On Minimax Estimating Hilbert Random Elements*

Konstantin Siemenikhin

Probability Theory Dep., Moscow Aviation Inst.

The problem of designing optimal algorithms for estimating random elements has received earlier considerable attention (Balakrishnan, 1976; Curtain and Pritchard, 1978; Ramm, 1996). Basically, these works deal with linear procedures and do not discuss the efficiency of nonlinear estimates. On the other hand, the nonlinear optimal estimation algorithms are essentially based on using the true distribution of the random elements involved. Nevertheless, these obstacles of optimal methods can be overcome by means a minimax approach (Verdú and Poor, 1984; Başar and Bernhard, 1991; Siemenikhin, 2003). Actually, even though the class of estimators contains all nonlinear transformations, the minimax estimate turns to be linear under very broad assumptions (Siemenikhin, 2003; Siemenikhin and Lebedev, 2004). In the finite-dimensional case, this result holds whenever the covariances of the model parameters are supposed to belong to a compact set. Furthermore, it turns out that for the uncertainty set under consideration the least favorable distribution is Gaussian. In this paper, the analogous results are proved for the infinite-dimensional model.

Using the technique of dual optimization we provide the sufficient conditions for the minimax estimate to be defined analytically via a solution of the dual optimization problem. Thus, if the least favorable covariance (i.e., the solution of the dual problem) is found, the minimax estimate should be designed as the optimal one. For numerical calculation of the minimax estimate we present the recursive algorithm which takes into account only finite-dimensional transformations of the observed random element.

Let X and Y be separable Hilbert spaces. Consider the problem of minimax estimating the random element $\xi \in X$ from the observable random element $\eta \in Y$ w.r.t. the mean-square criterion

$$\mathfrak{D}(\mathbf{F}, \mathsf{P}_\zeta) = \mathsf{E}\|\xi - \mathbf{F}\eta\|^2,$$

where $\mathbf{F}$ is an admissible estimation procedure and P_ζ is the distribution of the element $\zeta = (\xi, \eta) \in Z = X \times Y$.

Concerning ξ and η we have the following *a priori* information:

$$\mathsf{E}\|\xi\|^2 < \infty, \quad \mathsf{E}\xi = 0, \quad \mathsf{E}\|\eta\|^2 < \infty, \quad \mathsf{E}\eta = 0,$$

$$K_\zeta = \mathsf{cov}\{\zeta, \zeta\} = \begin{pmatrix} K_\xi & K_{\xi\eta} \\ K_{\eta\xi} & K_\eta \end{pmatrix} \in \mathcal{K},$$

*The work is supported in part by Russian Foundation for Basic Research under Grant No 05-01-00508

where $\mathcal{K}$ is a given set of selfadjoint positively semidefinite nuclear operators defined on Z. By $\mathcal{P}$ denote the family of all distributions P_ζ described above.

Definition 1. We say the X-valued random element $\widetilde{\xi}$ is an *admissible estimate*, if there exists a sequence of cylindrical mappings $\mathbf{F}_n\colon Y \to X$ such that for every $\mathsf{P}_\zeta \in \mathcal{P}$

$$\mathsf{E}\|\mathbf{F}_n\eta\|^2 < \infty, \quad \mathsf{E}\|\mathbf{F}_n\eta - \widetilde{\xi}\|^2 \to 0, \quad n \to \infty. \tag{1}$$

In this case we use the notation $\widetilde{\xi} = \mathbf{F}\eta$ and call the sequence $\mathbf{F} = \{\mathbf{F}_n\}$ the *admissible estimation procedure* (briefly $\mathbf{F} \in \mathfrak{F}$).

Definition 2. The admissible estimation procedure $\mathbf{F} = \{\mathbf{F}_n\}$ and corresponding estimate $\widetilde{\xi} = \mathbf{F}\eta$ are called *linear* if in (1) $\{\mathbf{F}_n\}$ are linear mappings. In this case we write $\mathbf{F} \in \mathfrak{L}$.

Definition 3. The admissible estimation procedure $\widehat{\mathbf{F}} = \{\widehat{\mathbf{F}}_n\}$ and corresponding estimate $\hat{\xi} = \widehat{\mathbf{F}}\eta$ are called minimax if

$$\widehat{\mathbf{F}} \in \arg\min_{\mathbf{F}\in\mathfrak{F}} \sup_{\mathsf{P}_\zeta\in\mathcal{P}} \mathfrak{D}(\mathbf{F}, \mathsf{P}_\zeta). \tag{2}$$

To formulate our main result we introduce the dual problem

$$\hat{K}_\zeta \in \arg\max_{K_\zeta\in\mathcal{K}} J(K_\zeta), \quad J(K_\zeta) = \inf_{\mathbf{F}\in\mathfrak{L}} \mathfrak{D}(\mathbf{F}, \mathsf{P}_\zeta). \tag{3}$$

The solution to the minimax estimation problem (2) is provided by the following theorem.

Theorem 1. *Suppose that $\mathcal{K}$ is convex, there exists a solution $\hat{K}_\zeta$ of the dual problem* (3), *and*

$$\exists\, c > 0\colon \quad \langle K_\eta y, y\rangle \le c^2 \langle \hat{K}_\eta y, y\rangle \quad \forall\, K_\zeta \in \mathcal{K} \quad \forall\, y \in Y. \tag{4}$$

If
a) $\widehat{\mathsf{P}}_\zeta$ *is the Gaussian distribution with zero mean and covariance* $\hat{K}_\zeta$;
b) $\widehat{\mathbf{F}}$ *is the best linear admissible estimation procedure given* $\mathsf{P}_\zeta = \widehat{\mathsf{P}}_\zeta$:

$$\widehat{\mathbf{F}} \in \arg\min_{\mathbf{F}\in\mathfrak{L}} \mathfrak{D}(\mathbf{F}, \widehat{\mathsf{P}}_\zeta),$$

then,
the pair $(\widehat{\mathbf{F}}, \widehat{\mathsf{P}}_\zeta)$ *forms a saddle point for the game* $(\mathfrak{D}, \mathfrak{F}, \mathcal{P})$:

$$\mathfrak{D}(\widehat{\mathbf{F}}, \mathsf{P}_\zeta) \le \mathfrak{D}(\widehat{\mathbf{F}}, \widehat{\mathsf{P}}_\zeta) \le \mathfrak{D}(\mathbf{F}, \widehat{\mathsf{P}}_\zeta) \quad \forall\, \mathbf{F} \in \mathfrak{F} \quad \forall\, \mathsf{P}_\zeta \in \mathcal{P}.$$

In addition, $\mathfrak{D}(\widehat{\mathbf{F}}, \widehat{\mathsf{P}}_\zeta) = J(\hat{K}_\zeta)$.

So, the linear estimate $\hat{\xi} = \widehat{\mathbf{F}}\eta$ is minimax on the class of all admissible estimates and the Gaussian law $\widehat{\mathsf{P}}_\zeta$ is the least favorable distribution on the family $\mathcal{P}$.

Condition (4) is analogous to that known to be the regularity condition in the finite-dimensional case. Actually, if $\dim X < \infty$, $\dim Y < \infty$, and $\mathcal{K}$ is bounded, then (4) is equivalent to the embedding $\mathrm{im}[K_\eta] \subseteq \mathrm{im}[\hat{K}_\eta]\ \forall\, K_\zeta \in \mathcal{K}$.

The iterative algorithm generating the minimax estimate is presented below.

Theorem 2. *Under the conditions of Theorem* 1, *we assume that the sequence* $\{g_n\} \subset Y$ *is a complete orthonormal system in* Y *w.r.t. the norm* $\|y\|_- = \langle \hat{K}_\eta y, y\rangle^{1/2}$. *Then the sequence*

$$\hat{\xi}_0 = 0, \quad \hat{\xi}_n = \hat{\xi}_{n-1} + \langle \eta, g_n\rangle \, \hat{K}_{\xi\eta} g_n, \quad n = 1, 2, \ldots,$$

converges to the minimax estimate $\hat{\xi}$ *whenever* $\mathsf{P}_\zeta \in \mathcal{P}$, *namely,*

$$\sup_{\mathsf{P}_\zeta \in \mathcal{P}} \mathsf{E}\|\hat{\xi}_n - \hat{\xi}\|^2 \leq c^2 \varepsilon_n^2, \quad \varepsilon_n^2 = \sum_{k>n} \|\hat{K}_{\xi\eta} g_k\|^2 \to 0, \quad n \to \infty;$$

$$\sup_{\mathsf{P}_\zeta \in \mathcal{P}} \mathsf{E}\|\xi - \hat{\xi}_n\|^2 \leq \Big(J(\hat{K}_\zeta)^{1/2} + c\,\varepsilon_n \Big)^2 .$$

References

Başar, T. and P. Bernhard (1991). *H_∞-optimal control and related minimax design problem. A game theoretic approach.* Birkhäuser. Boston.

Balakrishnan, A. V. (1976). *Applied functional analysis.* Springer-Verlag. N.Y.

Curtain, R. F. and A. J. Pritchard (1978). *Infinite dimensional linear system theory.* Springer-Verlag. N.Y.

Ramm, A. G. (1996). *Random fields estimation theory.* J. Wiley & Sons. N.Y.

Siemenikhin, K. V. (2003). Minimax estimation of random elements by the root-mean-square criterion. *Journal of Computer and Systems Sciences International* **42**(5), 670–682.

Siemenikhin, K. V. and M. V. Lebedev (2004). Minimax estimation of random elements: Theory and applications. In: *Proc. 43rd IEEE Conf. on Decision and Control (CDC'2004).* Bahamas, Nassau. pp. 3581–3586.

Verdú, S. and H. V. Poor (1984). On minimax robustness: A general approach and applications. *IEEE Trans. Inform. Theory* **IT-30**(2), 328–340.