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.1974.2.03

УДК 519.281

В. ОЛЬМАН

МИНИМАКСНОЕ ДОВЕРИТЕЛЬНОЕ ОЦЕНИВАНИЕ ПАРАМЕТРА РЕГРЕССИОННОЙ МОДЕЛИ ПРИ НОРМАЛЬНОЙ АДДИТИВНОЙ ПОМЕХЕ

Пусть задана регрессионная модель

$$\mathbf{y}=\mathbf{X}\theta+\mathbf{e}, \tag{1}$$

где $\mathbf{y}$ — реализация случайного n-мерного вектора $\mathbf{Y}$, $\mathbf{X}$ — матрица порядка $n\times m$ известных коэффициентов, $\mathbf{e}$ — реализация n-мерного гауссовского вектора $\mathcal{E}$ с нулевым вектором средних и единичной дисперсионной матрицей $\mathbf{I}_n$, θ — m-мерный вектор-параметр модели. Для простоты изложения будем считать, что $\det(\mathbf{X}^T\mathbf{X})\neq 0$.

Под доверительной процедурой $C(\cdot|\cdot)$, следуя [1], будем подразумевать подмножество декартова произведения выборочного и параметрического пространств:

$$\{(\theta\in\Omega^m,\ \mathbf{y}\in R^n):(\theta,\mathbf{y})\in C(\cdot/\cdot)\},$$

где R^n — n-мерное евклидово пространство, а Ω^m — фиксированное подмножество m-мерного евклидова пространства. За качество доверительного оценивания примем величину

$$\min_{\theta\in\Omega^m} P\{\theta\in C(\cdot/\mathbf{Y})\}, \tag{2}$$

где $P\{\cdot\}$ — вероятность события, порожденная распределением вектора $\mathbf{Y}(\mathbf{Y}=\mathbf{X}\theta+\mathcal{E})$ Минимаксной будем называть доверительную процедуру, максимизирующую критерий (2).

Построению минимаксных доверительных процедур для оценивания среднего многомерного нормального закона посвящено много работ. В качестве результатов, имеющих отношение к данной статье, можно выделить следующие:

1. Среди доверительных процедур фиксированной m-мерной лебеговской меры при каждом $\mathbf{y}\in R^n$, инвариантных относительно сдвига и вращения [2], шар с центром в оценке наименьших квадратов (ОНК) представляет собой единственную процедуру, максимизирующую критерий (2) при $\Omega^m=R^m$.
2. Среди всех доверительных процедур шар с центром в ОНК не является допустимой процедурой [3, 4], хотя на этой процедуре и реализуется максимум критерия (2) при $\Omega^m=R^m$.

В настоящей статье в качестве доверительных процедур будут рассматриваться центрированные линейной по наблюдениям статистикой множества

$$C_{\mathbf{D,T,u}}(\cdot/\cdot)=\{(\theta,\mathbf{y}):(\theta-\mathbf{Ty}-\mathbf{u})^{\mathrm{T}}\mathbf{D}(\theta-\mathbf{Ty}-\mathbf{u})\leqslant c^2\}, \qquad (3)$$

где $\mathbf{T}$ — произвольная неслучайная матрица порядка $m\times n$, $\mathbf{u}$ — произвольный неслучайный m-мерный вектор, $\mathbf{D}$ — некоторая неотрицательно определенная матрица порядка $m\times m$, а c^2 — фиксированная положительная постоянная. Описанные доверительные процедуры преобразуют критерий (2) к виду

$$\min_{\theta\in\Omega^m} P\{(\theta-\mathbf{TY}-\mathbf{u})^{\mathrm{T}}\mathbf{D}(\theta-\mathbf{TY}-\mathbf{u})\leqslant c^2\}. \qquad (4)$$

Задача заключается в максимизации критерия (4) по матрицам $\mathbf{T}$ и $\mathbf{D}$ и вектору $\mathbf{u}$ при условии, что m-мерный лебеговский объем $\mu(\cdot)$ множества $C_{\mathbf{D,T,u}}(\cdot/\mathbf{y})\,\forall\,\mathbf{y}\in R^n$ не превосходит фиксированной величины μ_c. Показывается, что в описанной задаче шар с центром в ОНК — единственная оптимальная процедура при $\Omega^m=R^m$. Приводятся частные решения задачи в случае, когда

$$\Omega^m=\{\theta\in R^m:(\theta-\theta_0)^{\mathrm{T}}\mathbf{A}(\theta-\theta_0)\leqslant 1\}, \qquad (5)$$

где θ_0 — известный m-мерный вектор и $\mathbf{A}$ — неотрицательно определенная матрица.

Так как вектор $\mathcal{E}$ распределен по нормальному закону, то согласно [5] задача сводится к нахождению

$$\max_{\mu(C_{\mathbf{D,T,u}}(\cdot/\mathbf{y}))\leqslant\mu_c} P\{\mathcal{E}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{DT}\mathcal{E}\} \qquad (6)$$

при $\mathbf{u}=0$ и $\mathbf{T}=(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$. Так как $\mu(C_{\mathbf{D,T,u}}(\cdot/\mathbf{y}))=\gamma_{c,m}\det(\mathbf{D})$, где

$$\gamma_{c,m}=\frac{\Gamma^m\left(\frac{1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)c^{2m}}\int_0^1 \exp(-u)\,u^{\frac{m-1}{2}}\,du$$ [6], очевидно, что (6) мажорируется величиной $\max\limits_{\mathbf{M}} P\{\mathcal{E}^{\mathrm{T}}\mathbf{M}\mathcal{E}\leqslant c^2\}$ при ограничении

$$\det(\mathbf{M})\leqslant\det(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mu_c/\gamma_{c,m}. \qquad (7)$$

Обозначим через $I_c(\mathbf{M})$ m-мерный эллипсоид, порожденный неотрицательно определенной матрицей $\mathbf{M}$: $I_c(\mathbf{M})=\{\mathbf{e}\in R^m:\mathbf{e}^{\mathrm{T}}\mathbf{Me}\leqslant c^2\}$. Пусть $\mathbf{M}_0$ единичная с точностью до множителя матрица, удовлетворяющая ограничению (7), т. е. $I_c(\mathbf{M}_0)$ — шар в R^m. Очевидно, лебеговские объемы множеств $S_1(\mathbf{M})=I_c(\mathbf{M}_0)\backslash I_c(\mathbf{M})$, $S_2(\mathbf{M})=I_c(\mathbf{M})\backslash I_c(\mathbf{M}_0)$ совпадают и $P\{S_1(\mathbf{M})\}\geqslant P\{S_2(\mathbf{M})\}$, следовательно,

$$\max_{\det(\mathbf{M})\leqslant\det(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mu_c/\gamma_{c,m}} P\{\mathcal{E}^{\mathrm{T}}\mathbf{M}\mathcal{E}\leqslant c^2\}=P\{\mathcal{E}^{\mathrm{T}}\mathbf{M}_0\mathcal{E}\leqslant c^2\}. \qquad (8)$$

Таким образом, критерий (6) реализуется при

$$\mathbf{D}_0=(\mathbf{X}^{\mathrm{T}}\mathbf{X})\left[\frac{\mu_c}{\det(\mathbf{X}^{\mathrm{T}}\mathbf{X})}\frac{1}{\gamma_{c,m}}\right]^{1/m}.$$

Доказанный результат формулирует

Теорема 1. *Среди доверительных процедур вида* (3) *ограниченного лебеговского объема для* $\forall\mathbf{y}\in R^n$ *эллипсоид с центром в ОНК и* *матрицей* $\mathbf{D}_0$ *представляет собой единственную минимаксную по критерию* (2) *доверительную процедуру.*

Замечания. 1. Если в качестве доверительных эллипсоидов рассматриваются эллипсоиды, которым соответствуют вырожденные матрицы $\mathbf{D}$ ранга $k < m$, причем k-мерный объем проекции этих эллипсоидов в R^k не превосходит μ_c, то $\mathbf{D}_0 = \left(\dfrac{\mu_c}{\prod\limits_{i=1}^{k} x_i^2} \dfrac{1}{\gamma_{c,k}}\right)^{1/k} \sum\limits_{i=1}^{k} x_i^2 \mathbf{a}_i \mathbf{a}_i^{\mathrm{T}}$ при условии, что $x_1^2 \geqslant x_2^2 \geqslant \cdots \geqslant x_m^2$ — собственные числа матрицы $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ и $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ — соответствующие им собственные векторы.

2. Из равенства (8) видно, что обычный $\mathbf{D}$-оптимальный план оптимален и в доверительном оценивании.

3. Докажем более общий факт: пусть в качестве доверительных процедур рассматриваются множества $C(\cdot/\mathbf{y})$ некоторой фиксированной лебеговской меры $\mu(C(\cdot/\mathbf{y})) = \mu < \infty$ и пусть положение их в пространстве параметра $\boldsymbol{\theta}$ фиксируется линейным преобразованием вектора наблюдений, т. е. $C(\cdot/\mathbf{y}) = C(\cdot/\mathbf{Ty}+\mathbf{u})$, причем доверительная процедура $C(\cdot/\cdot)$ инвариантна относительно сдвига $C(\cdot/\mathbf{Ty}+\mathbf{u}) = \mathbf{u} + C(\cdot/\mathbf{Ty})\ \forall \mathbf{u}$. Пусть также каждому множеству $C(\cdot/\mathbf{Ty}+\mathbf{u})$ соответствует вещественное число $K_c < \infty$ такое, что

$$(\boldsymbol{\theta} - \mathbf{Ty})^{\mathrm{T}} (\boldsymbol{\theta} - \mathbf{Ty}) \leqslant K_c \qquad \forall\, \boldsymbol{\theta} \in C(\cdot/\mathbf{Ty}). \tag{9}$$

Тогда эллипсоид с центром в ОНК и матрицей $\mathbf{D}_0$ будет единственным оптимальным доверительным множеством в смысле критерия (2).

Доказательство. Как видно, условие (9) накладывает ограничения на оптимальную матрицу $\mathbf{T} : \mathbf{TX} = \mathbf{I}_m$, следовательно:

$$\max_{\mathbf{T},\mathbf{u},C} \min_{\boldsymbol{\theta} \in R^m} P\{\boldsymbol{\theta} \in C(\cdot/\mathbf{TY}+\mathbf{u})\} =$$

$$= \max_{\mathbf{u},\ \mathbf{TX}=\mathbf{I}_m,\ C} \frac{1}{(2\pi)^{n/2}} \int\limits_{\{\mathbf{e} \in R^n\ :\ \mathbf{Te} \in C(\mathbf{u}/\cdot)\}} \exp(-\mathbf{e}^{\mathrm{T}}\mathbf{e}/2)\, d\mathbf{e}.$$

Нетрудно убедиться, что $\mathbf{T} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$, а далее аналогично рассуждениям в теореме убеждаемся в справедливости замечания.

Пусть теперь априори известно, что m-мерный параметр модели (1) принадлежит эллипсоиду $I_1(\mathbf{A}) = \{\boldsymbol{\theta} \in R^m : \boldsymbol{\theta}^{\mathrm{T}}\mathbf{A}\boldsymbol{\theta} \leqslant 1, \det(\mathbf{A}) \neq 0\}$, а матрица $\mathbf{D} = \mathbf{dd}^{\mathrm{T}}$, где $\mathbf{d}$ — известный m-мерный вектор. Задача заключается в максимизации по матрице $\mathbf{T}$ и вектору $\mathbf{u}$ критерия

$$\min_{\boldsymbol{\theta}^{\mathrm{T}}\mathbf{A}\boldsymbol{\theta} \leqslant 1} P\{(\mathbf{d}^{\mathrm{T}}\boldsymbol{\theta} - \mathbf{d}^{\mathrm{T}}\mathbf{Ty} - \mathbf{d}^{\mathrm{T}}\mathbf{u})^2 \leqslant c^2\}. \tag{10}$$

Обозначим вектор $\mathbf{T}^{\mathrm{T}}\mathbf{d}$ через $\mathbf{s}$, а $P\{(\mathbf{d}^{\mathrm{T}}\boldsymbol{\theta} - \mathbf{s}^{\mathrm{T}}\mathbf{Y} - \mathbf{d}^{\mathrm{T}}\mathbf{u})^2 \leqslant C^2\}$ через $B(\mathbf{s}, \boldsymbol{\theta})$. Нетрудно получить

$$B(\mathbf{s}, \boldsymbol{\theta}) = \frac{1}{(2\pi)^{n/2}} \int\limits_{(\mathbf{s}^{\mathrm{T}}\mathbf{e}+\mathbf{d}^{\mathrm{T}}\mathbf{u}+\mathbf{s}^{\mathrm{T}}\mathbf{X}\boldsymbol{\theta}-\mathbf{d}^{\mathrm{T}}\boldsymbol{\theta})^2 \leqslant c^2} \exp(-\mathbf{e}^{\mathrm{T}}\mathbf{e}/2)\, d\mathbf{e}.$$

Произведем ортогональное преобразование над вектором $\mathbf{e}$ такое, при котором первая строка преобразующей матрицы равна $\mathbf{s}^{\mathrm{T}} \dfrac{1}{\sqrt{\mathbf{s}^{\mathrm{T}}\mathbf{s}}}$, тогда

$$B(\mathbf{s},\theta)=\frac{1}{\sqrt{2\pi}}\int\limits_{(y\sqrt{\mathbf{s}^{\mathrm{T}}\mathbf{s}}+\mathbf{d}^{\mathrm{T}}\mathbf{u}+\mathbf{s}^{\mathrm{T}}\mathbf{X}\theta-\mathbf{d}^{\mathrm{T}}\theta)^2\leqslant c^2}\exp(-y^2/2)\,dy.$$

Очевидно, $\min\limits_{\theta^{\mathrm{T}}\mathbf{A}\theta\leqslant 1} B(\mathbf{s},\theta)$ достигается на векторе $\theta_{\mathbf{s}}$, реализующем $\max\limits_{\theta^{\mathrm{T}}\mathbf{A}\theta\leqslant 1}[\mathbf{d}^{\mathrm{T}}\mathbf{u}+(\mathbf{s}^{\mathrm{T}}\mathbf{X}-\mathbf{d}^{\mathrm{T}})\theta]^2$, откуда легко следует $\mathbf{u}=0$, а по неравенству Коши—Шварца

$$g^2(\mathbf{s})=\max_{\theta^{\mathrm{T}}\mathbf{A}\theta\leqslant 1}[(\mathbf{s}^{\mathrm{T}}\mathbf{X}-\mathbf{d}^{\mathrm{T}})\theta]^2=(\mathbf{s}^{\mathrm{T}}\mathbf{X}-\mathbf{d}^{\mathrm{T}})\mathbf{A}^{-1}(\mathbf{X}^{\mathrm{T}}\mathbf{s}-\mathbf{d}).$$

Следовательно,

$$\max_{\mathbf{s},\mathbf{u}}\min_{\theta^{\mathrm{T}}\mathbf{A}\theta\leqslant 1}B(\mathbf{s},\theta)=\max_{\mathbf{s}}\frac{1}{\sqrt{2\pi}}\int\limits_{-c+g(\mathbf{s})\leqslant y\sqrt{\mathbf{s}^{\mathrm{T}}\mathbf{s}}<c+g(\mathbf{s})}\exp(-y^2/2)\,dy.$$

Легко убедиться, что

$$\max_{\mathbf{s}}\int\limits_{-c+g(\mathbf{s})\leqslant y\sqrt{\mathbf{s}^{\mathrm{T}}\mathbf{s}}<c+g(\mathbf{s})}\cdot\ =$$

$$=\max_{\lambda}\max_{\mathbf{s}:\,\mathbf{s}^{\mathrm{T}}\mathbf{s}=d(\lambda)}\int\limits_{-c+g(\mathbf{s})\leqslant yd(\lambda)\leqslant c+g(\mathbf{s})}\exp(-y^2/2)\,dy,\tag{11}$$

где $d^2(\lambda)=\mathbf{d}^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X}+\lambda\mathbf{A})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X}+\lambda\mathbf{A})^{-1}\mathbf{d}$. Внутренний максимум выражения (11) достигается на векторе s_λ, который реализует $\min\limits_{\mathbf{s}^{\mathrm{T}}\mathbf{s}=d(\lambda)} g(\mathbf{s})$.

Применив метод множителей Лагранжа и использовав равенство $(\mathbf{X}\mathbf{A}^{-1}\mathbf{X}^{\mathrm{T}}+\lambda\mathbf{I}_n)^{-1}\mathbf{X}\mathbf{A}^{-1}=\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X}+\lambda\mathbf{A})^{-1}$ [7], получим

$$\mathbf{s}_\lambda=\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X}+\lambda\mathbf{A})^{-1}\mathbf{d}$$

и

$$g^2(\mathbf{s}_\lambda)=\lambda^2\mathbf{d}^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X}+\lambda\mathbf{A})^{-1}\mathbf{A}(\mathbf{X}^{\mathrm{T}}\mathbf{X}+\lambda\mathbf{A})^{-1}\mathbf{d}.$$

Пусть $\mathbf{L}$ — невырожденная квадратная матрица такая, что $\mathbf{X}^{\mathrm{T}}\mathbf{X}=\mathbf{L}^{\mathrm{T}}\mathbf{X}_0\mathbf{L}$ и $\mathbf{A}=\mathbf{L}^{\mathrm{T}}\mathbf{A}_0\mathbf{L}$, где $\mathbf{X}_0$ и $\mathbf{A}_0$ — диагональные матрицы с элементами $x_{01}^2, x_{02}^2, \ldots, x_{0m}^2$ и $a_{01}^2, a_{02}^2, \ldots, a_{0m}^2$. Обозначим вектор $(\mathbf{L}^{\mathrm{T}})^{-1}\mathbf{d}$ через $\mathbf{d}_0^{\mathrm{T}}=(d_{01}, d_{02}, \ldots, d_{0m})$, тогда

$$d^2(\lambda)=\sum_{i=1}^{m}d_{i0}^2\frac{x_{i0}^2}{(x_{i0}^2+\lambda a_{i0}^2)^2}$$

и

$$g^2(\mathbf{s}_\lambda)=\lambda^2\sum_{i=1}^{m}d_{i0}^2\frac{a_{i0}^2}{(x_{i0}^2+\lambda a_{i0}^2)^2}.$$

Согласно теореме Ньютона—Лейбница из (11) получаем, что оптимальный параметр λ должен удовлетворять уравнению

$$\exp\left(-\frac{[c+g(\mathbf{s}_\lambda)]^2}{2d^2(\lambda)}\right)\frac{g'(\mathbf{s}_\lambda)d(\lambda)-d'(\lambda)[c+g(\mathbf{s}_\lambda)]}{d^2(\lambda)}=$$

$$=\exp\left(-\frac{[-c+g(s_\lambda)]^2}{2d^2(\lambda)}\right)\frac{g'(s_\lambda)d(\lambda)-d'(\lambda)[-c+g(s_\lambda)]}{d^2(\lambda)} \tag{12}$$

или

$$\exp\left(-\frac{cg(s_\lambda)}{d^2(\lambda)}\right)=-1+\frac{2[g'(s_\lambda)d(\lambda)-d'(\lambda)g(s_\lambda)]}{g'(s_\lambda)d(\lambda)-d'(\lambda)g(s_\lambda)-cd'(\lambda)}. \tag{13}$$

Решение уравнения (13) относительно λ весьма затруднительно, поэтому имеет смысл решать его относительно c при фиксированном λ. Весьма простыми методами можно показать, что уравнение (13) относительно c имеет, кроме $c=0$, еще один корень c_λ. Таким образом, доказана

Теорема 2. *Критерий* (8) *при* $c=c_\lambda$ *максимизируется при* $\mathbf{u}=0$ *и* $\mathbf{T}=\mathbf{T}_{0\lambda}=(\mathbf{X}^T\mathbf{X}+\lambda\mathbf{A})^{-1}$.

Замечания. 1. Построение оптимального доверительного интервала при вырожденных матрицах $\mathbf{A}$ и $\mathbf{X}^T\mathbf{X}$ не содержит принципиальной трудности и может быть проведено с помощью методов, использованных в [7].

2. Если $\mathbf{X}^T\mathbf{X}=\alpha^2\mathbf{A}$ $(\alpha^2>0)$, то уравнение (13) превращается в квадратное относительно λ

$$\exp\left(-\frac{c\lambda(\alpha^2+\lambda)}{t\alpha^2}\right)=\frac{t-c}{t+c}, \quad \text{где } t^2=\sum_{i=1}^m\frac{d_{i0}^2}{a_{i0}^2}.$$

3. Если априорное ограничение на параметр θ задано в виде $(\theta-\theta_0)^T\mathbf{A}(\theta-\theta_0)\leqslant 1\,(\det\mathbf{A}\neq 0)$, где θ_0 — фиксированный m-мерный вектор, то максимум по матрице $\mathbf{T}$ и вектору $\mathbf{u}$ критерия

$$\min_{(\theta-\theta_0)^T\mathbf{A}(\theta-\theta_0)} P\{(\mathbf{d}^T\theta-\mathbf{d}^T\mathbf{T}\mathbf{Y}-\mathbf{d}^T\mathbf{u})^2\leqslant c_\lambda^2\}$$

достигается при $\mathbf{u}=\mathbf{u}_{\theta_0}=\theta_0-\mathbf{T}_{0\lambda}\mathbf{X}\theta_0$ и $\mathbf{T}_{0\lambda}=(\mathbf{X}^T\mathbf{X}+\lambda\mathbf{A})^{-1}\mathbf{X}^T$.

В заключение рассмотрим построение оптимальной доверительной процедуры $C_{\mathbf{D},\mathbf{T},\mathbf{u}}(\cdot/\cdot)$ при $\mathbf{D}=\mathbf{X}^T\mathbf{X}$ и $\Omega^m=\{\theta\in R^m:\theta^T\mathbf{X}^T\mathbf{X}\theta\leqslant r^2\ (r>0)\}$.

Несложными преобразованиями критерий (2) приводится к виду:

$$\min_{\theta^T\theta\leqslant r^2} P\{[\mathbf{T}_1\mathcal{E}+(\mathbf{T}_1\mathbf{X}_1-\mathbf{I}_m)\theta+(\mathbf{X}^T\mathbf{X})^{1/2}\mathbf{u}]^T[\mathbf{T}_1\mathcal{E}+(\mathbf{T}_1\mathbf{X}_1-\mathbf{I}_m)\theta+$$

$$+(\mathbf{X}^T\mathbf{X})^{1/2}\mathbf{u}]\leqslant c^2\}, \tag{14}$$

где $\mathbf{T}_1=(\mathbf{X}^T\mathbf{X})^{1/2}\mathbf{T}$, $\mathbf{X}_1=\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1/2}$.

Существуют ортогональные матрицы $\mathbf{L}_1$, $\mathbf{M}_1$ и $\mathbf{L}_2$, $\mathbf{M}_2$ соответственно порядков m и n такие, что $\mathbf{X}_1^T=\mathbf{M}_1\mathbf{I}_{m0}\mathbf{M}_2$, $\mathbf{T}_1=\mathbf{L}_1\mathbf{S}\mathbf{I}_{m0}\mathbf{L}_2$, где $\mathbf{I}_{m0}=(\mathbf{I}_m\,\vdots\,\mathbf{O}_{m,n-m})$ ($\mathbf{O}_{m,n-m}$ — нулевая матрица порядка $m\times(n-m)$), а $\mathbf{S}$ — диагональная матрица порядка $m\times m$, диагональные элементы которой $s_1, s_2, \ldots, s_m$ есть собственные числа матрицы $(\mathbf{T}_1\mathbf{T}_1^T)^{1/2}$. Таким образом, критерий (14) преобразуется к виду:

$$\min_{\theta^{\mathrm T}\theta\leqslant r^2} P\{[\mathcal{E}+(\mathbf{I}_{m0}\mathbf{L}\mathbf{I}_{m0}^{\mathrm T}\mathbf{M}-\mathbf{S}^{-1})\theta+\mathbf{d}]^{\mathrm T}\mathbf{S}^{-2}[\mathcal{E}+(\mathbf{I}_{m0}\mathbf{L}\mathbf{I}_{m0}^{\mathrm T}\mathbf{M}-\mathbf{S}^{-1})\theta+ +\mathbf{d}]\leqslant c^2\}, \tag{15}$$

где $\mathbf{d}=\mathbf{S}^{-1}\mathbf{L}_1^{\mathrm T}(\mathbf{X}^{\mathrm T}\mathbf{X})^{1/2}\mathbf{u}$, а $\mathbf{L}=\mathbf{L}_2\mathbf{M}_2^{\mathrm T}$ и $\mathbf{M}=\mathbf{M}_1^{\mathrm T}\mathbf{L}_1$.
Невырожденность матрицы $\mathbf{S}$ будет ясна из дальнейшего. Критерий (15) требует максимизации по элементам $s_1, s_2, \ldots, s_m$, матрицам $\mathbf{L}$ и $\mathbf{M}$ и вектору $\mathbf{d}$.

Очевидно, минимум в (15) всегда реализуется на векторе θ, компоненты которого совпадают по знаку с соответствующими компонентами вектора $\mathbf{d}$, а на таких векторах (15) максимизируется по $\mathbf{d}$ при $\mathbf{d}=0$. Заметим, что

$$\max_{\theta^{\mathrm T}\theta\leqslant r^2} \theta^{\mathrm T}(\mathbf{I}_{m0}\mathbf{L}\mathbf{I}_{m0}^{\mathrm T}\mathbf{M}-\mathbf{S}^{-1})^{\mathrm T}(\mathbf{I}_{m0}\mathbf{L}\mathbf{I}_{m0}^{\mathrm T}\mathbf{M}-\mathbf{S}^{-1})\theta\geqslant \max_{1\leqslant i\leqslant m}(a_i-s_i^{-1})^2 r^2,$$

где $a_1, a_2, \ldots, a_m$ — диагональные элементы матрицы $\mathbf{I}_{m0}\mathbf{L}\mathbf{I}_{m0}^{\mathrm T}\mathbf{M}$, так как наибольшее собственное число матрицы не меньше максимального из диагональных элементов. Но $a_i\leqslant 1$, $1\leqslant i\leqslant m$, а для оптимальной матрицы $\mathbf{S}$ $s_i\leqslant 1$, $1\leqslant i\leqslant m$, следовательно,

$$\min_{a_1,a_2,\ldots,a_m}\max_{1\leqslant i\leqslant m} r^2(a_i-s_i^{-1})^2=r^2\max_i(1-s_i^{-1})^2.$$

Обозначив $\max\limits_{1\leqslant i\leqslant m}(1-s_i^{-1})^2$ через $(1-z)^2$, получим

$$\max_{\mathbf{T},\mathbf{u}}\min_{\theta^{\mathrm T}\mathbf{X}^{\mathrm T}\mathbf{X}\theta\leqslant r^2} P\{(\theta-\mathbf{T}\mathbf{Y}-\mathbf{u})^{\mathrm T}\mathbf{X}^{\mathrm T}\mathbf{X}(\theta-\mathbf{T}\mathbf{Y}-\mathbf{u})\leqslant c^2\}\leqslant$$
$$\leqslant\max_{z\geqslant 1} P\left\{\frac{1}{z^2}[\mathcal{E}+r(1-z)\theta_0]^{\mathrm T}[\mathcal{E}+r(1-z)\theta_0]\leqslant c^2\right\}, \tag{16}$$

где $\theta_0^{\mathrm T}=(1, 0, \ldots, 0)$ — m-мерный вектор. Нетрудно проверить, что (16) становится равенством при $\mathbf{T}_z=\frac{1}{z}(\mathbf{X}^{\mathrm T}\mathbf{X})^{-1}\mathbf{X}^{\mathrm T}$. Таким образом, доказана

Теорема 3. *Максимум критерия* (12) *достигается при* $\mathbf{u}=0$ *и матрице* $\mathbf{T}=\mathbf{T}_{z_0}=\frac{1}{z_0}=(\mathbf{X}^{\mathrm T}\mathbf{X})^{-1}\mathbf{X}^{\mathrm T}$, *где* z_0 *максимизирует по* z *интеграл*

$$\int\limits_{\left\{\mathbf{e}\in R^n:\,\frac{1}{z^2}[\mathbf{e}+r(1-z)\theta_0]^{\mathrm T}[\mathbf{e}+r(1-z)\theta_0]\leqslant c^2\right\}}\exp(-\mathbf{e}^{\mathrm T}\mathbf{e}/2)\,d\mathbf{e}.$$

ЛИТЕРАТУРА

1. Wallace D., Ann. math. statist., **30**, 864—876 (1959).
2. Pitman E. J. G., Biometrika, **80**, No. 3, 4 (1939).
3. Stein C. M., J. Roy. Statist. Soc., **B 24**, 265—296 (1962).
4. Joshi V. M., Ann. math. statist., **38**, No. 6 (1967).

5. Ольман В., Изв. АН ЭССР, Физ. Матем., **23**, 135 (1974).
6. Фихтенгольц Г. М., Курс дифф. и интегр. исчисления, т. 3, М., 1958, с. 676.
7. Кукс Я., Изв. АН ЭССР, Физ. Матем., **21**, 73 (1972).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
4/IX 1973

V. OLMAN

REGRESSIOONIMUDELI PARAMEETRI MINIMAKSNE USALDUSHINNANG NORMAALSE ADITIIVSE VEA PUHUL

Vaadeldakse optimaalse usaldusellipsoidi konstrueerimist vaatlusvektori lineaarteisenduse abil tsentreeritud ellipsoidide klassis. See usaldusellipsoid maksimeerib lineaarse regressioonimudeli parameetri väärtuse ellipsoidi kuuluvuse minimaalse tõenäosuse. Esitatakse ülesande lahendus parameetri lubatavate väärtuste mõnede konkreetsete piirkondade kohta.

V. OLMAN

MINIMAX CONFIDENCE ESTIMATION OF LINEAR REGRESSION MODEL PARAMETER UNDER NORMAL ADDITIVE ERROR

The problem of the construction of optimal confidence ellipsoid, centered by linear transform of the normal vector of observations, is considered. As the criterion, minimal probability of covering a linear regression model parameter by ellipsoid overall possible values of this parameter, is used. Solutions of the problem for some specific sets of possible model parameter values are given.