

В. ОЛЬМАН

ОЦЕНИВАНИЕ КОЭФФИЦИЕНТОВ ЛИНЕЙНОЙ РЕГРЕССИИ КАК АНТАГОНИСТИЧЕСКАЯ ИГРА

(Представил Н. Алумяэ)

Рассмотрим регрессионную модель

$$Y = X\theta + \varepsilon, \tag{1}$$

где Y — n -мерный наблюдаемый случайный вектор, ε — n -мерный ненаблюдаемый случайный вектор, причем $E(\varepsilon | \theta) = 0$ и $E(\varepsilon \varepsilon^T | \theta) = \sigma^2 I_n$, I_n — единичная матрица порядка n , элементы матрицы X порядка $n \times m$ и σ^2 — известные коэффициенты, а θ — m -мерный вектор, подлежащий оцениванию.

Пусть априори известно, что оцениваемый вектор θ принадлежит m -мерному эллипсоиду $L(A) = \{\theta : \theta^T A \theta \leq 1\}$, где A — известная неотрицательно определенная матрица порядка m . Оценивание вектора θ с помощью статистики $\hat{\theta}(Y)$, с одной стороны, и выбор вектора θ , с другой стороны, можно рассматривать как игру двух лиц [1]: статистики $\hat{\theta}(Y)$ образуют класс стратегий первого игрока, а векторы θ , принадлежащие эллипсоиду $L(A)$, — класс стратегий второго игрока. Ограничим стратегии первого игрока линейными по наблюдениям статистиками, т. е.

$$\hat{\theta}(Y) = TY,$$

где T — произвольная матрица порядка $m \times n$, составленная из неслучайных элементов. За функцию игры примем квадратичный риск

$$r(\theta, T) = E\{(\theta - TY)^T P (\theta - TY)\},$$

где P — заданная неотрицательно определенная матрица порядка m и ранга $R(P) = k \leq m$. Таким образом, при использовании первым игроком матрицы T , а вторым — вектора θ величина $r(\theta, T)$ есть проигрыш первого игрока и выигрыш второго.

В [2-4] рассматривалась минимаксная задача, т. е. поиск матрицы T_0 такой, что

$$\inf_T \sup_{\theta \in L(A)} r(\theta, T) = \sup_{\theta \in L(A)} r(\theta, T_0).$$

В настоящей работе решается задача максимизации гарантированного выигрыша второго игрока, т. е.

$$\inf_T r(\theta, T) \rightarrow \sup_{\theta \in L(A)} \tag{2}$$

Кроме того, выводится условие, при котором игра является вполне определенной и которое позволяет решать минимаксную задачу более простым методом, чем это сделано в [2, 3].

Пусть p_1, \dots, p_k — ортонормированные собственные векторы мат-



рицы \mathbf{P} , соответствующие ненулевым собственным числам $\lambda_1, \dots, \lambda_k$.

Тогда $\mathbf{P} = \sum_{i=1}^k \lambda_i \mathbf{p}_i \mathbf{p}_i^T$, а так как [2]

$$r(\theta, \mathbf{T}) = \sigma^2 \operatorname{tr} \mathbf{T}^T \mathbf{P} \mathbf{T} + \theta^T (\mathbf{I}_m - \mathbf{T} \mathbf{X})^T \mathbf{P} (\mathbf{I}_m - \mathbf{T} \mathbf{X}) \theta,$$

то

$$r(\theta, \mathbf{T}) = \sigma^2 \sum_{i=1}^k \lambda_i \mathbf{p}_i^T \mathbf{T} \mathbf{T}^T \mathbf{p}_i + \sum_{i=1}^k \lambda_i (\mathbf{p}_i^T (\mathbf{I}_m - \mathbf{T} \mathbf{X}) \theta)^2. \quad (3)$$

Дифференцируя по векторам $\mathbf{T}^T \mathbf{p}_i$, $i = 1, 2, \dots, k$, выпуклую по ним функцию $r(\theta, \mathbf{T})$, получаем [5]

$$\frac{\partial r(\theta, \mathbf{T})}{\partial \mathbf{T}^T \mathbf{p}_i} = 2\sigma^2 \lambda_i \mathbf{T}^T \mathbf{p}_i + 2\lambda_i \mathbf{p}_i^T (\mathbf{I}_m - \mathbf{T} \mathbf{X}) \theta (-\mathbf{X} \theta)$$

и, следовательно, векторы, на которых достигается минимальное значение функции $r(\theta, \mathbf{T})$, представимы в виде

$$\mathbf{T}^T \mathbf{p}_i = (\sigma^2 \mathbf{I}_n + \mathbf{X} \theta \theta^T \mathbf{X}^T)^{-1} \mathbf{X} \theta \theta^T \mathbf{p}_i.$$

Так как [5]

$$(\sigma^2 \mathbf{I}_n + \mathbf{X} \theta \theta^T \mathbf{X}^T)^{-1} = 1/\sigma^2 \mathbf{I}_n - 1/\sigma^2 \mathbf{X} \theta \theta^T \mathbf{X}^T / (\sigma^2 + \theta^T \mathbf{X}^T \mathbf{X} \theta),$$

то

$$\mathbf{T}^T \mathbf{p}_i = \mathbf{X} \theta \theta^T \mathbf{p}_i / (\sigma^2 + \theta^T \mathbf{X}^T \mathbf{X} \theta). \quad (4)$$

Подставляя (4) в (3), получаем равенство

$$\inf_{\mathbf{T}} r(\theta, \mathbf{T}) = \sigma^2 \cdot \theta^T \mathbf{X}^T \mathbf{X} \theta \theta^T \mathbf{P} \theta / (\sigma^2 + \theta^T \mathbf{X}^T \mathbf{X} \theta)^2 +$$

$$+ \theta^T (\mathbf{I}_m - \mathbf{X}^T \mathbf{X} \theta \theta^T / (\sigma^2 + \theta^T \mathbf{X}^T \mathbf{X} \theta)) \mathbf{P} (\mathbf{I}_m - \theta \theta^T \mathbf{X}^T \mathbf{X} / (\sigma^2 + \theta^T \mathbf{X}^T \mathbf{X} \theta)) \theta,$$

которое легко преобразуется к равенству

$$\inf_{\mathbf{T}} r(\theta, \mathbf{T}) = \sigma^2 \theta^T \mathbf{P} \theta / (\sigma^2 + \theta^T \mathbf{X}^T \mathbf{X} \theta). \quad (5)$$

Таким образом, решение задачи сводится к поиску

$$\sup_{\theta \in L(\mathbf{A})} \sigma^2 \theta^T \mathbf{P} \theta / (\sigma^2 + \theta^T \mathbf{X}^T \mathbf{X} \theta). \quad (6)$$

Теорема 1. Для того чтобы выражение (6) было конечно, необходимо и достаточно включения

$$\operatorname{Ker} \mathbf{A} \cap \operatorname{Ker} \mathbf{X}^T \mathbf{X} \subseteq \operatorname{Ker} \mathbf{P}, \quad (7)$$

где, как обычно, для произвольной матрицы \mathbf{B} порядка $m \times m$ ядро $\operatorname{Ker} \mathbf{B} = \{\theta \in R^m : \mathbf{B} \theta = 0\}$.

Доказательство. Необходимость легко доказывается от противного. Пусть условие (7) не выполнено, т. е. существует вектор $\theta_0 \in R^m$ такой, что $\mathbf{A} \theta_0 = \mathbf{X}^T \mathbf{X} \theta_0 = 0$ и $\mathbf{P} \theta_0 \neq 0$. Выберем последовательность векторов $\theta_n = n \theta_0$, $n = 1, 2, \dots$. Тогда, очевидно, $\theta_n^T \mathbf{A} \theta_n < 1$ и

$$\theta_n^T \mathbf{P} \theta_n / (\sigma^2 + \theta_n^T \mathbf{X}^T \mathbf{X} \theta_n) = 1/\sigma^2 \cdot n^2 \theta_0^T \mathbf{P} \theta_0 \rightarrow \infty.$$

Докажем достаточность. Заметим, что если \mathbf{B} — ненулевая неотрицательно определенная матрица, то по теореме Куранта—Фишера [6]

$$\inf_{\theta \in \operatorname{Ker}^\perp \mathbf{B} \subset R^m} \theta^T \mathbf{B} \theta / \theta^T \theta = b_1, \quad (8)$$

где b_1 — минимальное положительное собственное число матрицы \mathbf{B} , а $\text{Ker}^\perp \mathbf{B}$ — ортогональное дополнение ядра $\text{Ker} \mathbf{B}$. Таким образом,

$$\begin{aligned} \sup_{\theta \in L(\mathbf{A})} \theta^T \mathbf{P} \theta / (\sigma^2 + \theta^T \mathbf{X}^T \mathbf{X} \theta) &\leq \sup_{\theta \in L(\mathbf{A})} \theta^T \mathbf{P} \theta / \theta^T (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X}) \theta \leq \\ &\leq \sup_{\theta \in R^m} \theta^T \mathbf{P} \theta / \theta^T (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X}) \theta = \\ &= \sup_{\theta_1, \theta_2 \in R^m} (\theta_1 + \theta_2)^T \mathbf{P} (\theta_1 + \theta_2) / (\theta_1 + \theta_2)^T (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X}) (\theta_1 + \theta_2), \end{aligned}$$

где $\theta_1 \in \text{Ker} (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X})$, $\theta_2 \in \text{Ker}^\perp (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X})$. В силу условия (7) теоремы и равенства (8) получаем

$$\begin{aligned} \sup_{\theta \in L(\mathbf{A})} \theta^T \mathbf{P} \theta / (\sigma^2 + \theta^T \mathbf{X}^T \mathbf{X} \theta) &\leq \sup_{\theta_2 \in \text{Ker}^\perp (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X})} \theta_2^T \mathbf{P} \theta_2 / \theta_2^T (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X}) \theta_2 \leq \\ &\leq [\sup_{\theta \in R^m} \theta^T \mathbf{P} \theta / \theta^T \theta] / [\inf_{\theta \in \text{Ker}^\perp (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X})} \theta^T (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X}) \theta / \theta^T \theta] < \infty, \end{aligned}$$

что полностью доказывает теорему.

Теорема 2. Пусть выполнено условие теоремы 1. Тогда

$$\sup_{\theta \in L(\mathbf{A})} \inf_{\mathbf{T}} r(\theta, \mathbf{T}) = \sigma^2 f_1,$$

где f_1 — максимальное из чисел, удовлетворяющих равенству

$$\mathbf{P} \theta = f_1 (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X}) \theta$$

для какого-либо $\theta \in \text{Ker}^\perp (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X})$.

Доказательство. Очевидно, что

$$\begin{aligned} \sup_{\theta \in L(\mathbf{A})} \theta^T \mathbf{P} \theta / (\sigma^2 + \theta^T \mathbf{X}^T \mathbf{X} \theta) &\leq \sup_{\theta \in L(\mathbf{A})} \frac{\theta^T \mathbf{P} \theta}{\theta^T (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X}) \theta} \leq \\ &\leq \sup_{\theta \in R^m} \frac{\theta^T \mathbf{P} \theta}{\theta^T (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X}) \theta} = f_1, \end{aligned} \quad (9)$$

причем последнее равенство реализуется на векторе θ_1 , удовлетворяющем равенству

$$\mathbf{P} \theta_1 = f_1 (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X}) \theta_1, \quad \theta_1 \in \text{Ker}^\perp (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X}).$$

Пусть $\mathbf{A} \theta_1 \neq 0$, тогда можно считать, что $\theta_1^T \mathbf{A} \theta_1 = 1$ и, таким образом, в (9) везде получаем равенства. Пусть теперь $\mathbf{A} \theta_1 = 0$. Выберем $\theta_n = n \theta_1$ и получим

$$\sigma^2 \theta_n^T \mathbf{P} \theta_n / (\sigma^2 + \theta_n^T \mathbf{X}^T \mathbf{X} \theta_n) \rightarrow \theta_1^T \mathbf{P} \theta_1 / \theta_1^T \mathbf{X}^T \mathbf{X} \theta_1 = f_1.$$

Таким образом, в обоих случаях неравенства (9) в действительности являются равенствами, что и доказывает теорему.

Замечание 1. Нетрудно показать, что число f_1 , определенное в теореме, является максимальным собственным числом матрицы $(\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X})^{+1/2} \mathbf{P} (\sigma^2 \mathbf{A} + \mathbf{X}^T \mathbf{X})^{+1/2}$, где $(\cdot)^+$ — обратная матрица Мура [5].

Определим теперь условия, при которых описанная игра является вполне определенной, т. е.

$$\inf_{\mathbf{T}} \sup_{\theta \in L(\mathbf{A})} r(\theta, \mathbf{T}) = \sup_{\theta \in L(\mathbf{A})} \inf_{\mathbf{T}} r(\theta, \mathbf{T}). \quad (10)$$

В частности, в [2] показано, что при $R(\mathbf{P}) = 1$ равенство (10) имеет место для любых матриц \mathbf{A} и $\mathbf{X}^T \mathbf{X}$. Рассмотрим случай, когда $R(\mathbf{P}) \geq 2$.

Для полной определенности игры необходимо и достаточно существования седловой точки [1], т. е. таких стратегий T_0 и $\theta_0 \in L(A)$, что

$$r(\theta, T_0) \leq r(\theta_0, T_0) \leq r(\theta_0, T) \quad \forall T, \theta \in L(A). \quad (11)$$

Из правого неравенства (11) получаем, что $T_0 = \theta_0 \theta_0^T X^T / (\sigma^2 + \theta_0^T X^T X \theta_0)$, а из левого неравенства — $\theta_0 = \theta_1$, где вектор θ_1 из теоремы 2. Таким образом, если игра определена, то стратегии $T_1 = \theta_1 \theta_1^T X^T / (\sigma^2 + \theta_1^T X^T X \theta_1)$ и θ_1 образуют седловую точку. Используя равенство $P\theta_1 = f_1(\sigma^2 A + X^T X)\theta_1$, получаем

$$\begin{aligned} r(\theta, T_1) &= \sigma^2 \theta_1^T X^T X \theta_1 \theta_1^T P \theta_1 / (\sigma^2 + \theta_1^T X^T X \theta_1)^2 + \theta^T [P - 2X^T X \theta_1 \theta_1^T P \times \\ &\quad \times (\sigma^2 + \theta_1^T X^T X \theta_1)^{-1} + \theta_1^T P \theta_1 \cdot X^T X \theta_1 \theta_1^T X^T X / (\sigma^2 + \theta_1^T X^T X \theta_1)^2] \theta = \\ &= \sigma^2 \cdot f_1^2 \theta_1^T X^T X \theta_1 / \theta_1^T P \theta_1 + \theta^T B \theta, \end{aligned}$$

где $B = P + (f_1^2 \sigma^4 A \theta_1 \theta_1^T A - P \theta_1 \theta_1^T P) / \theta_1^T P \theta_1$. Отсюда следует

Теорема 3. Для полной определенности игры необходимо и достаточно, чтобы

$$\sup_{\theta \in L(A)} \theta^T B \theta = \theta_1^T B \theta_1. \quad (12)$$

Замечание 2. Условие (12) эквивалентно тому, что вектор $A^{1/2} \theta_1$ является собственным вектором матрицы $A^{1/2} B A^{1/2}$, соответствующим максимальному собственному числу этой матрицы.

Замечание 3. Условие $A\theta_1 \neq 0$ необходимо для полной определенности игры, так как в противном случае $\theta_1^T B \theta_1 = 0$, что противоречит (12).

Замечание 4. Условие $\text{Ker } A \subseteq \text{Ker } P$ необходимо для полной определенности игры, так как в противном случае $\sup_{\theta \in L(A)} \theta^T B \theta = \infty$, что противоречит (12).

Более простым выглядит решение задачи об определенности игры при условии, сформулированном в [3], т. е. при существовании невырожденной матрицы C такой, что матрицы $\bar{A} = C^T A C$, $\bar{P} = C^T P C$, $\bar{X}^T X = C^T X^T X C$ являются диагональными с диагональными элементами $a_1^2, \dots, a_m^2, p_1^2, \dots, p_m^2, x_1^2, \dots, x_m^2$ соответственно. Очевидно, что задача с матрицами A, P и $X^T X$ эквивалентна задаче с матрицами \bar{A}, \bar{P} и $\bar{X}^T X$. Пусть для определенности

$$\max_{1 \leq i \leq m} p_i^2 / (\sigma^2 a_i^2 + x_i^2) = p_1^2 / (\sigma^2 a_1^2 + x_1^2) = f_1.$$

Тогда у вектора θ_1 первая компонента $1/a_1$ (см. замечание 3), остальные компоненты — нули, а сам вектор θ_1 удовлетворяет уравнению

$$\bar{P}\theta_1 = f_1(\sigma^2 \bar{A} + \bar{X}^T X)\theta_1.$$

Очевидно, что справедливость условия (12) надо проверять лишь на векторах $\theta \in \text{Ker } \bar{A}$ (см. замечание 4), удовлетворяющих условиям

$$\bar{B}\theta = f\bar{A}\theta, \quad \theta^T \bar{A}\theta = 1$$

при некотором $f \in R^1$. Число таких векторов совпадает с числом ненулевых диагональных элементов матрицы \bar{A} , и векторы имеют вид: $1/a_j$ ($a_j \neq 0, m \geq j \geq 1$) на j -м месте, а на остальных местах — нули. Тогда условие (12) эквивалентно тому, что при $a_j \neq 0$

$$f_1 \sigma^2 \geq |p_1 p_j / (a_1 a_j)|.$$

ЛИТЕРАТУРА

1. Wald, A. Statistical Decision Functions. Wiley, New York, 1950.
2. Кукс Я., Ольман В. Изв. АН ЭССР, Физ. Матем., 21, № 1, 66—72 (1972).
3. Кукс Я., Ольман В. Изв. АН ЭССР, Физ. Матем., 20, № 4, 480—482 (1971).
4. Rao, C. R. App. Statist., 4, № 6, 1023—1037, (1976).
5. Рао С. Р. Линейные статистические методы и их применения. М., «Наука», 1968.
6. Беллман Р. Введение в теорию матриц. М., «Наука», 1969.

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
12 октября 1982

V. OLMAN

LINEAARSE REGRESSIOONI KORDAJATE HINDAMINE KUI ANTAGONISTLIK MÄNG

Lineaarse regressioonimudeli kordajate hindamise ülesanne on formuleeritud mängu-teooria terminites. On antud sadulpunkti olemasolu tingimused, mis võimaldab mõnel juhul analüütiliselt leida esialgse minimaksülesande lahendi.

V. OLMAN

ESTIMATION OF LINEAR REGRESSION COEFFICIENTS AS AN ANTAGONISTIC GAME

Linear estimation of a parameter $\theta \in R^m$ of a linear regression model $EY = X\theta$ is considered as an antagonistic game with the utility-loss function $r(\theta, T) = E\{(\theta - TY)^T P(\theta - TY)/\theta\}$. The maxmin strategy in explicit form in case if an unknown parameter θ belongs to the fixed ellipsoid $\{\theta \in R^m : \theta^T A \theta \leq 1\}$ is given. Conditions for the existence of a saddle point are obtained that provide the solution of the minmax estimation problem. The solution of minmax estimation problem in case if $X^T X$, A and P have the same eigenvectors, is given as an example.