EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 28. KÖIDE
FÜÜSIKA * MATEMAATIKA. 1979, NR. 4

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 28
ФИЗИКА * МАТЕМАТИКА. 1979, № 4

https://doi.org/10.3176/phys.math.1979.4.13

УДК 519.22

V. VAATMANN

MINIMAX ESTIMATION OF RANDOM VECTORS

V. VAATMANN. JUHUSLIKE VEKTORITE MINIMAKSHINDAMINE

В. ВААТМАНН. МИНИМАКСНОЕ ОЦЕНИВАНИЕ СЛУЧАЙНЫХ ВЕКТОРОВ

(Presented by H. Aben)

1. Let $\xi=(\xi_1, \xi_2, \ldots, \xi_n)^T$ be an observable and $\eta=(\eta_1, \eta_2, \ldots, \eta_m)^T$ a non-observable real-valued random vector. Denote by $\mathbf{E}$ the operator of mathematical expectation.
Let

$$m_\xi=\mathbf{E}\xi, \quad m_\eta=\mathbf{E}\eta, \quad D_\xi=\mathbf{E}[(\xi-m_\xi)(\xi-m_\xi)^T],$$

$$D_{\eta\xi}=\mathbf{E}[(\eta-m_\eta)(\xi-m_\xi)^T], \qquad D_\eta=\mathbf{E}[(\eta-m_\eta)(\eta-m_\eta)^T].$$

Let $R_{m\times n}$ be the space of all $m\times n$-matrices and R_m the space of all m-vectors with norms $\|A\|=[\mathrm{tr}(AA^T)]^{1/2}$ and $\|a\|=(a^Ta)^{1/2}$, respectively.

We shall consider the problem of the linear estimation of the vector η by ξ so that the mean square error will be minimized. If m_η, m_ξ, $D_{\eta\xi}$ and D_ξ are known, the result may be found, e. g. in [1].

In the present paper the case is considered when we do not know m_η and $D_{\eta\xi}$ exactly, but a set $K\subset R_{m\times n}$ and a set $S\subset R_m$ are given such that $D_{\eta\xi}\in K$, $m_\eta\in S$. In this case we determine the linear minimax estimate for η.

Definition. *The linear estimate $\hat{\eta}=L_0\xi+c_0$ $(L_0\in R_{m\times n},\ c_0\in R_m)$ is called the minimax estimate with respect to the sets $K\subset R_{m\times n}$, $S\subset R_m$ if*

$$\max_{D_{\eta\xi}\in K, m_\eta\in S} \mathbf{E}[(L\xi+c-\eta)^T(L\xi+c-\eta)]\big|_{L=L_0,\ c=c_0}=$$

$$=\min_{L\in R_{m\times n},\, c\in R_m}\ \max_{D_{\eta\xi}\in K, m_\eta\in S} \mathbf{E}[(L\xi+c-\eta)^T(L\xi+c-\eta)]. \qquad (1)$$

A set $S_a\subset R_m$ is called symmetrical with respect to $a\in R_m$ if $S_a=S+a$, where $S\subset R_m$ is a symmetrical set with respect to all co-ordinate axes.

In the following we shall suppose for simplicity that $m_\xi=0$.

2. Let us formulate the main result of this paper.

Theorem. *Let $K\subset R_{m\times n}$ and $S_a\subset R_m$ be closed convex bounded sets and let S_a be symmetrical with respect to $a\in R_m$. Then $\hat{\eta}=L_0\xi+c_0$ is the linear minimax estimate for η if $c_0=a$ and $L_0=D^0_{\eta\xi}D^+_\xi$, where*

$$\operatorname{tr}(D^0_{\eta\xi} D^+_{\xi} D^0_{\xi\eta}) = \min_{D_{\eta\xi} \in K} \operatorname{tr}(D_{\eta\xi} D^+_{\xi} D_{\xi\eta})$$

and D^+_{ξ} denotes the pseudo-inverse of D_{ξ}.

P r o o f. Transform the expression on the right side of (1).

$$\begin{aligned} &\min_{L\in R_{m\times n},\, c\in R_m} \max_{D_{\eta\xi}\in K,\, m_\eta\in S_a} \mathbf{E}[(L\xi+c-\eta)^T(L\xi+c-\eta)] = \\ &= \min_{L\in R_{m\times n},\, c\in R_m} \max_{D_{\eta\xi}\in K,\, m_\eta\in S_a} \operatorname{tr} \mathbf{E}[(L\xi+c-\eta)(L\xi+c-\eta)^T] = \\ &= \min_{L\in R_{m\times n}} \max_{D_{\eta\xi}\in K} \operatorname{tr}(LD_\xi L^T - 2D_{\eta\xi}L^T + D_\eta) + \\ &+ \min_{c\in R_m} \max_{m_\eta\in S_a} [(c-m_\eta)^T(c-m_\eta)]. \end{aligned} \tag{2}$$

It is evident that the function on the right side of (2) is convex in $L \in R_{m\times n}$ and concave in $D_{\eta\xi} \in K$. Using Theorem 37.3 [2] and the fact that the rank of $D_{\eta\xi}$ cannot be greater than the rank of D_ξ, we can rewrite (2) as

$$\begin{aligned} &\min_{L\in R_{m\times n}} \max_{D_{\eta\xi}\in K} \operatorname{tr}(LD_\xi L^T - 2D_{\eta\xi}L^T + D_\eta) + \min_{c\in R_m} \max_{m_\eta\in S_a} [(c-m_\eta)^T(c-m_\eta)] = \\ &= \max_{D_{\eta\xi}\in K} \min_{L\in R_{m\times n}} \operatorname{tr}(LD_\xi L^T - 2D_{\eta\xi}L^T + D_\eta) + \min_{c\in R_m} \max_{m_\eta\in S_a} [(c-m_\eta)^T(c-m_\eta)] = \\ &= \max_{D_{\eta\xi}\in K} \min_{L\in R_{m\times n}} \operatorname{tr}[D_\eta - D_{\eta\xi}D^+_\xi D_{\xi\eta} + (LD_\xi - D_{\eta\xi})D^+_\xi(LD_\xi - D_{\eta\xi})^T] + \\ &+ \min_{c\in R_m} \max_{m_\eta\in S_a} [(c-m_\eta)^T(c-m_\eta)]. \end{aligned} \tag{3}$$

As $D^+_\xi \geqslant 0$, it follows that the minimax matrix L_0 is equal to $D^0_{\eta\xi} D^+_\xi$, where $D^0_{\eta\xi}$ minimizes $\operatorname{tr}(D_{\eta\xi}D^+_\xi D_{\xi\eta})$ subject to $D_{\eta\xi} \in K$. The corresponding minimum exists according to Theorem 2.1.2 [3].

It remains to prove $c_0 = a$. Using Theorem 32.3 [2], we have

$$\min_{c\in R_m} \max_{m_\eta\in S_a} [(c-m_\eta)^T(c-m_\eta)] = \min_{c\in R_m} [(c-m_\eta(c))^T(c-m_\eta(c))],$$

where $\operatorname{sgn}(m_\eta(c) - a) = -\operatorname{sgn}(c-a)$, $m_\eta(c)$ belongs to the boundary of S_a and $\operatorname{sgn}(b_1, b_2, \ldots, b_m)^T = (\operatorname{sgn} b_1, \operatorname{sgn} b_2, \ldots, \operatorname{sgn} b_m)^T$. Therefore $c_0 = a$. This completes the proof.

R e m a r k s:

1. It follows from the proof that the maximum mean square error of the linear minimax estimate $\hat{\eta} = L_0\xi + c_0$ is

$$\begin{aligned} &\max_{D_{\eta\xi}\in K,\, m_\eta\in S_a} \mathbf{E}[(L_0\xi+c_0-\eta)^T(L_0\xi+c_0-\eta)] = \\ &= \operatorname{tr}(D_\eta - D^0_{\eta\xi}D^+_\xi D^0_{\xi\eta}) + \max_{m_\eta\in S_a} [(a-m_\eta)^T(a-m_\eta)]. \end{aligned}$$

2. If m_ξ is known and not zero, we estimate η by $\xi - m_\xi$ and the minimax linear estimate is equal to $D_{\eta\xi} D_\xi^{+} (\xi - m_\xi) + a$.
3. If D_η is known, the Cramer—Rao inequality yields $D_{\eta\xi} \in K_1$, where $K_1 = \{A \in R_{m\times n} : A D_\xi^{+} A^T \leqslant D_\eta\}$. In this case we can take $K \cap K_1$ instead of K.
4. If η and ξ are two jointly Gaussian random vectors, then the linear minimax estimate for η by ξ is also the minimax estimate and equals $\mathbf{E}_m(\eta|\xi)$, where $\mathbf{E}_m$ is the conditional expectation with respect to the least favorable distribution of η and ξ, i. e.

$$\min_{g\in G} \max_{D_{\eta\xi}\in K,\, m_\eta \in S_a} \mathbf{E}[(g(\xi)-\eta)^T (g(\xi)-\eta)] =$$

$$= \min_{g\in G} \mathbf{E}_m[(g(\xi)-\eta)^T(g(\xi)-\eta)] = \mathbf{E}_m[(L_0\xi + c_0 - \eta)^T (L_0\xi + c_0 - \eta)],$$

where G is the set of all measurable functions.

3. Now let us consider a simple example to illustrate the results given above.

We consider the problem of the linear estimation of the m-vector η in the model

$$\xi = A\eta + \varepsilon, \tag{4}$$

where ξ is an n-vector of observations, A is an $n \times m$-matrix of known elements, ε is an n-vector of the noise, dependent on η.

Let m_η, D_η, D_ε be known and let $m_\varepsilon = 0$ $\|D_{\eta\varepsilon}\|^2 \leqslant c^2$, where c is a known scalar such that $c^2 < \|D_\eta A^T\|^2$.
We get

$$\|D_{\eta\xi} - D_\eta A^T\|^2 = \|D_{\eta\varepsilon}\|^2 \leqslant c^2.$$

It can be shown that the function $\operatorname{tr}(D_{\eta\xi} D_\xi^{+} D_{\xi\eta})$ is minimized subject to $\|D_{\eta\xi} - D_\eta A^T\|^2 \leqslant c^2$ by

$$\overset{0}{D}_{\eta\xi} = [\operatorname{tr}(D_\eta A^T D_\xi^{+} A D_\eta)]^{-1/2} \{[\operatorname{tr}(D_\eta A^T D_\xi^{+} A D_\eta)]^{1/2} - c[\operatorname{tr} D_\xi^{+}]^{1/2}\} D_\eta A^T. \tag{5}$$

Therefore the linear minimax estimate for η by ξ in the model (4) is $\hat{\eta} = \overset{0}{D}_{\eta\xi} D_\xi^{+} \xi + (I - \overset{0}{D}_{\eta\xi} D_\xi^{+} A) m_\eta$, where $\overset{0}{D}_{\eta\xi}$ is determined by (5), $D_\xi = A D_\eta A^T + D_\varepsilon$ and I is the identity matrix.

If $c^2 \geqslant \|D_\eta A^T\|^2$, then $\overset{0}{D}_{\eta\xi} = 0$ and $\hat{\eta} = m_\eta$.

REFERENCES

1. Липцер Р. Ш., Ширяев А. Н., Статистика случайных процессов, М., «Наука», 1974.
2. Рокафеллар Р., Выпуклый анализ, М., «Мир», 1973.
3. Лоран П.-Ж., Аппроксимация и оптимизация, М., «Мир», 1975.

Academy of Sciences of the Estonian SSR,
Institute of Cybernetics

Received
June 21, 1979