, ул. С. Демьянчука 4, корп. 2

Об эволюции основ метода наименьших квадратов на основе принципа максимума информации по Фишеру

Расмотрены вопросы эволюции метода наименьших квадратов на основе замены гауссовского принципа наибольшего веса фишеровским принципом максимума информации. Приведены рабочие формулы для практической реализации процедур метода наименьших квадратов с целью максимизации информации по Фишеру.

ПРО ЕВОЛЮЦІЮ ПОЛОЖЕНЬ МЕТОДУ НАЙМЕНШИХ КВАДРАТІВ НА ОСНОВІ ПРИЦИПУ МАКСИМУМУ ІНФОРМАЦІЇ ЗА ФІШЕРОМ, Джунь Й. В. — Розглянуто питання еволюції методу найменших квадратів на основі заміни гауссівського принципу найбільшої ваги фішерівським принципом максимуму інформації. Наведено робочі формули для практичної реалізації процедур методу найменших квадратів з метою максимізації інформації за Фішером.

ON THE EVOLUTION OF CONCEPTS OF THE LEAST-SQUARES METHOD ON THE BASIS OF THE FISHER PRINCIPLE OF MAXIMUM INFORMATION, by Dzhun I. V. — We consider the question on the evolution of the least-squares method on the basis of the substitution of the Fisher principle of maximum information for the Gaussian greatest weight principle. Working formulas for practical realization of the least-squares method procedures with the aim of the Fisher maximization of information are given.

К. Ф. Гаусс в работе [I] дал обоснование метода наименьших квадратов (МНК), которое базируется на принципе наибольшего веса. После него разработкой теории МНК занимались Ф. В. Бессель, П. А. Ганзен, Й. Ф. Энке, В. Йордан, Ф. Р. Гельмерт, А. А. Марков. Однако их работы не изменили основ МНК и касались лишь его деталей и задач его практического применения [5, с. 168]. А. М. Лежандр [18], впервые назвавший принцип минимума суммы квадратов поправок методом наименьших квадратов, предупреждал, что прежде чем применить этот метод, нужно тщательно просмотреть все наблюдения и выбросить те из них, которые есть или кажутся аномальными. Но проблема выбраковки наблюдений оказалась значительно более сложной, чем ожидалось. Впервые ощутили это астрономы Ф. В. Бессель [10], С. Ньюком [18, 19], С. Л. Дулитл [11, 12], Х. Р. Хюльме и Л. С. Т. Симс [13]. Суть проблемы состояла в том, что не только число аномальных результатов оказывалось на практике значительно большим, чем это следовало из закона Гаусса, но и сама форма действительных распределений ошибок была не совсем гауссовой и имела от последней стандартные отличия. Эту проблему позднее успешно разрешил Г. Джеффрис в работах [14—18]. Он показал практическую и теоретическую несостоятельность закона ошибок Гаусса при условии, если число многократных наблюдений $n > 500$. В работе [16] он рекомендует аномальные наблюдения не отбрасывать, а сохранять, придавая им меньшие веса p, вычисляемые по формуле

$$\frac{1}{p} \quad 1 \quad \frac{\quad}{\quad}^{2} \frac{m^2}{2(m \quad 0.5)^3}\,, \tag{1}$$

где — ошибка наблюдения, и m — параметры распределения Пирсона VII типа, которое он предложил для аппроксимации реальных распределений ошибок наблюдений большого объёма. Упомянутое предложение оказалось очень удачным. Это подтвердили многочисленные астрометрические исследования [2—4, 6, 8]. Эмпирические распределения ошибок действительно чаще всего соответствуют распределению Пирсона VII типа.

Но предложив более совершенный закон ошибок и систему весов (1), Джеффрис на этом остановился и не предпринимал попыток эволюции основ МНК. На наш взгляд, для такой эволюции наиболее полезным является понятие информации, введенное Фишером. Оно должно удовлетворять таким требованиям [20]:

1. Информация, содержащаяся в выборке, должна возрастать пропорционально числу наблюдений: удвоение их числа должно удваивать количество информации, если ошибки измерений независимы.

2. Информация должна быть связана исключительно с задачами астрометрического эксперимента. Данные, не имеющие отношения к проверяемой гипотезе, не должны иметь какой-либо информации.

3. Информация должна быть связана с точностью: чем выше точность наблюдений, тем больше информации мы получаем. Для астрономии именно точность инструментальных средств является решающим показателем ее развития.

Очевидно, что понятие, имеющее названные свойства, является очень ценным при постановке и анализе астрономических наблюдений.

Обозначим через $y(x_i;\ _1,\ _2,\ldots)$ плотность вероятности, зависящую от параметров для n наблюдений x_i, которые мы рассматриваем как n случайных величин. Функция максимального правдоподобия для фиксированных значений будет иметь вид

$$L \quad \prod_{i\ 1}^{n} y(x_i \quad _1, \quad _2, \ldots). \tag{2}$$

По определению Фишера объём информации об искомом параметре , содержащийся в наблюдении x, выражается формулой

$$I_x(\ ) \qquad \frac{\ln L}{\ }^{2} y(x_i, \ )dx. \tag{3}$$

Информация (3) имеет ряд ценных математических свойств, описанных в работе [20], а, также свойство, позволяющее существенно упростить вычисления:

$$\mathrm{E} \ \frac{\ln L}{\ }^{2} \qquad \mathrm{E} \ \frac{^2 \ln L}{\ ^2}, \tag{4}$$

где Е — оператор математического ожидания.

При -мерном   выражение (3) превращается в информационную матрицу размера   , элемент которой имеет вид

$$I_x(\ _i \ _j) \qquad \frac{^2 \ln L}{\ _i \ _j} y(x_i; \ _1, \ _2, \ldots)dx. \tag{5}$$

Пусть переменная x распределена нормально с известной дисперсией   и неизвестным средним a:

$$y(x_i; a, \ ) \quad \frac{1}{\sqrt{2 \ }} \exp \quad \frac{(x \quad a)^2}{2 \ ^2} \quad . \tag{6}$$

Тогда в соответствии с формулой (4) информация, которая содержится в одном наблюдении x, равна

$$I_1(a) \qquad \frac{^2}{a^2} \quad \ln \ \sqrt{2 \ } \quad \frac{(x \quad a)^2}{2 \ ^2} \ y(x, \ , a)dx \quad \frac{1}{\ ^2}. \tag{7}$$

Как видим из формулы (7), в случае нормального распределения ошибок вес наблюдения — это мера количества информации по Фишеру.

В случае n наблюдений имеем для распределения (6)

$$L \quad \prod_{i\ 1}^{n} y(x_i; \ , a).$$

В соответствии с (4) получим

$$I_n(a) \qquad \frac{^2}{a^2} \qquad \ln \ \sqrt{2 \ } \quad \frac{(x \quad a)^2}{2 \ ^2} \quad f(x; \ , a)dx \tag{8}$$

$$= n \ I_1(a) \quad \frac{n}{\ ^2}.$$

Соотношение (8) ясно выражает суть практических усилий астрономов: увеличение количества информации достигается двумя путями повышением точности и увеличением числа наблюдений.

Для закона ошибок Джеффриса, т. е. для модифицированного им распределения Пирсона VII типа [15]

$$y(x_i;\ ,\ _{\mathrm{VII}},m)=\frac{(m\quad 1)}{\sqrt{2\ (m\quad 0.5)}\ (m\quad 0.5)}$$

$$\frac{1}{_{\mathrm{VII}}}\ 1\quad \frac{0.5m^2}{(m\quad 0.5)^3}\ \frac{x}{_{\mathrm{VII}}}\ ^{2\quad m} \qquad (9)$$

имеем такое значение информации по Фишеру для случая n наблюдений, которое получено в предположении, что $_{\mathrm{VII}}$ и m известны:

$$I_n(\ )\quad E\ \frac{^2\ln L}{^2}\quad \frac{n}{^2_{\mathrm{VII}}}\ \frac{m^3}{(m\quad 0.5)^2(m\quad 1)}. \qquad (10)$$

Распределение (9) при m является нормальным. В этом случае соотношение (10) становится идентичным выражению (8). В отличие от результата (8), формула (10) демонстрирует один парадокс: информация увеличивается не только с увеличением n и уменьшением $^2_{\mathrm{VII}}$, она увеличивается и с увеличением уклонения распределения Пирсона VII типа от закона Гаусса, увеличиваясь с уменьшением m.

Нынешний уровень развития астрометрии, отличающийся все большими объёмами наблюдений, требует выхода на качественно новые уровни процедур МНК, учитывающих максимизацию фишеровской информации. Как практически осуществить эту максимизацию в соответствии с отношением (10)? Очевидно, для этого нужно получить эффективные оценки параметров $_{\mathrm{VII}}$ и m в (10) для закона ошибок Джеффриса (9). Кроме того, нужно знать еще алгоритм реализации принципа максимума информации по Фишеру. Этот принцип легко осуществить на практике, не меняя обычных процедурных схем классического МНК и внося лишь незначительные изменения в его программное обеспечение. Практически реализация принципа максимума информации по Фишеру может быть осуществена в три этапа.

Вначале применяют классический МНК, основываясь на принципе наибольшего веса Гаусса. На этом же этапе получаем разности «observation calculation» $_i$ и определяем несмещенные оценки их асимметрии и эксцесса [7, с. 423]:

$$A\quad \frac{\sqrt{n(n\quad 1)}}{n\quad 2}\frac{_3}{^3},$$

$$\frac{n\quad 1}{(n\quad 2)(n\quad 3)}\ (n\quad 1)\ \frac{_4}{_2^2}\quad 3\quad 6\ , \qquad (11)$$

где $_r$— центральные выборочные моменты порядка r; $\sqrt{\ }$.

Далее определяем дисперсии ^2_A и ^{2} статистик A и по формулам [7]:

$$^2_A \quad \frac{6n(n \quad 1)}{(n \quad 2)(n \quad 1)(n \quad 3)}, \qquad\qquad (12)$$

$$^2 \quad \frac{24n(n \quad 1)^2}{(n \quad 3)(n \quad 2)(n \quad 3)(n \quad 5)}.$$

На основе (12) строим 90 % доверительные интервалы для A и по формулам

$$A \quad 1.645 \quad _A, \qquad\qquad 1.645 \quad . \qquad\qquad (13)$$

Если оба доверительные интервалы (13) накрывают нуль, можно ограничиться применением классического МНК. В дальнейших приближениях нет необходимости. Полученные результаты можно считать окончательными.

Мы не рассматриваем здесь патологические случаи применения классического МНК, когда доверительный интервал для A не накрывает нуль, или когда весь доверительный интервал для находится в отрицательной области. Эти случаи требуют специального анализа.

Ко второму этапу вычислений приступают в том случае, когда доверительный интервал для A накрывает нуль, а весь доверительный интервал для находится в области положительных значений. Такого рода распределения остаточных ошибок удовлетворительно представляются кривой Пирсона VII типа (9), которую напишем в более компактном виде:

$$y(\ _i; \ , \ _{\mathrm{VII}}, m) \quad \frac{c}{_{\mathrm{VII}}} R^{\ m}, \qquad\qquad (14)$$

где $_i$ — значения остаточных ошибок,

$$c \qquad (m \quad 1)[\sqrt{2 \ (m \quad 0.5)} \ (m \quad 0.5)]^{\ 1},$$

$$R \quad 1 \quad (x \quad a)^2 \ \ ^2_{\mathrm{VII}} 0.5M^{\ 1}, \qquad M \quad (m \quad 0.5)^3 m^{\ 2}.$$

Затем находим эффективные оценки параметров распределения (14), дифференцируя функцию максимального правдоподобия L по $_{\mathrm{VII}}$ и m, используя систему уравнений, предложенную Джеффрисом [15] и видоизмененную в работе [4]:

$$\frac{\ln L}{} \quad \frac{m}{M \ ^2_{\mathrm{VII}}} \ _{i \ 1}^{\ n} R_i^{\ 1}(\ _i \qquad) \quad 0, \qquad\qquad (15)$$

$$\frac{\ln L}{_{\mathrm{VII}}} \quad \frac{n}{_{\mathrm{VII}}} \quad \frac{m}{M \ ^3_{\mathrm{VII}}} \ _{i \ 1}^{\ n} R_i^{\ 1}(\ _i \qquad)^2 \quad 0, \qquad\qquad (16)$$

$$\frac{\ln L}{m} \quad n \ \ _0 \ \ _{i \ 1}^{\ n} \ln R_i \ \ _{i \ 1}^{\ n} M_1 R_i^{\ 1} \ \frac{_i}{_{\mathrm{VII}}}^{\ 2} \quad 0, \qquad\qquad (17)$$

где $_0 =$ $(m + 1)$ $(m + 0.5)$ – [2(m , m пси-функция, $M_1 = 0.5m^2(m + 1)(m \quad 0.5)^{-4}$.

Система уравнений (15) 17) решается методом приближений. В первом приближении принимается $=$ $_i / n$, $^2_{\mathrm{VII}} = 0.933$ $_2$, $m = 4$.

С целью вычисления веса ошибки напишем следующее соотношение, предполагая, что L зависит только от , а значения $_{\mathrm{VII}}$ и m для распределения (14) известны:

$$\frac{\ln L}{} \qquad \frac{y\,(\ \ _i;\ ,\ _{\mathrm{VII}}, m)}{y(\ \ _i;\ ,\ _{\mathrm{VII}}, m)}. \tag{18}$$

Умножая числитель и знаменатель в (18) на $_i$ получаем оценку для , предполагающую применение метода последовательных приближений:

$$_i \qquad \frac{_i\, p(\ \ _i \qquad _{i\ 1})}{p(\ \ _i \qquad _{i\ 1})}. \tag{19}$$

В формуле (19) индекс i показывает номер приближения, а веса $p(\ \ _i$ определяются по формуле

$$p(\ \ _i \qquad _{i\ 1}) \qquad \frac{y\,(\ \ _i;\ ,\ _{\mathrm{VII}}, m)}{(\ \ _i \qquad _{i\ 1})\, y(\ \ _i;\ ,\ _{\mathrm{VII}}, m)}. \tag{20}$$

Если придерживаться второго постулата МНК, введенного Гауссом, и считать нулевыми математические ожидания величин , то формула для определения весов приобретает окончательный вид:

$$p(\ \ _i) \qquad \frac{y\,(\ \ _i;\ ,\ _{\mathrm{VII}}, m)}{_i\, y(\ \ _i;\ ,\ _{\mathrm{VII}}, m)} \qquad \frac{m \quad 0.5}{m}^{3} \quad ^{2}_{\mathrm{VII}} \quad \frac{^2_i}{2m}^{1}, \tag{21}$$

где $_{\mathrm{VI}}$ и m оценки параметров распределения Пирсона VII типа для остаточных погрешностей, полученные из решения системы (15) (17).

Практика показывает, что незначительные отличия $_i / n$ от нуля не оказывают какого-либо существенного воздействия на значения весов $p(\ \ _i)$.

После получения весов (21) приступают к третьему этапу вычислений. Для этого каждое уравнение ошибок $_i$ в схемах МНК нормируют, умножая его на вес $p(\ \ _i)$, вычисленный по формуле (21), и получают второй раз оценки искомых параметров, но уже при условии $^2_i\, p(\ \ _i) = \min$.

Искомые параметры (поправки) получают по формуле

$$l_j \quad D_{l_j} / D, \tag{22}$$

где D детерминант системы нормальных уравнений, D_{l_j} определители соответствующих неизвестных l_j, $j = 1, 2, ..., k$.

Стандартные ошибки найденных значений l_j получим из соотношений

$$
{}^{2}_{l_j} \qquad {}^{2}_{0}\frac{A_{jj}}{D}, \qquad (23)
$$

$$
{}^{2}_{0} \qquad \frac{{}^{2}_{i}\, p(\ \ _{i})}{n \quad k},
$$

где A_{jj} миноры соответствующих диагональных элементов системы нормальных уравнений.

В менее ответственных случаях веса $p(\ \ _{i})$ остаточных погрешностей $_{i}$ можно получать значительно проще, без решения системы (15)—(17), используя следующую весовую функцию, впервые полученную в работе [2]:

$$
p(\ \ _{i}) \qquad \frac{5 \quad 6}{2 \ {}^{2} \ \ _{2} \qquad {}^{2}_{i}}, \qquad (24)
$$

$$
_{2} \qquad 3,
$$

где — эксцес распределения остаточных погрешностей, полученный еще на первом этапе вычислений.

В результате нашего исследования можно сделать следующие выводы.

1. При обработке астрономических наблюдений очень важно учитывать реальный закон распределения разностей «observation calculation», так как информация по Фишеру увеличивается не только с уменьшением , но и с уменьшением параметра распределения Пирсона VII типа. При этом заметим, что для закона Гаусса $m =$.

2. С увеличением числа наблюдений разности «observation calculation» приобретают все более отчетливо выраженный негауссов характер. Эти разности почти всегда следуют распределению Пирсона VII типа. Они имеют, как правило, на порядок и более различающиеся веса, что обусловлено непрерывным изменением условий астрономических наблюдений. Поэтому важность предложеной нами эволюционной схемы МНК будет возрастать вместе с возрастанием объёмов выборок в астрометрических экспериментах.

Наше рассмотрение можно закончить словами П. Т. Шардена о значении эволюции [9]: «Что такое эволюция теория, система, гипотеза?... Нет, нечто гораздо большее, чем все это: она основное условие, которому должны подчиняться и удовлетворять все теории, гипотезы, системы, если они хотят быть разумными и искренними».

1. *Гаусс К. Ф.* Избранные геодезические сочинения / Под ред. Г. В. Багратуни. Пер. с лат. и нем. — М.: Изд-во геодез. лит., 1957.—Т. 1. Способ наименьших квадратов.—234 с.

2. *Джунь И. В.* Анализ параллельных широтных наблюдений, выполненных по общей программе: Автореф. дис. ... канд. физ-мат. наук. — Киев: Ин-т математики АН Украины, 1974.—19 с.

3. *Джунь И. В*. Распределение Пирсона VII типа ошибок лазерных наблюдений ИСЗ // Кинематика и физика небес. тел.—1991.—**7**, № 3.—С. 82—91.

4. *Джунь И. В*. Математическая обработка астрономической и космической информации при негауссовых ошибках наблюдений. Автореф. дис. ... д-ра физ.-мат. наук. — Киев: ГАО АН Украины, 1992.—46 с.

5. *Кемниц Ю. В*. Теория ошибок измерений. 2-е изд., перераб. и доп. М.: Недра, 1967.—176 с.

6. *Король А. К*. Склонения ярких и слабых фундаментальных звезд в единой системе. — Киев: Наук. думка, 1969.—234 с.

7. *Крамер Г*. Математические методы статистики. — М.: Мир, 1978.—648 с.

8. *Харин А. С., Яцкив Я. С.* Изучение ошибок наблюдений Голосеевского каталога звезд широтных программ // Астрометрия и астрофизика.—1970.—Вып. 10.—С. 34—43.

9. *Шарден П. Т.* Феномен человека. — М.: Наука, 1987.—286 с.

10. *Bessel F. W*. Fundamenta astronomiae. — Königsberg, 1818.

11. *Doolittle C. L*. Results of observation with the zenith-telescope and the Wharton reflex zenith tube // Astron. J.—1910.—**26**, N 608.—P. 12—24.

12. *Doolittle C. L*. Results of observation with the zenith-telescope and the Wharton reflex zenith tube // Astron. J.—1912.—**27**, N 641.—P. 46—62.

13. *Hulme H. R., Syms L. S. T*. The law of error and the combination of observations // Mon. Notic. Roy. Astron. Soc.—1939.—**99,** N 8.—P. 642—658.

14. *Jeffreys H.* The law of error and the combination of observations // Phil. Trans. Roy. Soc. London A.—1937.—N 237.—P. 231—271.

15. *Jeffreys H*. The law of error in the Greenwich variation of latitude observations // Mon. Notic. Roy Astron. Soc.—1939.—**99**, N 9.—P. 703—709.

16. *Jeffreys H.* Theory of Probability. Sec. Edition. — Oxford, 1940.—468 p.

17. *Legendre A. M*. On the method of least sguares (1805) Translated from the french in “A source book in mathematic” / Ed. D. E. Smith. — New York: Dover Publication Inc., 1989.—P. 576—579.

18. *Newcomb S. A.* Generalized theory of the combination of observations so as to obtain the best result // Amer. J. Math.—1886.—N 1/14.—P. 343—366.

19. *Newcomb S*. Researches of the motion of the Moon // Astron. Pap. Publ. US Naut. Office.—1912.—**9**.—P. 1—249.

20. *Statistical* methods in experimental physics / W. T. Eadie, D. Dryard, F. E. James, M. Roos, B. Sadoulet. — Cern, Geneva and University of Helsinki, Amsterdam — London: North-Holland Publisching Company, 1971.—330 p.

Поступила в редакцию 11.02.11