дачи программирования, предлагается вместо U_p использовать множество

$$U_{\gamma pf} = \{x : m_x + k_a \sigma_x \leqslant 0,\ x \geqslant 0,$$

$$k_a = (1/(1 - K_\gamma^2/2f))\,[K_p + \sqrt{K_p^2 - (1 - K_\gamma^2/2f)\,(K_p^2 - K_\gamma^2/n)}\,]\},$$

совпадающее с U_p с вероятностью γ.

Автор благодарит Т. Тобиаса за ряд ценных замечаний.

ЛИТЕРАТУРА

1. Charnes A., Cooper W. W., Op. Res., 11, 18 (1963).
2. Symonds C. H., Op. Res., 15, 495 (1967).
3. Рао С. Р., Линейные статистические методы и их применение, М., 1968.
4. Owen D. B., Technometrics, 10, 445 (1968).
5. Easterling R. G., J. Am. Stat. Ass., 64, 1031 (1969).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
7/V 1971

EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 20. KÖIDE
FÜÜSIKA * MATEMAATIKA. 1971, NR. 4

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 20
ФИЗИКА * МАТЕМАТИКА. 1971, № 4

https://doi.org/10.3176/phys.math.1971.4.21

УДК 519.281

Я. КУКС, В. ОЛЬМАН

МИНИМАКСНАЯ ЛИНЕЙНАЯ ОЦЕНКА КОЭФФИЦИЕНТОВ РЕГРЕССИИ. II

J. KUKS, V. OLMAN. REGRESSIOONIKOEFITSIENTIDE LINEAARNE MINIMAKSHINNANG. II

J. KUKS, V. OLMAN. A MINIMAX LINEAR ESTIMATOR OF REGRESSION COEFFICIENTS. II

В статистических исследованиях весьма часто встречается задача оценивания неизвестного m-мерного вектор-параметра θ в модели

$$E\{\mathbf{Y}\,|\,\theta\} = \mathbf{X}\theta, \quad E\{(\mathbf{Y} - \mathbf{X}\theta)(\mathbf{Y} - \mathbf{X}\theta)^{\mathsf{T}}\,|\,\theta\} = \sigma^2 \mathbf{I}_n, \tag{1}$$

где $E\{\mathbf{U}|\theta\}$ — математическое ожидание случайной матрицы $\mathbf{U}$ при фиксированном значении параметра θ; $\mathbf{Y}$ — случайный n-мерный вектор наблюдений; $\mathbf{X}$ — известная матрица порядка $n \times m$; $\mathbf{I}_n$ — единичная матрица порядка $n \times n$; σ^2 — известный положительный коэффициент. Ограничимся рассмотрением линейных по наблюдениям оценок $\hat{\theta} = \mathbf{TY}$ параметра θ, где $\mathbf{T}$ — матрица порядка $m \times n$. За критерий качества оценки примем супремум функции риска $E\{(\hat{\theta} - \theta)^{\mathsf{T}} \mathbf{P} (\hat{\theta} - \theta)\,|\,\theta\}$ по параметру θ, принадлежащему некоторому априорно заданному эллипсоиду с центром в нуле, т. е. функцию

$$Q(\mathbf{T}) = \sup_{\theta^{\mathsf T}\mathbf{A}\theta\leqslant 1} E\{(\mathbf{TY}-\theta)^{\mathsf T}\mathbf{P}(\mathbf{TY}-\theta)\,|\,\theta\},\tag{2}$$

где $\mathbf{P}$ и $\mathbf{A}$ — заданные симметричные неотрицательно определенные матрицы порядка $m\times m$.

Решим задачу определения оптимальной по критерию (2) *оценки, если выполнены следующие условия.*

1. Строки матрицы $\mathbf{P}$ представимы в виде линейных комбинаций строк матриц $\mathbf{X}^{\mathsf T}\mathbf{X}$ и $\mathbf{A}$. Это условие, как показано в работе [1], необходимо и достаточно для существования матрицы $\mathbf{T}$, при которой функция $Q(\mathbf{T})$ конечна.

2. Существует такая невырожденная матрица $\mathbf{C}$ порядка $m\times m$, что матрицы $\mathbf{C}^{\mathsf T}\mathbf{X}^{\mathsf T}\mathbf{XC}$, $\mathbf{C}^{\mathsf T}\mathbf{PC}$ и $\mathbf{C}^{\mathsf T}\mathbf{AC}$ диагональны. Это требование выполняется, если, например, две из матриц $\mathbf{X}^{\mathsf T}\mathbf{X}$, $\mathbf{P}$ и $\mathbf{A}$ совпадают с точностью до множителя.

Используя определение модели (1), получим

$$Q(\mathbf{T}) = \sup_{\theta^{\mathsf T}\mathbf{A}\theta\leqslant 1} \theta^{\mathsf T}(\mathbf{I}_m-\mathbf{X}^{\mathsf T}\mathbf{T}^{\mathsf T})\,\mathbf{P}\,(\mathbf{I}_m-\mathbf{TX})\,\theta+\sigma^2\,\mathrm{Sp}\,\mathbf{T}^{\mathsf T}\mathbf{PT}.\tag{3}$$

Легко показать [1], что среди матриц, минимизирующих $Q(\mathbf{T})$, существует матрица вида $\mathbf{T}=\mathbf{SX}^{\mathsf T}$, где $\mathbf{S}$ — матрица порядка $m\times m$. Таким образом, проблема сводится к минимизации по матрице $\mathbf{S}$ функции

$$Q_1(\mathbf{S}) = \sup_{\theta^{\mathsf T}\mathbf{A}\theta\leqslant 1} \theta^{\mathsf T}(\mathbf{I}_m-\mathbf{X}^{\mathsf T}\mathbf{XS}^{\mathsf T})\,\mathbf{P}\,(\mathbf{I}_m-\mathbf{SX}^{\mathsf T}\mathbf{X})\,\theta+\sigma^2\,\mathrm{Sp}\,\mathbf{S}^{\mathsf T}\mathbf{PSX}^{\mathsf T}\mathbf{X}$$

или по матрице $\mathbf{R}=\mathbf{C}^{-1}\mathbf{S}(\mathbf{C}^{\mathsf T})^{-1}$ функции

$$Q_2(\mathbf{R}) = \sup_{\beta^{\mathsf T}\mathbf{B}\beta\leqslant 1} \beta^{\mathsf T}(\mathbf{I}_m-\mathbf{WR}^{\mathsf T})\,\mathbf{F}\,(\mathbf{I}_m-\mathbf{RW})\,\beta+\sigma^2\,\mathrm{Sp}\,\mathbf{R}^{\mathsf T}\mathbf{FRW},$$

где $\beta=\mathbf{C}^{-1}\theta$, а $\mathbf{W}=\mathbf{C}^{\mathsf T}\mathbf{X}^{\mathsf T}\mathbf{XC}$, $\mathbf{F}=\mathbf{C}^{\mathsf T}\mathbf{PC}$, $\mathbf{B}=\mathbf{C}^{\mathsf T}\mathbf{AC}$ — диагональные матрицы с диагональными элементами w_i^2, f_i^2, b_i^2 $(i=1,\ldots,m)$ соответственно. Обозначим через β_i m-мерный вектор, i-я координата которого равна единице, а все остальные нули, $i=1,\ldots,m$, и через $l_1^2,\ldots,l_m^2$ — диагональные элементы матрицы $\mathbf{R}^{\mathsf T}\mathbf{FR}$. Пусть w_i, f_i, b_i, l_i — неотрицательные квадратные корни соответственно чисел w_i^2, f_i^2, b_i^2, l_i^2, $i=1,\ldots,m$. Тогда, используя неравенство Коши—Шварца, имеем

$$Q_2(\mathbf{R}) \geqslant \lim_{\varepsilon\to+0}\ \sup_{i\in\{1,\ldots,m\}} [(\beta_i^{\mathsf T}\mathbf{F}\beta_i-2\beta_i^{\mathsf T}\mathbf{WR}^{\mathsf T}\mathbf{F}\beta_i+\beta_i^{\mathsf T}\mathbf{WR}^{\mathsf T}\mathbf{FRW}\beta_i)/(b_i^2+\varepsilon)]+$$

$$+\sigma^2\sum_{i=1}^m l_i^2w_i^2 \geqslant \lim_{\varepsilon\to+0}\ \sup_{i\in\{1,\ldots,m\}} [(f_i-l_iw_i^2)^2/(b_i^2+\varepsilon)]+\sigma^2\sum_{i=1}^m l_i^2w_i^2 =$$

$$=\varphi(l_1,\ldots,l_m).\tag{4}$$

Обозначим через $l_1^*,\ldots,l_m^*$ те значения переменных $l_1,\ldots,l_m$, при которых минимизируется функция $\varphi(l_1,\ldots,l_m)$. Не умаляя общности, можно считать, что $b_i>0$, $w_i>0$ для $i=1,\ldots,r$; $b_i>0$, $w_i=0$ для $i=r+1,\ldots,s$; $b_i=0$, $w_i>0$ для $i=s+1,\ldots,t$; $b_i=0$, $w_i=0$ для $i=t+1,\ldots,m$; $f_i/b_i\leqslant f_{i+1}/b_{i+1}$ для $i=1,\ldots,r-1$ и $f_i/b_i\leqslant f_s/b_s$ для $i=r+1,\ldots,s$ $(1\leqslant r\leqslant s\leqslant t\leqslant m)$. При $w_i=0$ минимизируемая функция от l_i не зависит и можно выбрать $l_i^*=0$, $i=r+1,\ldots,s,\ t+1,\ldots,m$. Поскольку из условия 1 следует, что $f_i=0$, $i=t+1,\ldots,m$, то для конечности функции $\varphi(l_1,\ldots,l_m)$ необходимо и достаточно, чтобы $f_i=l_iw_i^2$, $i=s+1,\ldots,t$. Следовательно, $l_i^*=f_i/w_i^2$, $i=s+1,\ldots,t$ и

$$\min_{l_1,\ldots,l_m} \varphi(l_1,\ldots,l_m) = \min_{l_1,\ldots,l_r} \{ \sup_{i\in\{1,\ldots,s\}} [(f_i - l_i w_i^2)^2/b_i^2] + \sigma^2 \sum_{i=1}^{r} l_i^2 w_i^2\} + \sigma^2 \sum_{i=s+1}^{t} f_i^2/w_i^2.$$

Пусть $c^* = \sup_{i\in\{1,\ldots,s\}} (f_i - l_i^* w_i^2)/b_i$. Тогда, очевидно, $l_i^* = \max[0, (f_i - c^* b_i)/w_i^2]$, $i = 1, \ldots, r$. Нетрудно показать, что функция

$$\psi(c) = c^2 + \sigma^2 \sum_{i=1}^{r} w_i^2 [\max[0, (f_i - cb_i)/w_i^2]]^2$$

на интервале $[(f_s/b_s)\operatorname{sign}(s-r), \infty)$ достигает минимального значения в точке c^*. Вычисление c^* не представляет трудности, так как $\psi(c)$ является строго выпуклой функцией и на интервале $[f_{j-1}/b_{j-1}, f_j/b_j]$ представляет собой параболу

$$\psi_j(c) = c^2 + \sigma^2 \sum_{i=j}^{r} (f_i - cb_i)^2/w_i^2, \quad j = 1, \ldots, r,$$

где $f_0 = 0$, $b_0 = 1$ и $\psi(c) = c^2$ при $c \geqslant f_r/b_r$.

Пусть $\mathbf{R}^*$ — диагональная матрица порядка $m \times m$ с i-м диагональным элементом $r_i^* = l_i^*/f_i$, если $f_i \neq 0$, и $r_i^* = 0$, если $f_i = 0$, $i = 1, \ldots, m$. Подставляя матрицу $\mathbf{T}^* = \mathbf{C}\mathbf{R}^*\mathbf{C}^\mathrm{T}\mathbf{X}^\mathrm{T}$ в выражение (3), легко проверить, что $Q(\mathbf{T}^*) = \min_{l_1,\ldots,l_m} \varphi(l_1, \ldots, l_m)$ и, следовательно, $Q(\mathbf{T}^*) = \min_{\mathbf{T}} Q(\mathbf{T})$.

Отметим, что в частном случае, когда $\mathbf{P} = g\mathbf{A}$, где $g > 0$ (а следовательно, $f_j/b_j = g$, $j = 1, \ldots, r$), на интервале $0 \leqslant c \leqslant g$ имеем

$$\psi(c) = c^2 + \sigma^2 \sum_{i=1}^{r} (g - c)^2 b_i^2/w_i^2$$

и

$$c^* = \max[f_s/b_s, (g\sigma^2 \sum_{i=1}^{r} b_i^2/w_i^2)/(1 + \sigma^2 \sum_{i=1}^{r} b_i^2/w_i^2)].$$

Рассматривая оценки вида $\hat{\theta} = \mathbf{t} + \mathbf{T}\mathbf{Y}$, где $\mathbf{t}$ — m-мерный вектор, и определяя критерий качества оценки функцией

$$q(\mathbf{t}, \mathbf{T}) = \sup_{(\theta-\theta_0)^\mathrm{T} \mathbf{A}(\theta-\theta_0) \leqslant 1} E\{(\mathbf{t} + \mathbf{T}\mathbf{Y} - \theta)^\mathrm{T} \mathbf{P} (\mathbf{t} + \mathbf{T}\mathbf{Y} - \theta) \mid \theta\},$$

можно показать, что $q(\mathbf{t}^*, \mathbf{T}^*) = \min_{\mathbf{t},\mathbf{T}} q(\mathbf{t}, \mathbf{T})$, где $\mathbf{t}^* = (\mathbf{I}_m - \mathbf{T}^*\mathbf{X})\theta_0$, а матрица $\mathbf{T}^*$ определена выше.

ЛИТЕРАТУРА

1. Кукс Я., Изв. АН ЭССР, Физ. Матем., 21 (в печати).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
21/VI 1971