EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 21. KÖIDE
FÜÜSIKA * MATEMAATIKA. 1972, NR. 1

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 21
ФИЗИКА * МАТЕМАТИКА. 1972, № 1

https://doi.org/10.3176/phys.math.1972.1.09

УДК 519.281

Я. КУКС

МИНИМАКСНАЯ ОЦЕНКА КОЭФФИЦИЕНТОВ РЕГРЕССИИ

В настоящей работе рассматривается решение несколько более общей, чем в работах [1, 2], минимаксной задачи оценивания коэффициентов регрессии в модели

$$E\{\mathbf{Y} \mid \theta\} = \mathbf{X}\theta, \quad \operatorname{Cov}\{\mathbf{Y} \mid \theta\} = \sigma^2 \mathbf{I}_n, \tag{1}$$

где $E\{\mathbf{Y}|\theta\}$ и $\operatorname{Cov}\{\mathbf{Y}|\theta\}$ — соответственно математическое ожидание и ковариационная матрица n-мерного случайного вектора $\mathbf{Y}$ при фиксированном значении m-мерного вектор-параметра θ; $\mathbf{X}$ — известная матрица порядка $n \times m$ и ранга l; $\mathbf{I}_n$ — единичная матрица порядка $n \times n$; σ^2 — известный положительный коэффициент. Ограничимся рассмотрением оценок $\hat{\theta}$ параметра θ вида $\hat{\theta} = \mathbf{t} + \mathbf{T}\mathbf{Y}$, где $\mathbf{t}$ — m-мерный вектор, а $\mathbf{T}$ — матрица порядка $m \times n$. Пусть $\mathbf{P}$ — некоторая симметричная неотрицательно определенная матрица порядка $m \times m$ и ранга r. Определим функцию риска в виде

$$R(\mathbf{t}, \mathbf{T}, \theta) = E\{(\mathbf{t} + \mathbf{T}\mathbf{Y} - \theta)^{\mathrm{T}} \mathbf{P} (\mathbf{t} + \mathbf{T}\mathbf{Y} - \theta) \mid \theta\} \tag{2}$$

и за критерий качества оценки примем функцию

$$Q(\mathbf{t}, \mathbf{T}) = \sup_{\mathbf{A}\theta \in \Omega_k} R(\mathbf{t}, \mathbf{T}, \theta), \tag{3}$$

где $\mathbf{A}$ — матрица порядка $k \times m$, ранг которой равен k; Ω_k — замкнутое ограниченное множество в k-мерном евклидовом пространстве R^k. *Проблема заключается в определении вектора* $\mathbf{t}^*$ *и матрицы* $\mathbf{T}^*$*, при которых достигает минимума функция* $Q(\mathbf{t}, \mathbf{T})$.

С учетом определения модели (1) функция риска (2) приводится к виду

$$R(\mathbf{t}, \mathbf{T}, \theta) = \mathbf{t}^{\mathrm{T}}\mathbf{P}\mathbf{t} - 2\mathbf{t}^{\mathrm{T}}\mathbf{P}(\mathbf{I}_m - \mathbf{T}\mathbf{X})\theta + \theta^{\mathrm{T}}(\mathbf{I}_m - \mathbf{X}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}})\mathbf{P}(\mathbf{I}_m - \mathbf{T}\mathbf{X})\theta + \sigma^2 \operatorname{Sp} \mathbf{T}^{\mathrm{T}}\mathbf{P}\mathbf{T}. \tag{4}$$

Среди матриц, минимизирующих функцию (3) по $\mathbf{T}$, существует матрица вида $\mathbf{T} = \mathbf{S}\mathbf{X}^{\mathrm{T}}$, где $\mathbf{S}$ — матрица порядка $m \times m$. Действительно, любую матрицу $\mathbf{T}$ порядка $m \times n$ можно представить в виде $\mathbf{T} = \mathbf{S}\mathbf{X}^{\mathrm{T}} + \mathbf{S}_1\mathbf{Z}^{\mathrm{T}}$, где $\mathbf{S}_1$ — матрица порядка $m \times (n-l)$; $\mathbf{Z}^{\mathrm{T}}$ — матрица порядка $(n-l) \times n$ и $\mathbf{Z}^{\mathrm{T}}\mathbf{X} = 0$. В выражении (4) слагаемое $\sigma^2 \operatorname{Sp} \mathbf{T}^{\mathrm{T}}\mathbf{P}\mathbf{T}$ принимает минимальное значение при нулевой матрице $\mathbf{S}_1$, а остальные слагаемые от матрицы $\mathbf{S}_1$ не зависят. Следовательно,

$$\min_{\mathbf{t},\mathbf{T}} Q(\mathbf{t}, \mathbf{T}) = \min_{\mathbf{t},\mathbf{S}} Q(\mathbf{t}, \mathbf{S}\mathbf{X}^{\mathrm{T}}) = \min_{\mathbf{t},\mathbf{S}} Q_1(\mathbf{t}, \mathbf{S}),$$

где

$$Q_1(\mathbf{t}, \mathbf{S}) = \sup_{\mathbf{A}\theta\in\Omega_k} [\mathbf{t}^T\mathbf{P}\mathbf{t} - 2\mathbf{t}^T\mathbf{P}(\mathbf{I}_m - \mathbf{S}\mathbf{X}^T\mathbf{X})\theta + \theta^T(\mathbf{I}_m - \mathbf{X}^T\mathbf{X}\mathbf{S}^T)\mathbf{P}(\mathbf{I}_m - \mathbf{S}\mathbf{X}^T\mathbf{X})\theta] + \sigma^2 \operatorname{Sp}\mathbf{X}\mathbf{S}^T\mathbf{P}\mathbf{S}\mathbf{X}^T. \tag{5}$$

Пусть $\mathbf{U}$ — произвольная матрица порядка $i \times j$. Тогда существует единственная [3] матрица $\mathbf{U}^+$ порядка $j \times i$, удовлетворяющая соотношениям

$$\mathbf{U}\mathbf{U}^+\mathbf{U} = \mathbf{U}, \quad \mathbf{U}^+\mathbf{U}\mathbf{U}^+ = \mathbf{U}^+, \quad (\mathbf{U}^+\mathbf{U})^T = \mathbf{U}^+\mathbf{U}, \quad (\mathbf{U}\mathbf{U}^+)^T = \mathbf{U}\mathbf{U}^+.$$

Через $\mathbf{U}^-$ будем обозначать матрицу порядка $j \times i$, удовлетворяющую равенству $\mathbf{U}\mathbf{U}^-\mathbf{U} = \mathbf{U}$. Пусть $\mathbf{B}$ — матрица порядка $(m-k) \times m$, строки которой образуют базис $(m-k)$-мерного многообразия в R^m, ортогонального к многообразию, порожденному строками матрицы $\mathbf{A}$. Нетрудно проверить, что

$$\mathbf{A}^+\mathbf{A} + \mathbf{B}^+\mathbf{B} = \mathbf{I}_m.$$

Обозначим через $\mathbf{p}_i$, $i = 1, \ldots, r$, собственные векторы матрицы $\mathbf{P}$, соответствующие ее положительным собственным значениям λ_i, $i = 1, \ldots, r$, тогда

$$Q_1(\mathbf{t}, \mathbf{S}) = \sup_{\mathbf{A}\theta\in\Omega_k} \{\sum_{i=1}^{r} \lambda_i [\mathbf{t}^T\mathbf{p}_i - \mathbf{p}_i^T(\mathbf{I}_m - \mathbf{S}\mathbf{X}^T\mathbf{X})(\mathbf{A}^+\mathbf{A} + \mathbf{B}^+\mathbf{B})\theta]^2\} + \sigma^2 \operatorname{Sp}\mathbf{X}\mathbf{S}^T(\sum_{i=1}^{r} \lambda_i \mathbf{p}_i\mathbf{p}_i^T)\mathbf{S}\mathbf{X}^T. \tag{6}$$

Из представления (6) нетрудно видеть, что, поскольку условие $\mathbf{A}\theta \in \Omega_k$ не накладывает никаких ограничений на $(m-k)$-мерный вектор $\mathbf{B}\theta$, функция $Q_1(\mathbf{t}, \mathbf{S})$ принимает конечное значение при тех и только тех матрицах $\mathbf{S}$, которые для всех $i = 1, \ldots, r$ удовлетворяют равенству

$$\mathbf{p}_i^T\mathbf{S}\mathbf{X}^T\mathbf{X}\mathbf{B}^+ = \mathbf{p}_i^T\mathbf{B}^+. \tag{7}$$

Уравнение

$$\mathbf{u}_i^T\mathbf{X}^T\mathbf{X}\mathbf{B}^+ = \mathbf{p}_i^T\mathbf{B}^+ \tag{8}$$

тогда и только тогда разрешимо относительно вектора $\mathbf{u}_i$, когда существует такой вектор η_i, что

$$\mathbf{p}_i^T = \eta_i^T(\mathbf{X}^T\mathbf{X} + \mathbf{A}^+\mathbf{A}). \tag{9}$$

Действительно, из равенства (8) следует, что $\mathbf{p}_i^T = \mathbf{p}_i^T(\mathbf{A}^+\mathbf{A} + \mathbf{B}^+\mathbf{B}) = \mathbf{p}_i^T\mathbf{A}^+\mathbf{A} + \mathbf{u}_i^T\mathbf{X}^T\mathbf{X}\mathbf{B}^+\mathbf{B} = \mathbf{p}_i^T\mathbf{A}^+\mathbf{A} - \mathbf{u}_i^T\mathbf{X}^T\mathbf{X}\mathbf{A}^+\mathbf{A} + \mathbf{u}_i^T\mathbf{X}^T\mathbf{X}$, т. е. вектор $\mathbf{p}_i$ принадлежит многообразию, порожденному строками матриц $\mathbf{A}$ и $\mathbf{X}$. Но такое же многообразие порождают строки матрицы $\mathbf{X}^T\mathbf{X} + \mathbf{A}^+\mathbf{A} = (\mathbf{X}^T \mid \mathbf{A}^+\mathbf{A})(\mathbf{X}^T \mid \mathbf{A}^+\mathbf{A})^T$. Если справедливо представление (9), то $\mathbf{p}_i^T\mathbf{B}^+ = \eta_i^T\mathbf{X}^T\mathbf{X}\mathbf{B}^+$, так как $\mathbf{A}\mathbf{B}^+ = \mathbf{A}\mathbf{B}^+\mathbf{B}\mathbf{B}^+ = \mathbf{A}\mathbf{B}^T\mathbf{B}^{+T}\mathbf{B}^+ = 0$. Вводя обозначение $\mathbf{H} = \mathbf{I}_m - \mathbf{X}^T\mathbf{X}\mathbf{B}^+(\mathbf{X}^T\mathbf{X}\mathbf{B}^+)^-$, общее решение уравнения (8) получаем в виде [3]

$$\mathbf{u}_i^T = \eta_i^T + \nu_i^T\mathbf{H}, \tag{10}$$

где ν_i — произвольный m-мерный вектор. Система уравнений относительно матрицы $\mathbf{S}$

$$\mathbf{p}_i^{\mathrm{T}}\mathbf{S} = \eta_i^{\mathrm{T}} + \nu_i^{\mathrm{T}}\mathbf{H}, \quad i = 1, \ldots, r, \tag{11}$$

всегда допускает решение в силу линейной независимости векторов $\mathbf{p}_i$, $i = 1, \ldots, r$. *Следовательно, принадлежность векторов* $\mathbf{p}_i$, $i = 1, \ldots$ $\ldots, r$, *линейному многообразию, порожденному строками матриц* $\mathbf{X}$ *и* $\mathbf{A}$, *является необходимым и достаточным условием для существования вектора* $\mathbf{t}$ *и матрицы* $\mathbf{S}$, *при которых значение функции* $Q_1(\mathbf{t}, \mathbf{S})$ *конечно.* В последующем изложении считаем это условие выполненным.

Таким образом, проблема минимизации $Q_1(\mathbf{t}, \mathbf{S})$ сводится к оптимальному выбору векторов $\mathbf{t}$ и ν_i, $i = 1, \ldots, r$. Используя (7), (9) и (11), получаем

$$\begin{aligned}\mathbf{p}_i^{\mathrm{T}}(\mathbf{I}_m - \mathbf{S}\mathbf{X}^{\mathrm{T}}\mathbf{X})\theta &= \mathbf{p}_i^{\mathrm{T}}(\mathbf{I}_m - \mathbf{S}\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{A}^{+}\mathbf{A}\theta = \\ &= \eta_i^{\mathrm{T}}\mathbf{A}^{+}\mathbf{A}\theta - \nu_i^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{A}^{+}\mathbf{A}\theta,\end{aligned} \tag{12}$$

$$\begin{aligned}\operatorname{Sp}\mathbf{X}\mathbf{S}^{\mathrm{T}}\Big(\sum_{i=1}^{r}\lambda_i\mathbf{p}_i\mathbf{p}_i^{\mathrm{T}}\Big)\mathbf{S}\mathbf{X}^{\mathrm{T}} &= \sum_{i=1}^{r}\lambda_i\mathbf{p}_i^{\mathrm{T}}\mathbf{S}\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{S}^{\mathrm{T}}\mathbf{p}_i = \\ &= \sum_{i=1}^{r}\lambda_i[\eta_i^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\eta_i + 2\nu_i^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X}\eta_i + \nu_i^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{H}^{\mathrm{T}}\nu_i].\end{aligned} \tag{13}$$

Вводя обозначения $\beta = \mathbf{A}\theta$, $\mathbf{z} = (\mathbf{t}^{\mathrm{T}}, \nu_1^{\mathrm{T}}, \ldots, \nu_r^{\mathrm{T}})^{\mathrm{T}}$ и подставляя (12) и (13) в (6), получаем для минимизируемой функции следующее выражение:

$$Q_2(\mathbf{z}) = \sup_{\beta\in\Omega_k} R_2(\mathbf{z}, \beta), \tag{14}$$

где

$$\begin{aligned}R_2(\mathbf{z}, \beta) = \sum_{i=1}^{r}\lambda_i[&(\mathbf{t}^{\mathrm{T}}\mathbf{p}_i - \eta_i^{\mathrm{T}}\mathbf{A}^{+}\beta + \nu_i^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{A}^{+}\beta)^2 + \\ &+ \sigma^2(\eta_i^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\eta_i + 2\nu_i^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X}\eta_i + \nu_i^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{H}^{\mathrm{T}}\nu_i)].\end{aligned} \tag{15}$$

Выражение в квадратных скобках в формуле (15) при фиксированном векторе β, как легко проверить, есть выпуклая функция по $\mathbf{z}$ для всех $i = 1, \ldots, r$, а следовательно, и функция $Q_2(\mathbf{z}) = \sup_{\beta\in\Omega_k} R_2(\mathbf{z}, \beta)$ выпукла по $\mathbf{z}$. Для ее минимизации по вектору z в общем случае нам кажется наиболее подходящим метод, изложенный в работе [4] и состоящий в построении последовательности $\{\mathbf{z}^j\}$, определяемой соотношениями

$$\mathbf{z}^{j+1} = \mathbf{z}^j - \varrho_j\mathbf{c}^j/\|\mathbf{c}^j\|, \quad \varrho_j \geqslant 0, \quad \varrho_j \to 0, \quad \sum_{j=0}^{\infty}\varrho_j = \infty,$$

где вектор $\mathbf{c}^j$ — опорный к $Q_2(\mathbf{z})$ в точке $\mathbf{z}^j = (\mathbf{t}^{j\mathrm{T}}, \nu_1^{j\mathrm{T}}, \ldots, \nu_r^{j\mathrm{T}})^{\mathrm{T}}$.

Можно показать, что функция $Q_2(\mathbf{z})$ удовлетворяет условию теоремы 2 работы [4] и, следовательно, $\lim_{j\to\infty} Q_2(\mathbf{z}^j) = \min_{\mathbf{z}} Q_2(\mathbf{z})$ при любом значении $\mathbf{z}^0$. По-видимому, хорошее начальное приближение $\mathbf{z}^0$ можно получить, включая Ω_k в некоторый эллипсоид и используя результаты работы [1]. Опорным к выпуклой функции $Q_2(\mathbf{z})$ в точке $\mathbf{z}^j$ называется вектор $\mathbf{c}^j$, удовлетворяющий условию $Q_2(\mathbf{z}) - Q_2(\mathbf{z}^j) \geqslant \mathbf{c}^{j\mathrm{T}}(\mathbf{z} - \mathbf{z}^j)$. Нетрудно проверить, что в настоящем случае такому условию удовлетворяет вектор

$$c^j = [\partial R_2(\mathbf{z}, \beta^{*j})/\partial \mathbf{z}]_{\mathbf{z}=\mathbf{z}^j},$$

т. е. вектор частных производных функции (15) по координатам вектора $\mathbf{z}$ в точке $\mathbf{z} = \mathbf{z}^j$ при фиксированном векторе $\beta = \beta^{*j}$, где β^{*j} — любой такой вектор из Ω_k, что $R_2(\mathbf{z}^j, \beta^{*j}) = \sup\limits_{\beta\in\Omega_k} R_2(\mathbf{z}^j, \beta)$.

Таким образом, единственная трудность в решении задачи — указать вектор $\beta^* \in \Omega_k$, на котором при фиксированном векторе $\mathbf{z}$ достигает максимума квадратичная выпуклая по β функция $R_2(\mathbf{z}, \beta)$. Каким образом решается эта проблема и разрешима ли она вообще — зависит от свойств множества Ω_k и матрицы $\mathbf{P}$.

Замечание 1. Допустим, что множество Ω_k является симметричным относительно некоторого вектора β_0, т. е. $\beta \in \Omega_k \Rightarrow (\beta_0 - (\beta - \beta_0)) \in \Omega_k$. Пусть $\Omega_{k\beta_0}$ — такое множество, что $\beta \in \Omega_k \Leftrightarrow (\beta - \beta_0) \in \Omega_{k\beta_0}$, тогда $(\beta - \beta_0) \in \Omega_{k\beta_0} \Leftrightarrow -(\beta - \beta_0) \in \Omega_{k\beta_0}$. Для матриц $\mathbf{S}$, удовлетворяющих условию (7), функция (6) представима в виде

$$Q_1(\mathbf{t}, \mathbf{S}) = \sup_{(\beta-\beta_0)\in\Omega_{k\beta_0}} \sum_{i=1}^{r} \lambda_i [\mathbf{p}_i^{\mathrm{T}}\mathbf{t} - \mathbf{p}_i^{\mathrm{T}}(\mathbf{I}_m - \mathbf{S}\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{A}^{+}\beta_0 -$$
$$- \mathbf{p}_i^{\mathrm{T}}(\mathbf{I}_m - \mathbf{S}\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{A}^{+}(\beta - \beta_0)]^2 + \sigma^2 \operatorname{Sp} \mathbf{X}\mathbf{S}^{\mathrm{T}}\Big(\sum_{i=1}^{r} \lambda_i \mathbf{p}_i \mathbf{p}_i^{\mathrm{T}}\Big)\mathbf{S}\mathbf{X}^{\mathrm{T}},$$

откуда легко следует, что оптимальным значением вектора $\mathbf{t}$ при любой матрице $\mathbf{S}$ является

$$\mathbf{t}_{\mathbf{S}} = (\mathbf{I}_m - \mathbf{S}\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{A}^{+}\beta_0.$$

Следовательно, задача сводится к минимизации по $\mathbf{S}$ функции $Q_1(\mathbf{t}_{\mathbf{S}}, \mathbf{S})$ или, что равносильно, к минимизации по вектору $\nu = (\nu_1^{\mathrm{T}}, \ldots, \nu_r^{\mathrm{T}})^{\mathrm{T}}$ функции

$$Q_3(\nu) = \sup_{\beta\in\Omega_k} R_3(\nu, \beta), \tag{16}$$

где

$$R_3(\nu, \beta) = \sum_{i=1}^{r} \lambda_i \{[(\eta_i^{\mathrm{T}} - \nu_i^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{A}^{+}(\beta - \beta_0)]^2 +$$
$$+ \sigma^2(\eta_i^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\eta_i + 2\nu_i^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X}\eta_i + \nu_i^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{H}^{\mathrm{T}}\nu_i)\}. \tag{17}$$

Для минимизации функции $Q_3(\nu)$ можно использовать описанный выше итерационный процесс.

Замечание 2. Если Ω_k — многогранник в k-мерном евклидовом пространстве, то супремум по $\beta \in \Omega_k$ функции $R_2(\mathbf{z}, \beta)$ (функции $R_3(\nu, \beta)$) при фиксированном векторе $\mathbf{z}$ (векторе ν) достигается в одной из вершин многогранника.

Замечание 3. При оценивании функции $f(\mathbf{p}) = \mathbf{p}^{\mathrm{T}}\theta$ статистикой $\hat{f}(\mathbf{p}) = \mathbf{p}^{\mathrm{T}}(\mathbf{t} + \mathbf{T}\mathbf{Y})$, где $\mathbf{p}$ — m-мерный вектор, в некоторых случаях естественным является критерий оценки $\sup\limits_{\mathbf{A}\theta\in\Omega_k} E\{[f(\mathbf{p}) - \hat{f}(\mathbf{p})]^2 \mid \theta\}$, совпадающий с критерием (3) при $\mathbf{P} = \mathbf{p}\mathbf{p}^{\mathrm{T}}$.

Пусть $\mathbf{P} = \mathbf{p}\mathbf{p}^{\mathrm{T}} = \lambda_1 \mathbf{p}_1 \mathbf{p}_1^{\mathrm{T}}$, тогда супремум по $\beta \in \Omega_k$ функции $R_2(\mathbf{z}, \beta)$ (функции $R_3(\nu, \beta)$) при фиксированном векторе $\mathbf{z}$ (векторе ν) достигается в точке, где линейная функция $L(\beta) = \eta_1^{\mathrm{T}}\mathbf{A}^{+}\beta - \nu_1^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{A}^{+}\beta$ принимает либо максимальное, либо минимальное значение на множестве Ω_k.

Таким образом, задача определения максимума квадратичной функции сводится к отысканию максимумов линейных функций $L(\beta)$ и

$-L(\beta)$ на множестве Ω_k, а если Ω_k — многогранник, то к задаче линейного программирования.

Замечание 4. Допустим, что $\mathbf{A}\theta \in \Omega_k \Leftrightarrow (\theta-\theta_0)^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}(\theta-\theta_0) \leqslant c$, где $c > 0$. Множество Ω_k симметрично относительно вектора $\beta_0 = \mathbf{A}\theta_0$ и согласно замечанию 1 определение оптимальной оценки сводится к минимизации функции (16). Пусть β^* — такой вектор, что $R_3(\nu,\beta^*) = \sup\limits_{\beta\in\Omega_k} R_3(\nu,\beta)$. Тогда $(\beta^*-\beta_0)/c^{1/2}$ является, очевидно, собственным вектором матрицы $\sum\limits_{i=1}^{r}\lambda_i\mathbf{A}^{+\mathrm{T}}(\eta_i^{\mathrm{T}}-\nu_i^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X})^{\mathrm{T}}(\eta_i^{\mathrm{T}}-\nu_i^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{A}^{+}$, соответствующим ее максимальному собственному числу, и вычисление вектора β^* при фиксированном векторе ν не представляет принципиальной трудности.

Кроме того, если $\mathbf{P} = \mathbf{p}\mathbf{p}^{\mathrm{T}}$ или существует такая невырожденная матрица $\mathbf{C}$ порядка $m\times m$, что матрицы $\mathbf{C}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{C}$, $\mathbf{C}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{C}$ и $\mathbf{C}^{\mathrm{T}}\mathbf{P}\mathbf{C}$ являются диагональными, то простые способы определения оптимальной по критерию (3) оценки указаны соответственно в работах [1] и [2].

Замечание 5. Допустим $\mathbf{P} = \mathbf{p}\mathbf{p}^{\mathrm{T}} = \lambda_1\mathbf{p}_1\mathbf{p}_1^{\mathrm{T}}$. Пусть для $i,j = 1,\ldots,s$ матрица $\mathbf{A}_i$ порядка $k_i\times m$ и ранга k_i удовлетворяет условию $\mathbf{A}_i\mathbf{A}_j^{\mathrm{T}} = 0$, если $i\neq j$, и Ω_{k_i} — замкнутое ограниченное симметричное относительно вектора $\beta_{i,0}$ множество в k_i-мерном евклидовом пространстве. Матрицу $\mathbf{A}$ и множество Ω_k определим следующим образом:

$$\mathbf{A} = (\mathbf{A}_1^{\mathrm{T}} \vdots \ldots \vdots \mathbf{A}_s^{\mathrm{T}})^{\mathrm{T}} \quad \text{и} \quad \mathbf{A}\theta\in\Omega_k \Leftrightarrow \mathbf{A}_i\theta\in\Omega_{k_i}, \quad i=1,\ldots,s.$$

Нетрудно проверить, что $\mathbf{A}^{+} = (\mathbf{A}_1^{+} \vdots \ldots \vdots \mathbf{A}_s^{+})$ и, следовательно, $\mathbf{A}^{+}\mathbf{A} = \sum\limits_{i=1}^{s}\mathbf{A}_i^{+}\mathbf{A}_i$. Введем обозначения $\theta_0 = \sum\limits_{i=1}^{s}\mathbf{A}_i^{+}\beta_{i,0}$, $\beta_0 = \mathbf{A}\theta_0$, $\beta_i = \mathbf{A}_i\theta$, $i=1,\ldots,s$, и $\beta = \mathbf{A}\theta = (\beta_1^{\mathrm{T}},\ldots,\beta_s^{\mathrm{T}})^{\mathrm{T}}$. Очевидно, $\beta_0 = (\beta_{1,0}^{\mathrm{T}},\ldots,\beta_{s,0}^{\mathrm{T}})^{\mathrm{T}}$ и множество Ω_k симметрично относительно вектора β_0. Тем самым определение оптимальной оценки сводится к минимизации функции (16). При минимизации (16) на каждом шаге описанного выше итерационного процесса требуется определить супремум по β функции $R_3(\nu,\beta)$ при фиксированном векторе ν. При сделанных предположениях

$$\sup_{\beta\in\Omega_k}[(\eta_1^{\mathrm{T}}-\nu_1^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{A}^{+}(\beta-\beta_0)]^2 =$$

$$= \sup_{\mathbf{A}\theta\in\Omega_k}[(\eta_1^{\mathrm{T}}-\nu_1^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{A}^{+}\mathbf{A}(\theta-\theta_0)]^2 =$$

$$= \sup_{\mathbf{A}_i\theta\in\Omega_{k_i},\, i=1,\ldots,s}[(\eta_1^{\mathrm{T}}-\nu_1^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X})\sum_{i=1}^{s}\mathbf{A}_i^{+}\mathbf{A}_i(\theta-\theta_0)]^2 =$$

$$= \sup_{\beta_i\in\Omega_{k_i},\, i=1,\ldots,s}[\sum_{i=1}^{s}(\eta_1^{\mathrm{T}}-\nu_1^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{A}_i^{+}(\beta_i-\beta_{i,0})]^2 =$$

$$= [\sum_{i=1}^{s}\sup_{\beta_i\in\Omega_{k_i}}(\eta_1^{\mathrm{T}}-\nu_1^{\mathrm{T}}\mathbf{H}\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{A}_i^{+}(\beta_i-\beta_{i,0})]^2.$$

Следовательно,

$$R_3(\nu, \beta^*) = \sup_{\beta \in \Omega_k} R_3(\nu, \beta), \quad \text{если} \quad \beta^* = (\beta_1^{*\mathrm{T}}, \ldots, \beta_s^{*\mathrm{T}})^{\mathrm{T}},$$

где β_i^* такой вектор, что

$$(\eta_1^{\mathrm{T}} - \nu_1^{\mathrm{T}} \mathbf{HX}^{\mathrm{T}}\mathbf{X}) \mathbf{A}_i^{+} (\beta_i^* - \beta_{i,0}) =$$

$$= \sup_{\beta_i \in \Omega_{k_i}} (\eta_1^{\mathrm{T}} - \nu_1^{\mathrm{T}} \mathbf{HX}^{\mathrm{T}}\mathbf{X}) \mathbf{A}_i^{+} (\beta_i - \beta_{i,0}), \quad i = 1, \ldots, s.$$

В некоторых случаях приведение вычисления вектора β^* к вычислению векторов β_i^*, $i = 1, \ldots, s$, дает существенное упрощение в решении задачи. Так, например, если для $i = 1, \ldots, s$

$$\mathbf{A}_i\theta \in \Omega_{k_i} \Leftrightarrow (\theta - \theta_{i,0})^{\mathrm{T}} \mathbf{A}_i^{\mathrm{T}} \mathbf{A}_i (\theta - \theta_{i,0}) \leqslant c_i, \quad \text{где} \quad c_i \geqslant 0,$$

то $\beta_{i,0} = \mathbf{A}_i\theta_{i,0}$, $\beta_i^* - \beta_{i,0} = \dfrac{\mathbf{A}_i^{+\mathrm{T}} (\eta_1 - \mathbf{X}^{\mathrm{T}}\mathbf{XH}^{\mathrm{T}}\nu_1) c_i^{1/2}}{[(\eta_1^{\mathrm{T}} - \nu_1^{\mathrm{T}}\mathbf{HX}^{\mathrm{T}}\mathbf{X}) \mathbf{A}_i^{+} \mathbf{A}_i^{+\mathrm{T}} (\eta_1 - \mathbf{X}^{\mathrm{T}}\mathbf{XH}^{\mathrm{T}}\nu_1)]^{1/2}}$

при неравенстве нулю знаменателя последнего выражения и вектор β_i^* произволен в противном случае.

ЛИТЕРАТУРА

1. Кукс Я., Ольман В., Изв. АН ЭССР, Физ. Матем., **21**, 66 (1972).
2. Кукс Я., Ольман В., Изв. АН ЭССР, Физ. Матем., **20**, 480 (1971).
3. Рао С. Р., Линейные статистические методы и их применения, М., 1968.
4. Поляк Б. Т., Докл. АН СССР, **174**, 33 (1967).

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
21/VI 1971

J. KUKS

REGRESSIOONIKOEFITSIENTIDE MINIMAKSHINNANG

Optimaalset hinnangut lineaarse regerssioonmudeli kordajate m-dimensionaalsele vektorile θ otsitakse konstantse vektori ja vaatluste suhtes lineaarse vektorfunktsiooni summana esitatavate hinnangute klassist. Optimiseerimise eesmärk on minimiseerida ruutriskifunktsiooni ülemist raja kõigi nende vektorite θ hulgal, mille projektsioon Eukleidilise ruumi R^m alamruumi $R^{k'}$ kuulub antud suletud tõkestatud hulka $\Omega'_k \subset R^{k'}$. Näidatakse, et probleem on lahendatav iteratsiooniprotsessi abil siis, kui suvalise fikseeritud hinnangu korral on võimalik ära näidata riskifunktsiooni maksimumpunkt hulgal Ω'_k. Erijuhtude jaoks on antud probleemi täielik lahendus.

J. KUKS

A MINIMAX ESTIMATOR OF REGRESSION COEFFICIENTS

This paper is concerned with finding an optimum estimator of an m-dimensional vector θ of coefficients in a linear regression model within the family of estimators which may be represented as a sum of a constant vector and a linear vector-statistic. The optimality criterion is to minimize the supremum of a quadratic risk function on a set of all these vectors θ, whose projection on a subspace $R^{k'}$ of the Euclidean space R^m belongs to a given closed bounded set $\Omega'_k \subset R^{k'}$. It is shown that the problem may be solved by means of an iterative minimization process if it is possible, for any fixed estimator, to indicate a point of maximum of the risk function on the set Ω'_k. For some particular cases, the full solution of the problem is given.