EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 23. KÖIDE
FÜÜSIKA * MATEMAATIKA. 1974, NR. 2

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 23
ФИЗИКА * МАТЕМАТИКА. 1974, № 2

https://doi.org/10.3176/phys.math.1974.2.05

УДК 519.281

В. ОЛЬМАН

ОБ ОДНОМ ХАРАКТЕРИСТИЧЕСКОМ СВОЙСТВЕ НОРМАЛЬНОГО ЗАКОНА

Рассмотрим линейную регрессионную модель

$$\mathbf{Y}=\mathbf{X}\theta+\mathcal{E}, \tag{1}$$

где $\mathbf{Y}$ n-мерный случайный вектор, $\mathbf{X}$ — матрица порядка $n\times m$ известных коэффициентов, θ — фиксированный m-мерный параметр модели, $\mathcal{E}$ — n-мерный вектор из независимых одинаково распределенных компонент с нулевым средним и единичной дисперсией. Ради простоты изложения будем считать, что $\det(\mathbf{X}^T\mathbf{X})\neq 0$.

Основные результаты по характеризации нормального закона оптимальностью линейной оценки приведены в [1, 2], где все характеризации проводятся при некоторой фиксированной функции потери. В настоящей статье выводится характеристическое свойство нормального закона, заключающееся в минимаксности оценки наименьших квадратов (ОНК) в классе линейных оценок для некоторого широкого класса функций потери.

Итак, пусть в схеме (1) надо оценить неизвестный вектор θ по наблюденной реализации случайного вектора $\mathbf{Y}$. Ограничимся рассмотрением только линейных неоднородных оценок $\hat{\theta}$, т. е. оценок вида $\hat{\theta}=\mathbf{T}\mathbf{Y}+\mathbf{u}$, где $\mathbf{T}$ — неслучайная постоянная матрица порядка $m\times n$, а $\mathbf{u}$ — неслучайный постоянный m-мерный вектор. Пусть качество оценивания измеряется величиной

$$\sup_{\theta\in R^m} EL((\theta-\mathbf{T}\mathbf{Y}-\mathbf{u})^T\mathbf{D}(\theta-\mathbf{T}\mathbf{Y}-\mathbf{u})), \tag{2}$$

где $L(s)$ принадлежит классу $\mathfrak{L}$ неубывающих неотрицательных функций, определенных на положительной полуоси, для которых существуют хотя бы одна матрица $\mathbf{T}(L)$ и один вектор $\mathbf{u}(L)$ с конечным значением критерия (2); $\mathbf{D}$ — неотрицательно определенная матрица порядка m и ранга k, а R^m — m-мерное евклидово пространство. Тогда верна.

Теорема 1. *Если вектор $\mathcal{E}$ в схеме (1) распределен по нормальному закону, то критерий (2) для всех функций $L(s)\in\mathfrak{L}$ и любой матрицы $\mathbf{D}$ минимизируется по матрице $\mathbf{T}$ и вектору $\mathbf{u}$ при $\mathbf{T}_0=(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ и $\mathbf{u}_0=0$, т. е. на ОНК.*

Доказательство. Представим матрицу $\mathbf{D}$ в виде $\mathbf{D}=\mathbf{A}^T\mathbf{D}_0\mathbf{A}$, где $\mathbf{A}$ — ортогональная матрица, а $\mathbf{D}_0$ — диагональная матрица, главный минор которой $\mathbf{D}_0^k$ порядка k имеет ранг k. Обозначив $\mathbf{AT}=\mathbf{M}$,

$XA^T = N$, $v = Au$, $\eta = A\theta$ и сделав замену $\mathcal{E} = Y - X\theta$, преобразуем (2) к виду

$$\sup_{\eta \in R^m} EL((\eta - M\mathcal{E} - MN\eta - v)^T D_0 (\eta - M\mathcal{E} - MN\eta - v)). \qquad (3)$$

Разобьем матрицы M и N на две подматрицы следующим образом: $M^T = (M_1^T \vdots M_2^T)$, $N = (N_1 \vdots N_2)$, где M_1^T, N_1 — матрицы порядка $n \times k$, M_2^T, N_2 — матрицы порядка $n \times (m-k)$, а векторы η^T и v^T представим в виде

$$\eta^T = (\eta_1^T \vdots \eta_2^T), \qquad v^T = (v_1^T \vdots v_2^T),$$

где η_1, v_1 — k-мерные, а η_2, v_2 — $(m-k)$-мерные векторы. Использовав блочное представление, перепишем (3) в виде

$$\sup_{\eta_1 \in R^k,\, \eta_2 \in R^{m-k}} EL((\eta_1 - M_1\mathcal{E} - M_1N_1\eta_1 - v_1 - M_1N_2\eta_2)^T D_0^k \times$$
$$\times (\eta_1 - M_1\mathcal{E} - M_1N_1\eta_1 - M_1N_2\eta_2 - v_1)). \qquad (4)$$

Очевидно, матрица M_1, минимизирующая (4), в силу неограниченности векторов η_1 и η_2 должна удовлетворять условиям:

$$M_1N_1 = I_k, \qquad M_1N_2 = O_{k,n-k}, \qquad (5)$$

где I_k — единичная матрица порядка k, а $O_{k,n-k}$ нулевая матрица порядка $k \times (n-k)$. Согласно [3] (теорема 1), минимум критерия (4) по вектору v_1 реализуется при $v_1 = 0$. Таким образом, задача сводится к минимизации интеграла

$$\int_{R^n} L(e^T M_1^T D_0^k M_1 e) \exp(-e^T e/2)\, de \qquad (6)$$

по матрице M_1, удовлетворяющей условиям (5). Но матрица M_1 представима в виде

$$M_1 = S_1N_1^T + S_2N_2^T + S_3N_3^T, \qquad (7)$$

где S_1, S_2, S_3 — матрицы соответственно порядков $k \times k$, $k \times (m-k)$ и $k \times (n-m)$, а столбцы матрицы N_3 образуют ортогональное дополнение к линейному многообразию, порожденному столбцами матрицы N. Интеграл (6) зависит только от ненулевых собственных чисел матрицы $M_1^T D_0^k M_1$, которые совпадают с собственными числами матрицы $(D_0^k)^{1/2} M_1 M_1^T (D_0^k)^{1/2}$. Но так как

$$(D_0^k)^{1/2} M_1 M_1^T (D_0^k)^{1/2} =$$
$$= (D_0^k)^{1/2} ((S_1N_1^T + S_2N_2^T)(S_1N_1^T + S_2N_2^T)^T + S_3N_3^T N_3 S_3)(D_0^k)^{1/2},$$

то используя минимакс-теорему Фишера [4], получаем, что (6) минимизируется по S_3 при $S_3 = O_{k,n-m}$. Из условий (5) и представления (7) получаем два матричных уравнения относительно неизвестных матриц S_1 и S_2:

$$S_1N_1^T N_1 + S_2N_2^T N_1 = I_k, \qquad S_1N_1^T N_2 + S_2N_2^T N_2 = O_{k,m-k}.$$

Нетрудно показать, что единственным решением системы матричных уравнений будут матрицы:

$$S_{10}=(N_1^T N_1 - N_1^T N_2(N_2^T N_2)^{-1} N_2^T N_1)^{-1},$$

$$S_{20}=-(N_1^T N_1 - N_1^T N_2(N_2^T N_2)^{-1} N_2^T N_1)^{-1} N_1^T N_2(N_2^T N_2)^{-1}.$$

Воспользовавшись методами вычисления обратной матрицы для блочной и для суммы двух матриц ([2], с. 44), можно показать, что матрица $S_{10}N_1^T + S_{20}N_2^T$ представляет собой верхние k строк матрицы $\mathbf{M}_0 = (\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T$. Следовательно, критерий (3) минимизируется на $\mathbf{M}_0$, а (2) — на $\mathbf{A}^T\mathbf{M}_0 = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$, что и требовалось доказать.

Будем говорить, что плотность распределения вектора $\mathcal{E}$ в схеме (1) обладает свойством (A), если ОНК минимизирует критерий (2) для всех функций потери $L(s) \in \mathfrak{L}$ и любых матриц $\mathbf{X}$ и $\mathbf{D}$.

Теорема 2. *В классе дифференцируемых ограниченных плотностей распределения вектора* $\mathcal{E}$ *в схеме* (1) *только нормальная плотность обладает свойством* (A).

Доказательство. Обозначим плотность распределения компоненты вектора $\mathcal{E}$ через f_ε и примем $\mathbf{D} = \mathbf{I}_m$. По свойству (A)

$$\min_{\mathbf{T},\mathbf{u}} \max_{\theta\in R^m} EL((\theta-\mathbf{TY}-\mathbf{u})^T(\theta-\mathbf{TY}-\mathbf{u})) = EL(\mathcal{E}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-2}\mathbf{X}^T\mathcal{E}). \quad (8)$$

С помощью последнего равенства доказывается

Лемма. *Всякая плотность* f_ε, *обладающая свойством* (A), *симметрична.*

Доказательство. Взяв в качестве $\mathbf{X}^T$ матрицу $(\mathbf{H} \,\vdots\, \mathbf{O}_{m,n-m})$, где $\mathbf{H}$ — диагональная матрица порядка m с единицей в левом верхнем углу и с остальными диагональными элементами h, и обозначив компоненты вектора $\mathcal{E}$ через $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, получим

$$EL((\varepsilon_1+u)^2+h^2\sum_{i=2}^{m}\varepsilon_i^2) \geqslant EL(\varepsilon_1^2+h^2\sum_{i=2}^{m}\varepsilon_i^2)\ \forall\, u\in R^1. \quad (9)$$

Рассмотрим неравенство (9) на функциях

$$L_c(s)=\begin{cases}0, & |s|\leqslant c,\\ 1, & |s|>c,\end{cases}$$

где c — любое положительное число. Продифференцировав левую часть неравенства (9) по u в точке $u=0$, получим

$$\int\limits_{\varepsilon_1^2+h^2\sum_{i=2}^{m}\varepsilon_i^2\geqslant c^2} f'_\varepsilon(\varepsilon_1)f_\varepsilon(\varepsilon_2)\ldots f_\varepsilon(\varepsilon_n)\,d\mathbf{e}=0\ \forall\, h.$$

Дифференцирование под знаком интеграла (как и все далее встречающиеся дифференцирования) возможно благодаря ограниченности, дифференцируемости и существованию конечного второго момента плотности f_ε. Переходя к пределу в последнем равенстве по $h\to 0$, получаем $\int\limits_{\varepsilon_1^2\geqslant c^2} f'_\varepsilon(\varepsilon_1)\,d\varepsilon_1=0$, откуда $f_\varepsilon(-c)=f_\varepsilon(c)$. Лемма доказана.

Матрица $\mathbf{T}$, на которой реализуется равенство (8), представима в виде $\mathbf{T} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T + \mathbf{S}\mathbf{M}^T$, где $\mathbf{S}$ — матрица порядка $m\times(n-m)$, а $\mathbf{M}$ — матрица полного ранга порядка $n\times(n-m)$ и $\mathbf{M}^T\mathbf{X} = \mathbf{O}_{n-m,m}$. Обозначим строки матрицы $\mathbf{S}$ через $\mathbf{s}_1^T, \mathbf{s}_2^T, \ldots, \mathbf{s}_m^T$, а строки матрицы $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ через $\mathbf{x}_1^T, \mathbf{x}_2^T, \ldots, \mathbf{x}_m^T$. Тогда

$$\mathscr{E}^{T}T^{T}T\mathscr{E}=\mathscr{E}^{T}X(X^{T}X)^{-2}X^{T}\mathscr{E}+\sum_{i=1}^{m}(s_i^{T}M^{T}\mathscr{E})^2+2\sum_{i=1}^{m}(\mathscr{E}^{T}x_i)(s_i^{T}M^{T}\mathscr{E}). \quad (10)$$

Подставляя (10) в равенство (8) и дифференцируя $EL(\mathscr{E}^{T}T^{T}T\mathscr{E})$ по векторам $s_1, s_2, \ldots, s_m$, получаем, что для выполнения условия (A) необходимо

$$\int_{R^n} L'(e^{T}X(X^{T}X)^{-2}X^{T}e)\,e^{T}x_iM^{T}e\prod_{j=1}^{n}f_\varepsilon(\varepsilon_j)\,de=0,\quad i=1, 2, \ldots, m,$$

где $L'(s)$ — произвольная неотрицательная функция, т. е.

$$M^{T}\int_{R^n} L'(e^{T}X(X^{T}X)^{-2}X^{T}e)\,ee^{T}\prod_{j=1}^{n}f_\varepsilon(\varepsilon_i)\;deX(X^{T}X)^{-1}=O_{n-m,m}. \quad (11)$$

Можно показать ([4], с. 71), что существуют ортогональные матрицы $\mathbf{A}$ и $\mathbf{B}$ соответственно порядков m и n такие, что $X^{T}=A(X_0 \mid O_{m,n-m})B$, где X_0 — диагональная матрица с положительными диагональными элементами $x_1^{-1}, x_2^{-1}, \ldots, x_m^{-1}$. Так как $M^{T}X=O_{n-m,m}$ и матрица $M^{T}M$ невырожденна, то из (11) нетрудно получить

$$\int_{R^n} L'\Big(\sum_{i=1}^{m}\varepsilon_i^2x_i^2\Big)\varepsilon_i\varepsilon_j\prod_{r=1}^{n}f_\varepsilon(b_r^{T}e)\,de=0,\quad 1\leqslant i\leqslant m,\quad m+1\leqslant j\leqslant n, \quad (12)$$

при любом наборе векторов $b_1, b_2, \ldots, b_n$, составляющих ортогональную матрицу. Отсюда следует

$$\int_{R^2} L'(\varepsilon_1^2x_1^2)\,\varepsilon_1\varepsilon_2f_\varepsilon(\varepsilon_1\cos\alpha-\varepsilon_2\sin\alpha)\times$$
$$\times f_\varepsilon(\varepsilon_1\sin\alpha+\varepsilon_2\cos\alpha)\,d\varepsilon_1\,d\varepsilon_2=0\ \ \forall\alpha,\ x_1>0, \quad (13)$$

(здесь α — параметр ортогональной матрицы второго порядка). Рассмотрим функциональное уравнение (13) на функциях $L'_c(s)=1-L_c(sx_1^2)$. Подставляя $L'_c(s)$ в (13) и дифференцируя обе части по c, получаем

$$\int_{-\infty}^{\infty}\varepsilon_2f_\varepsilon(c\cos\alpha-\varepsilon_2\sin\alpha)f_\varepsilon(c\sin\alpha+\varepsilon_2\cos\alpha)\,d\varepsilon_2+$$
$$+\int_{-\infty}^{\infty}\varepsilon_2f_\varepsilon(-c\cos\alpha-\varepsilon_2\sin\alpha)f_\varepsilon(-c\sin\alpha+\varepsilon_2\cos\alpha)\,d\varepsilon_2=0\ \ \forall c,\ \alpha,$$

а в силу четности плотности f_ε, доказанной в лемме,

$$\int_{-\infty}^{\infty}\varepsilon_2f_\varepsilon(s-t\varepsilon_2)f_\varepsilon(t+\varepsilon_2s)\,d\varepsilon_2=0\ \ \forall s,\ t. \quad (14)$$

Дифференцируя уравнение (14) по s в точке $s=0$, получаем

$$\int_{-\infty}^{\infty}\varepsilon_2f'_\varepsilon(-\varepsilon_2t)f_\varepsilon(t)\,d\varepsilon_2+\int_{-\infty}^{\infty}\varepsilon_2^2f_\varepsilon(-t\varepsilon_2)f'_\varepsilon(t)\,d\varepsilon_2=0.$$

Так как $\int_{-\infty}^{\infty}\varepsilon_2f'_\varepsilon(\varepsilon_2)\,d\varepsilon_2=-1$, а $\int_{-\infty}^{\infty}\varepsilon_2^2f_\varepsilon(\varepsilon_2)\,d\varepsilon_2=1$ по условию теоремы, то приходим к дифференциальному уравнению относительно плотности $f_\varepsilon(t)$

$$f'_\varepsilon(t) = -tf_\varepsilon(t),$$

единственным решением которого будет

$$f_\varepsilon(t) = 1/\sqrt{2\pi}\exp(-t^2/2).$$

ЛИТЕРАТУРА

1. Каган А. М., Линник Ю. В., Рао С. Р., Характеризационные задачи математической статистики, М., 1972.
2. Рао С. Р., Линейные статистические методы и их применения, М., 1968.
3. Ольман В., Изв. АН ЭССР, Физ. Матем., 23, 127 (1974).
4. Fischer E., Monatsh. Math. Physik, 16, 234—249 (1905).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
4/IX 1973

V. OLMAN

ÜHEST NORMAALJAOTUSE KARAKTERISTLIKUST OMADUSEST

Näidatakse, et lineaarse regressioonimudeli parameetri vähimruutude hinnang on lineaarsete hinnangute klassis suvaliste plaanimaatriksite ja suvaliste, ainult hinnangu ja parameetri vahelisest kaugusest sõltuvate kaofunktsioonide korral minimakshinnang siis ja ainult siis, kui sõltumatute vaatluste vektor allub normaaljaotusele.

V. OLMAN

ON THE CHARACTERISTIC PROPERTY OF THE NORMAL LAW

It is shown that the least squares estimator of a linear regression model parameter is minimax in a class of linear estimators under any plan matrices and any non-decreasing loss functions which depend only on the distance between an estimator and parameter if and only if the vector of independent observations has a normal distribution.