Math-Net.Ru
Общероссийский математический портал

А. Л. Брудно, К дисперсионному обоснованию метода наименьших квадратов,
Матем. сб., 1957, том 85, номер 1, 37–48

https://www.mathnet.ru/sm5071

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
https://www.mathnet.ru/rus/agreement

Параметры загрузки:
IP: 18.97.14.84
16 января 2026 г., 22:36:36

К дисперсионному обоснованию метода наименьших квадратов

А. Л. Брудно (Москва)

§ 1. Введение. Пусть y является однородной линейной функцией

$$y = \sum_{k=1}^{n} x_k c_k \tag{1.1}$$

n переменных

$$x_1, \ldots, x_n$$

с фиксированными, но неизвестными коэффициентами c_k. Пусть, далее, заданы m комплексов

$$x_{i1}, \ldots, x_{ni} \quad (i = 1, \ldots, m)$$

значений переменных $x_1, \ldots, x_n$ и измеряются m чисел

$$y_i = \sum_{k=1}^{n} x_{ki} c_k \quad (i = 1, \ldots, m). \tag{1.2}$$

По разным причинам (в том числе из-за ошибок, допущенных при измерениях) вместо значений y_i измерения дают значения

$$\eta_i = y_i + \Delta_i. \tag{1.3}$$

Будем предполагать выполненными следующие условия:

1.1) Добавки Δ_i являются независимыми случайными величинами с нормальным распределением, математическими ожиданиями $\mathbf{M}\Delta_i = 0$ и одинаковыми дисперсиями $\mathbf{D}^2\Delta_i = \Delta^2$.

1.2) Значения x_{ki}, y_i дают возможность однозначно определить из уравнений (1.2) функцию (1.1), т. е. найти коэффициенты c_k. Это значит, что ранг матрицы коэффициентов x_{ki} уравнений (1.2) равен n.

Пусть, наконец, задан комплекс

$$x_{10}, \ldots, x_{n0} \tag{1.4}$$

и требуется наилучшим образом указать значение

$$y_0 = \sum_{k=1}^{n} x_{k0} c_k \tag{1.5}$$

по уже известным x_{ki}, η_i, x_{k0}.

§ 2. Метод наименьших квадратов. Согласно методу наименьших квадратов (Гаусс) рекомендуется принять за приближенное значение числа y_0 величину θ_0

$$\theta_0 = \sum_{k=1}^{n} x_{k0}\zeta_k, \tag{2.1}$$

где коэффициенты ζ_k имеют значения, обращающие в минимум величину

$$\sum_{i=1}^{m}\left(\eta_i - \sum_{k=1}^{n} x_{ki}\zeta_k\right)^2. \tag{2.2}$$

Замечание. Коэффициенты ζ_k, обращающие в минимум величину (2.2), можно найти из системы линейно независимых уравнений

$$\sum_{k=1}^{n}\left(\sum_{i=1}^{m} x_{ki}x_{si}\right)\zeta_k = \sum_{i=1}^{m} x_{si}\eta_i \quad (s = 1, \ldots, n). \tag{2.3}$$

По этому и *дальнейшим* вопросам метода наименьших квадратов отсылаю читателя к заметке [1], изложение которой я не могу существенно улучшить.

§ 3. Постановка задачи. Обозначим через

$$\theta = \theta([x_{k0}], [x_{ki}], [\eta_i]) \quad (k = 1, \ldots, n;\ i = 1, \ldots, m) \tag{3.1}$$

функцию коэффициентов $[x_{k0}]$, матрицы $[x_{ki}]$ и случайных значений $[\eta_i]$.

Обозначим через Θ множество всех функций $\theta([x_{k0}], [x_{ki}], [\eta_i])$, обладающих следующим свойством: каковы бы ни были неизвестные, но фиксированные коэффициенты $[c_k]$, математическое ожидание

$$\mathbf{M}\theta = y_0. \tag{3.2}$$

Имеет место

Теорема. *Функция Гаусса принадлежит множеству* Θ,

$$\theta_0 \in \Theta. \tag{3.3}$$

Если $\theta \in \Theta$ *и для каких-нибудь фиксированных* $[x_{k0}]$, $[x_{ki}]$, $[c_k]$ *функция* $\theta \neq \theta_0$ *на множестве положительной меры значений* $[\eta_i]$, *то для этих* $[x_{k0}]$, $[x_{ki}]$, $[c_k]$ *будет*

$$\mathbf{D}^2\theta_0 < \mathbf{D}^2\theta. \tag{3.4}$$

Может быть, не лишним будет подчеркнуть, что по раз навсегда заданным точкам $\bar{x}_0 = [x_{k0}]$, $x_i = [x_{ki}]$ и по полученным при помощи измерений числам $[\eta_i]$ мы ищем наилучшее приближение для значения y_0 функции y *в одной единственной заранее заданной точке* $\bar{x}_0$.

Сформулированную теорему доказали Ф. Н. Давид и И. Нейман [2] в предположении, что Θ содержит только линейные функции переменных $[\eta_i]$ (на зависимость θ от $\bar{x}_0$, $\bar{x}_i$ ограничений не накладывается). Затем эту теорему доказал В. В. Петров [3] в более широких предположениях. Содержание настоящей заметки составляет доказательство этой же теоремы без ограничений. Математическое ожидание понимается в смысле интеграла Лебега.

§ 4. Доказательство первой части Теоремы. Надо показать, что функция Гаусса θ_0 (см. § 2) входит в класс $\boldsymbol{\Theta}$ и имеет конечную дисперсию. Это доказано самим Гауссом, легко проверяется и широко известно (см., например, [1]).

§ 5. Доказательство второй части Теоремы. Доказательство второй части Теоремы мы разобьем на пункты.

5.1. Из аргументов $\theta([x_{k0}], [x_{ki}], [\eta_i])$ опустим, для краткости, $[x_{k0}]$, $[x_{ki}]$, подставим

$$\eta_i = y_i + \Delta_i = \sum_{k=1}^{n} x_{ki} c_k$$

и рассмотрим математическое ожидание

$$\mathbf{M}\theta = \left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^m \int \theta\left(\left[\sum_{k=1}^{n} x_{ki} c_k + \Delta_i\right]\right) \prod_{i=1}^{m} e^{-\Delta_i^2 : 2\Delta^2} d\Delta_i. \qquad (5.1.1)$$

По условиям Теоремы интеграл Лебега существует.

В m-мерном пространстве $O\Delta_1, \ldots, \Delta_m$ введем новую ортонормированную систему координат $\overline{Ou_1}, \ldots, \overline{Ou_m}$. Для этого рассмотрим линейно независимые (см. 1.2)) векторы

$$\sum_{i=1}^{m} x_{ki} \overline{O\Delta_i} \quad (k = 1, \ldots, n), \qquad (5.1.2)$$

где $\overline{O\Delta_i}$ — орты соответствующих осей. Последние $m-n$ ортов $\overline{Ou_m}, \ldots, \overline{Ou_{n+1}}$ выберем (попарно ортогональными) в $m-n$-мерном пространстве, ортогональном n-мерному линейному пространству векторов (5.1.2). Остальные n ортов $\overline{Ou_n}, \ldots, \overline{Ou_1}$ последовательно выберем так, чтобы они были ортогональны уже выбранным ортам и, кроме того, чтобы

$$\overline{Ou_s} \perp \sum_{i=1}^{m} x_{ki} \overline{O\Delta_i} \quad (k = 1, \ldots, s-1;\ s = n, n-1, \ldots, 2). \qquad (5.1.3)$$

Положение $\overline{Ou_1}$ тем самым уже определено. Заметим, что

$$\overline{Ou_k} \text{ не ортогонально } \sum_{i=1}^{m} x_{ki} \overline{O\Delta_i} \quad (k = 1, \ldots, n). \qquad (5.1.4)$$

Действительно пространство, порожденное векторами

$$\overline{Ou_m}, \ldots, \overline{Ou_{s+1}}, \qquad (5.1.5)$$

ортогонально подпространству, порожденному

$$\sum_{i=1}^{m} x_{1i} \overline{O\Delta_i}, \ldots, \sum_{i=1}^{m} x_{si} \overline{O\Delta_i} \qquad (5.1.6)$$

($s = 1, \ldots, n$), и вместе эти подпространства порождают все m-мерное пространство $O\Delta_1, \ldots, \Delta_m$. Вектор $\overline{Ou_s}$ ортогонален векторам (5.1.5); следовательно, он лежит в подпространстве, порожденном векторами (5.1.6). Но согласно (5.1.3) вектор $\overline{Ou_s}$ ортогонален всем векторам (5.1.6), кроме последнего. Так как $\overline{Ou_s} \neq 0$, то утверждение (5.1.4) доказано.

Переход от старых осей к новым будем обозначать при помощи линейного оператора L, так что

$$[u_i] = L[\Delta_j], \quad [\Delta_i] = L^{-1}[u_j], \quad u_i = L_i[\Delta_j], \quad \Delta_i = L_i^{-1}[u_j]. \tag{5.1.7}$$

Перейдем к новым координатам в интеграле (5.1.1). Имеем:

$$\mathbf{M}\theta = \left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^m \int \theta\left(\left[\sum_{k=1}^{n} x_{ki}c_k + L_i^{-1}[u_j]\right]\right) \prod_{i=1}^{m} e^{-u_i^2 : 2\Delta^2} du_i, \tag{5.1.8}$$

ибо сумма квадратов $\Delta_1^2 + \ldots + \Delta_m^2 = u_1^2 + \ldots + u_m^2$, в силу ортонормированности старой и новой систем координат.

Введем функцию f, полагая

$$\theta([\eta_i]) = f(L[\eta_j]). \tag{5.1.9}$$

Функция f зависит еще и от невыписанных аргументов $[x_{k0}]$, $[x_{ki}]$. Заменим функцию θ на f в подынтегральном выражении (5.1.8):

$$\theta\left(\left[\sum_{k=1}^{n} x_{ki}c_k + L_i^{-1}[u_j]\right]\right) = f\left(L\left[\sum_{k=1}^{n} x_{ki}c_k + L_i^{-1}[u_j]\right]\right) =$$

$$= f\left(\left[\sum_{k=1}^{n} c_k L_i[x_{kj}] + u_i\right]\right) =$$

$$= f\left(u_1 + \sum_{k=1}^{n} y_{1k}c_k,\ u_2 + \sum_{k=2}^{n} y_{2k}c_k,\ \ldots,\ u_n + y_{nn}c_n,\ u_{n+1},\ \ldots,\ u_m\right); \tag{5.1.10}$$

здесь введены коэффициенты

$$y_{ki} = L_i[x_{kj}] \tag{5.1.11}$$

и учтено, что, согласно соотношениям ортогональности (см. (5.1.3)),

$$L_i[x_{kj}] = 0 \tag{5.1.12}$$

в случае $i \leqslant n$ для $k < i$ и в случае $i \geqslant n$ для всех k.

Особенно важно для дальнейшего, что преобразование L и, следовательно, коэффициенты y_{ki} не зависят от неизвестных коэффициентов c_k. Наконец, отметим, что

$$y_{kk} \neq 0 \quad (k = 1, \ldots, n), \tag{5.1.13}$$

в силу (5.1.4).

5.2. Подставив правую часть (5.1.10) в (5.1.8), получим:

$$\mathbf{M}\theta = \left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^m \int f\left(\left[u_i + \sum_{k=i}^{n} y_{ki}c_k\right]\right) \prod_{i=1}^{m} e^{-u_i^2 : 2\Delta^2} du_i \qquad (5.2.0)$$

(при исчезновении слагаемых сумма $\sum y_{ki}c_k$ считается равной нулю). По условию Теоремы для $\theta \in \boldsymbol{\Theta}$

$$\mathbf{M}\theta = \sum_{k=1}^{n} x_{k0}c_k$$

тождественно по всем значениям $[c_k]$. Для удобства дальнейших выкладок введем функцию $\mathfrak{L}_0$

$$\mathfrak{L}_0(c_1, \ldots, c_n) \equiv \sum_{k=1}^{n} \lambda_{k0}c_k = \sum_{k=1}^{n} x_{k0}c_k$$

и отметим, что $\mathfrak{L}_0$ — вполне определенная однородная линейная функция своих аргументов с коэффициентами, не зависящими от $c_1, \ldots, c_n$. Собирая три последних равенства, получаем:

$$\frac{1}{\sqrt{2\pi\Delta^2}} \int e^{-u_1^2 : 2\Delta^2} du_1 \left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^{m-1} \int f\left(\left[u_i + \sum_{k=i}^{n} y_{ki}c_k\right]\right) \prod_{i=2}^{m} e^{-u_i^2 : 2\Delta^2} du_i =$$
$$= \mathfrak{L}_0(c_1, \ldots, c_n). \qquad (5.2.1)$$

По известной теореме Фубини для каждого набора $[c_k]$ внутренний интеграл существует для почти всех значений u_1. Обозначим его через $\mathfrak{L}_1$:

$$\mathfrak{L}_1 = \left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^{m-1} \int f\left(\left[u_i + \sum_{k=i}^{n} y_{ki}c_k\right]\right) \prod_{i=2}^{m} e^{-u_i^2 : 2\Delta^2} du_i. \qquad (5.2.2)$$

$\mathfrak{L}_1$ является функцией $u_1 + \sum_{k=1}^{n} y_{k1}c_k; c_2, c_3, \ldots, c_n$. Подставив $\mathfrak{L}_1$ в (5.2.1), получим:

$$\frac{1}{\sqrt{2\pi\Delta^2}} \int e^{-u_1^2 : 2\Delta^2} \mathfrak{L}_1\left(u_1 + \sum_{k=1}^{n} y_{k1}c_k;\ c_2, \ldots, c_n\right) du_1 = \mathfrak{L}_0(c_1, \ldots, c_n). \qquad (5.2.3)$$

При фиксированных параметрах $c_2, \ldots, c_n$ это—свертка, и, согласно § 6 (см. ниже), единственная возможность для $\mathfrak{L}_1$ — та, что $\mathfrak{L}_1$ — линейная функция своего первого аргумента (с точностью до значений на множестве меры нуль):

$$\mathfrak{L}_1(t_1;\ c_2, \ldots, c_n) = t_1A_1 + B_1, \qquad (5.2.4)$$

где константы A_1, B_1 зависят от $c_2, \ldots, c_n$, но не от t_1. Для вычисления A_1 и B_1 подставим (5.2.4) в (5.2.3), что дает:

$$A_1 \sum_{k=1}^{n} y_{k1} c_k + B_1 = \mathfrak{L}_0 (c_1, \ldots, c_n) = \sum_{k=1}^{n} \lambda_{k0} c_k \tag{5.2.5}$$

тождественно по всем $c_1, \ldots, c_n$. Из тождественности по c_1 (от которого не зависят A_1, B_1) получаем:

$$A_1 = \lambda_{10} : y_{11}, \quad B_1 = \sum_{k=2}^{n} (\lambda_{k0} - A_1 y_{k1}) c_k. \tag{5.2.6}$$

Но $y_{11} \neq 0$ (см. (5.1.13)) и, следовательно, $\mathfrak{L}_1 (t_1; c_2, \ldots, c_n)$ вслед за $\mathfrak{L}_0$ является определенной однородной линейной функцией своих аргументов с коэффициентами λ_{k1}, не зависящими от $[c_k]$:

$$\mathfrak{L}_1 (t_1; c_2, \ldots, c_n) = \lambda_{11} t_1 + \sum_{k=2}^{n} \lambda_{k1} c_k. \tag{5.2.7}$$

5.3. Подставив значение $\mathfrak{L}_1$ из (5.2.7) в (5.2.2) (ср. с (5.2.1)), получим:

$$\frac{1}{\sqrt{2\pi\Delta^2}} \int e^{-u_2^2 : 2\Delta^2} du_2 \left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^{m-2} \int f\left(t_1; u_2 + \sum_{k=2}^{n} y_{k2} c_k, \ldots\right) \prod_{i=3}^{m} e^{-u_i^2 : 2\Delta^2} du_i = \\ = \mathfrak{L}_1 (t_1; c_2, \ldots, c_n). \tag{5.3.1}$$

Положим (ср. с (5.2.2))

$$\mathfrak{L}_2 = \left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^{m-2} \int f\left(t_1; u_2 + \sum_{k=2}^{n} y_{k2} c_k, \ldots\right) \prod_{i=3}^{m} e^{-u_i^2 : 2\Delta^2} du_i. \tag{5.3.2}$$

$\mathfrak{L}_2$ является функцией t_1; $u_2 + \sum_{k=2}^{n} y_{k2} c_k$; $c_3, \ldots, c_n$. Подставляя выражение (5.3.2) в (5.3.1), получаем (ср. с (5.2.3)):

$$\frac{1}{\sqrt{2\pi\Delta^2}} \int e^{-u_2^2 : 2\Delta^2} \mathfrak{L}_2 \left(t_1; u_2 + \sum_{k=2}^{n} y_{k2} c_k; c_3, \ldots, c_n\right) du_2 = \\ = \mathfrak{L}_1 (t_1; c_2, \ldots, c_n) \tag{5.3.3}$$

тождественно по всем t_1; $c_2, \ldots, c_n$, в частности тождественно по всевозможным значениям c_2. При фиксированных параметрах t_1; $c_3, \ldots, c_n$ этот интеграл является сверткой, и, по соображениям, изложенным относительно свертки (5.2.3), $\mathfrak{L}_2$ (с точностью до значений на множестве меры нуль) — линейная функция своего второго аргумента:

$$\mathfrak{L}_2 (t_1, t_2; c_3, \ldots, c_n) = t_2 A_2 + B_2, \tag{5.3.4}$$

где константы A_2, B_2 не зависят от t_2. Для вычисления констант подставляем (5.3.4) в (5.3.3) и после выкладок, аналогичных (5.2.4)—(5.2.7), находим, что $\mathfrak{L}_2$ является определенной однородной линейной функцией своих аргументов с коэффициентами, не зависящими от $[c_k]$.

5.4. Повторив n раз процесс, первые два шага которого проведены в п. п. 5.2 и 5.3, мы в результате получим (ср. с (5.2.1) и (5.3.1)):

$$\left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^{m-n}\int f(t_1, \ldots, t_n;\ u_{n+1}, \ldots, u_m)\prod_{i=n+1}^{m} e^{-u_i^2:2\Delta^2}\,du_i = \mathfrak{L}_n(t_1, \ldots, t_n), \quad (5.4.1)$$

где $\mathfrak{L}_n$ — определенная однородная линейная функция своих аргументов с коэффициентами, *не зависящими от* $[c_k]$.

Таким образом, в Θ входят только такие функции θ, для которых f

$$f(L[\eta_j]) = \theta([\eta_i]),$$

будучи подставлена в (5.4.1), дает определенную линейную однородную функцию $\mathfrak{L}_n$ (с точностью до значений на множестве меры нуль).

5.5. Рассмотрим дисперсию (ср. (5.1.1))

$$\mathbf{D}^2\theta = \left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^m \int \theta^2([\eta_i])\prod_{i=1}^{m} e^{-\Delta_i^2:2\Delta^2}\,d\Delta_i - (\mathbf{M}\theta)^2. \quad (5.5.1)$$

Нет нужды дополнительно предполагать конечность этого интеграла. Действительно, подынтегральное выражение в (5.5.1) неотрицательно, и интеграл Лебега, конечный или бесконечный, существует. Для функции θ_0 Гаусса дисперсия конечна. Поэтому при доказательстве того, что дисперсия любой другой функции $\theta \in \Theta$ больше дисперсии θ_0, можно, без ограничения общности, исключить из класса Θ функции с бесконечной дисперсией.

Преобразуем интеграл (5.5.1) так же, как мы преобразовали (5.1.1), т. е. перейдем к ортам $[\overline{Ou_i}]$ и функции f (см. (5.1.1) — (5.2.0)). Кроме того, заменим $\mathbf{M}\theta$ его значением (1.5); в результате получим:

$$\mathbf{D}^2\theta = \left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^m \int f^2\left(\left[u_i + \sum_{k=i}^{m} y_{ki}c_k\right]\right)\prod_{i=1}^{m} e^{-\Delta_i^2:2\Delta^2}\,d\Delta_i - \left(\sum_{k=1}^{n} x_{k0}c_k\right)^2. \quad (5.5.2)$$

При каждом наборе значений $[c_k]$ среди функций f, удовлетворяющих уравнению (5.4.1), нам надо найти функцию, дающую минимум выражению (5.5.2).

5.6. Покажем, что минимизирующей является функция $f_* = \mathfrak{L}_n$.

В самом деле, функция $f_* = \mathfrak{L}_n$, очевидно, удовлетворяет уравнению (5.4.1). Пусть теперь

$$+f(t_1, \ldots, t_n;\ u_{n+1}, \ldots, u_m) = \mathfrak{L}_n(t_1, \ldots, t_n) + \\ +\varphi(t_1, \ldots, t_n;\ u_{n+1}, \ldots, u_m)$$

удовлетворяет уравнению (5.4.1).

Тогда

$$\int \varphi(t_1, \ldots, t_n; u_{n+1}, \ldots, u_m) \prod_{i=n+1}^{m} e^{-u_i^2 : 2\Delta^2} du_i = 0 \qquad (5.6.1)$$

при всех $t_1, \ldots, t_n$. Далее,

$$f\left(\left[u_i + \sum_{k=i}^{n} y_{ki} c_k\right]_{i=1}^{i=m}\right) = \mathfrak{L}_n\left(\left[u_i + \sum_{k=i}^{n} y_{ki} c_k\right]_{i=1}^{i=n}\right) + \varphi\left(\left[u_i + \sum_{k=i}^{n} y_{ki} c_k\right]_{i=1}^{i=m}\right).$$

Подставляя это выражение для f в (5.5.2), получаем:

$$\left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^m \int (\mathfrak{L}_n + \varphi)^2 \prod_{i=1}^{m} e^{-u_i^2 : 2\Delta^2} du_i - \left(\sum_{k=1}^{n} x_{k0} c_k\right)^2 =$$

$$= \left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^m \int \mathfrak{L}_n^2 \prod_{i=1}^{m} e^{-u_i^2 : 2\Delta^2} du_i - \left(\sum_{k=1}^{n} x_{k0} c_k\right)^2 +$$

$$+ \left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^m \int \varphi^2 \prod_{i=1}^{m} e^{-u_i^2 : 2\Delta^2} du_i +$$

$$+ 2\left(\frac{1}{\sqrt{2\pi\Delta^2}}\right)^m \int \mathfrak{L}_n \prod_{i=1}^{n} e^{-u_i^2 : 2\Delta^2} du_i \int \varphi \prod_{i=n+1}^{m} e^{-u_i^2 : 2\Delta^2} du_i. \qquad (5.6.2)$$

В правой части (5.6.2) первые два члена представляют величину (5.5.2) для $f = \mathfrak{L}_n$. Третий член положителен, если φ отлична от нуля на множестве положительной меры. Последний член равен нулю, в силу (5.6.1). Утверждение 5.6, таким образом, доказано.

5.7. Возвращаясь к функции θ, получаем:

$$\theta_*([\eta_i]) = f_*(L[\eta_j]) = \mathfrak{L}_n(L[\eta_j]). \qquad (5.7.1)$$

Мы нашли функцию $\theta_* \in \Theta$, которая имеет минимальную дисперсию в классе Θ при неизвестных, но фиксированных $[c_k]$. Оказалось, что θ_* не зависит от коэффициентов $[c_k]$.

5.8. Мы доказали, что $\theta_*([\eta_i])$ является линейной однородной функцией своих аргументов. Поэтому, не уменьшая общности Теоремы, можно предполагать класс Θ состоящим только из линейных функций. В этом предположении Теорема была доказана Ф. Н. Давид и И. Нейманом [2]. Воспроизведем доказательство их теоремы. Статью [2] я достать не смог, поэтому нижеследующее доказательство может отличаться от первоначального. От доказательства этой же теоремы в [4] оно отличается заведомо.

5.9. Рассмотрим линейную однородную функцию

$$\theta([\eta_i]) = \sum_{i=1}^{m} a_i \eta_i. \qquad (5.9.1)$$

Так как

$$\eta_i = y_i + \Delta_i = \sum_{k=1}^{n} x_{ki} c_k + \Delta_i \text{ и } \mathbf{M}\Delta_i = 0, \tag{5.9.2}$$

то

$$\mathbf{M}\theta = \sum_{k=1}^{n} c_k \sum_{i=1}^{m} x_{ki} a_i. \tag{5.9.3}$$

Если $\mathbf{M}\theta$ равняется $\sum_{k=1}^{n} x_{k0} c_k$ при всевозможных $[c_k]$, то

$$\sum_{i=1}^{m} x_{ki} a_i = x_{k0} \quad (k = 1, \ldots, n). \tag{5.9.4}$$

Так как Δ_i независимы, $\mathbf{M}\Delta_i = 0$ и $\mathbf{D}^2\Delta_i = \Delta^2$, то

$$\mathbf{D}^2\theta = \Delta^2 \sum_{i=1}^{m} a_i. \tag{5.9.5}$$

Нам нужно минимизировать (5.9.5) при условиях (5.9.4). Будем искать по Лагранжу безусловный экстремум функции

$$F = \Delta^2 \sum_{i=1}^{m} a_i^2 - 2 \sum_{k=1}^{n} \lambda_k \left(\sum_{i=1}^{m} x_{ki} a_i - x_{k0} \right).$$

Это даст m уравнений $\frac{\partial F}{\partial a_i} = 0$, т. е.

$$a_i = \sum_{k=1}^{n} x_{ki} \lambda_k : \Delta^2 \quad (i = 1, \ldots, m). \tag{5.9.6}$$

Введем векторные обозначения:

$$\bar{a} = [a_i], \quad \bar{\lambda} = [\lambda_k], \quad \bar{x}_0 = [x_{k0}], \tag{5.9.7}$$

и матричное: $X = [x_{ki}]$. Здесь $k = 1, \ldots, n$; $i = 1, \ldots, m$; k — номер *строки*, i — номер *столбца*. Обозначим, как и раньше, через θ_* минимизирующую функцию θ, удовлетворяющую соотношениям (5.9.1), (5.9.4), (5.9.6). Перепишем в векторных обозначениях (5.9.1), (5.9.4) и (5.9.6):

$$\theta_* = (\bar{a}, \bar{\eta}), \quad X\bar{a} = \bar{x}_0, \quad \bar{a} = X^* \bar{\lambda} : \Delta^2, \tag{5.9.8}$$

где X^* — матрица, получающаяся из X заменой строк столбцами. Умножив последнее равенство на X, получаем:

$$\begin{gathered} X\bar{a} = XX^*\bar{\lambda} : \Delta^2, \quad \bar{x}_0 = XX^*\bar{\lambda} : \Delta^2, \\ \bar{\lambda} = (XX^*)^{-1}\bar{x}_0 \cdot \Delta^2, \quad \bar{a} = X^*(XX^*)^{-1}\bar{x}_0, \end{gathered} \tag{5.9.9}$$

где XX^* — квадратная симметричная невырожденная матрица, встречавшаяся в системе (2.3). Таким образом,

$$\theta_* = (X^*(XX^*)^{-1}\overline{x}_0, \overline{\eta}) = (\overline{x}_0, (XX^*)^{-1}X\overline{\eta}). \tag{5.9.10}$$

Введем еще вектор $\overline{\zeta} = [\zeta_k]$ $(k = 1, \ldots, n)$ и перепишем в векторных обозначениях равенства (2.1) и (2.2), определяющие функцию θ_0 Гаусса:

$$\theta_0 = (\overline{\zeta}, \overline{x}_0), \quad XX^*\overline{\zeta} = X\overline{\eta}. \tag{5.9.11}$$

Из этих уравнений получаем:

$$\overline{\zeta} = (XX^*)^{-1}X\overline{\eta}, \quad \theta_0 = \left((XX^*)^{-1}X\overline{\eta}, \overline{x}_0\right). \tag{5.9.12}$$

Сравнивая (5.9.10) и (5.9.12), находим: $\theta_0 = \theta_*$. Тем самым доказана теорема Ф. Н. Давид и И. Неймана [2] и закончено доказательство нашей теоремы.

§ 6. О свертке. *Пусть для почти всех y*

$$\frac{1}{\sqrt{2\pi\Delta^2}} \int_{-\infty}^{+\infty} f(y-\tau)\, e^{-\tau^2 : 2\Delta^2} d\tau = a + by, \tag{6.1}$$

где интеграл понимается в смысле Лебега. Тогда $f(y) = a + by$ для почти всех y.*

Сведение этого утверждения к известным теоремам об интеграле Фурье мы разобьем на пункты.

6.1. В равенстве (6.1) последовательно произведем подстановки

$$f(y) = a + by + \varphi(y), \quad y = x\Delta, \quad \tau = \tau_1\Delta,$$

$$x - \tau_1 = t, \quad \varphi(t\Delta)\exp(-t^2 : 2) = g(t).$$

При этом равенство (6.1) последовательно преобразуется в следующие:

$$\int_{-\infty}^{+\infty} \varphi(y-\tau)\, e^{-\tau^2 : 2\Delta^2} d\tau = 0, \qquad \int_{-\infty}^{+\infty} \varphi[(x-\tau_1)\Delta]\, e^{-\tau_1^2 : 2} d\tau_1 = 0,$$

$$\int_{-\infty}^{+\infty} \varphi(t\Delta)\, e^{-(x-t)^2 : 2} dt = 0, \qquad \int_{-\infty}^{+\infty} \varphi(t\Delta)\, e^{-t^2 : 2} e^{xt} e^{-x^2 : 2} dt = 0,$$

$$\int_{-\infty}^{+\infty} g(t)\, e^{xt} dt = 0.$$

Таким образом, утверждение § 6 примет следующий вид:

Если для почти всех x

$$\int_{-\infty}^{+\infty} g(t)\, e^{xt} dt = 0, \tag{6.1.1}$$

то $g(x) = 0$ почти всюду.

* Т. е. интеграл считается существующим, если конечен интеграл от абсолютного значения подынтегральной функции. Все встречающиеся в дальнейшем интегралы понимаются в этом же смысле.

6.2. Рассмотрим интеграл

$$\int_{-\infty}^{+\infty} (1+|t|)\,|g(t)|\cdot|e^{zt}|\,dt. \tag{6.2.1}$$

Покажем, что он равномерно сходится во всякой конечной области Q комплексного переменного $z=x+iy$. По условию (6.1) можно найти два числа a, b такие что:

$$a<0<b,$$

$$a+1<x<b-1 \text{ для всех } z=x+iy\in Q,$$

$$\int_{-\infty}^{+\infty} |g(t)|\,e^{xt}dt<\infty \text{ при } x=a \text{ и } x=b.$$

Введем функцию

$$r(t)=\begin{cases} |g(t)|\exp(at) & \text{при } t\leqslant 0,\\ |g(t)|\exp(bt) & \text{при } t>0.\end{cases}$$

Для $z\in Q$ функция $r(t)$ абсолютно мажорирует подынтегральную функцию (6.2.1) и

$$\int_{-\infty}^{+\infty} |r(t)|\,dt=\int_{-\infty}^{0} |g(t)|\,e^{at}dt+\int_{0}^{+\infty} |g(t)|\,e^{bt}dt<\infty.$$

Утверждение 6.2, таким образом, доказано.

6.3. В частности, показано, что интегралы Лебега

$$G(z)=\int_{-\infty}^{+\infty} g(t)\,e^{zt}dt \text{ и } \int_{-\infty}^{+\infty} g(t)\,te^{zt}dt$$

существуют и равномерно сходятся в любой конечной области комплексного переменного z. Но второй интеграл является формальной производной первого и условия для законности дифференцирования под знаком интеграла выполнены. Следовательно, $G(z)$ является аналитической функцией. По условию (6.1.1) $G(z)=0$ для почти всех точек действительной оси $z=x$. Следовательно, $G(z)\equiv 0$ для всех комплексных z. В частности,

$$\int_{-\infty}^{+\infty} |g(t)|\,dt<\infty \text{ и } \int_{-\infty}^{+\infty} g(t)\,e^{iyt}dt=0 \text{ для всех } y. \tag{6.3.1}$$

6.4. Согласно частному случаю теоремы 120 из [4] (стр. 217), из (6.3.1) следует, что $g(x)=0$ для почти всех действительных x, что и требовалось доказать.

Лаборатория управляющих машин АН СССР

(Поступило в редакцию 25/V 1956 г.)

Литература

1. А. Н. Колмогоров, К обоснованию метода наименьших квадратов, Успехи матем. наук, т. **I**, вып. **1** (**11**) (1946), 57—70.
2. F. N. David, J. Neyman, Extension of the Markoff theorem on least squares, Statistickal Research Memoirs, **II** (1938), 105—116.
3. В. В. Петров, О методе наименьших квадратов и его экстремальных свойствах, Успехи матем. наук, т. **IX**, вып. **1** (**59**) (1954), 41—62.
4. Е. Титчмарш, Введение в теорию интегралов Фурье, М.—Л., Гостехиздат, 1948.