ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
21,1 (1991), pp. 33–42

A. GRZYBOWSKI (Częstochowa)

MINIMAX STATE ESTIMATION FOR LINEAR STOCHASTIC SYSTEMS WITH AN UNCERTAIN PARAMETER

0. Introduction. Various papers are devoted to the problems of Bayes and/or minimax state estimation for stochastic systems (see [1], [3]–[7], [10]). The estimates have been derived under different assumptions about the uncertainty of the system. Major types of uncertainty are connected with the statistical characterization of the process and observation disturbances. The problem of state estimation for systems with uncertain first or second order statistics of the disturbances have been studied e.g. in [1] (the Bayes approach), and [3]–[7], [10] (the minimax approach).

In our paper we deal with the problem of state estimation for discrete time stochastic systems with an additive, time-invariant, random parameter. The parameter may have different values for different realizations of our process but its value is the same at every moment through a fixed realization. Various practical processes have got such a parameter. For instance some devices have parameters which are being stabilized when the device is being switched on, but every time we switch the device on the parameters can be stabilized on a different level. Another example is when the random parameter changes "very slowly" in comparison with a single realization of the controlled process. Then sometimes we can assume that for a fixed realization the parameter has the same value at every moment.

In the sequel we will assume that the parameter has an unknown distribution belonging to a given class of distributions. As usual the class will be called an *uncertainty class*.

Systems with random, time-invariant parameters are discussed in various papers. The Bayes approach to the problem of parameter estimation for such systems is presented in [8].

1985 *Mathematics Subject Classification*: 93E10.
Key words and phrases: state estimation, Bayes estimation, discrete time systems.

The minimax estimation problem discussed in our paper involves Bayesian estimation and statistical decision theory.

1. Preliminary remarks and notations. Throughout the paper, Greek letters indicate matrices, bold letters indicate vectors. We shall also use the following notations:

$\mathbf{o} =$ the zero vector of appropriate dimension,
α^T $(\mathbf{a}^T) =$ the transpose of a matrix α (a vector $\mathbf{a}$),
$\|\mathbf{a}\|^2 =$ $\mathbf{a}^T\mathbf{a}$ for every vector $\mathbf{a}$,
$\operatorname{tr}\alpha =$ the trace of the matrix α,
$\mathfrak{N}(\mathbf{m}, \Sigma) =$ the Gaussian distribution with the mean vector $\mathbf{m}$ and the covariance matrix Σ,
$\mathcal{P}_{\mathbf{X}|\mathbf{Y}} =$ the conditional distribution of a random vector $\mathbf{X}$ given a random vector $\mathbf{Y}$,
$f_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) =$ the density function of the distribution $\mathcal{P}_{\mathbf{X}|\mathbf{Y}}$.

Let $\alpha_n, \alpha_{n-1}, \ldots, \alpha_0$ be an arbitrary sequence of square matrices having the same dimensions. We denote the product $\alpha_n\alpha_{n-1}\ldots\alpha_k$ by $\alpha_{n,k}$. If $k > n$ then $\alpha_{n,k}$ is the identity matrix of an appropriate dimension.

Let $\mathbf{X}$ and $\mathbf{Y}$ be random vectors of known dimensions. Suppose that the vectors depend on a random parameter V. Let $E(\cdot \mid V = v)$ denote the conditional expectation with respect to (w.r.t.) the distributions of the vectors $\mathbf{X}$ and $\mathbf{Y}$ when the parameter V is known to be equal to v.

Let $\mathbf{d}(\mathbf{Y})$ be an estimate of $\mathbf{X}$ based on $\mathbf{Y}$. The *risk function* $R(v, \mathbf{d})$ connected with the estimate $\mathbf{d}$ is defined by

$$R(v, \mathbf{d}) = E(\|\mathbf{X} - \mathbf{d}\|^2 | V = v).$$

Let the parameter V have distribution $\mathcal{D}$. In this case we define the *Bayes risk* as usual, i.e.

$$r(\mathcal{D}, \mathbf{d}) = E_{\mathcal{D}}R(v, \mathbf{d}) = E\|\mathbf{X} - \mathbf{d}\|^2. \tag{1}$$

Denote the set of estimates $\mathbf{d}$ for which the Bayes risk $r(\mathcal{D}, \mathbf{d})$ exists by $\Delta_{\mathcal{D}}$.

The estimate $\widetilde{\mathbf{d}}$ which satisfies the condition

$$r(\mathcal{D}, \widetilde{\mathbf{d}}) = \inf_{\mathbf{d}\in\Delta_{\mathcal{D}}} r(\mathcal{D}, \mathbf{d})$$

is called a *Bayes estimate* (w.r.t. $\mathcal{D}$).

Let $\mathcal{G}$ be an uncertainty class in our problem and let $\Delta_{\mathcal{G}}$ denote the set of estimates $\mathbf{d}$ for which the Bayes risk $r(\mathcal{D}, \mathbf{d})$ exists for each $\mathcal{D} \in \mathcal{G}$.

The estimate $\mathbf{d}^*$ which satisfies the condition

$$\sup_{\mathcal{D}\in\mathcal{G}} r(\mathcal{D}, \mathbf{d}^*) = \inf_{\mathbf{d}\in\Delta_{\mathcal{G}}} \sup_{\mathcal{D}\in\mathcal{G}} r(\mathcal{D}, \mathbf{d})$$

is called a *$\mathcal{G}$-minimax estimate*.

The following lemma shows the relation between the Bayes and minimax estimates and is often used in order to prove that an estimate is minimax.

LEMMA. *Let* $\{\mathcal{D}_k\}_{k=1}^{\infty}$, $\mathcal{D}_k \in \mathcal{G}$, *be a sequence of distributions of* V *and let* $\{\mathbf{d}_k\}_{k=1}^{\infty}$ *and* $\{r(\mathcal{D}_k, \mathbf{d}_k)\}_{k=1}^{\infty}$ *be the corresponding sequences of Bayes estimates and Bayes risks. If* $\mathbf{d}^*$ *is an estimate for which the Bayes risk* $r(\mathcal{D}, \mathbf{d}^*)$ *satisfies*

$$\sup_{\mathcal{D}\in\mathcal{G}} r(\mathcal{D}, \mathbf{d}^*) \leq \limsup_{k\to\infty} r(\mathcal{D}_k, \mathbf{d}_k) \tag{2}$$

then $\mathbf{d}^*$ *is a* $\mathcal{G}$*-minimax estimate.*

The lemma is a slight generalization of Theorem 6.5.2 in [11]. For control policies such a version of the theorem can be found in [9].

The following corollary is also well known.

COROLLARY. *If the estimate* $\widehat{\mathbf{d}}$ *is Bayes w.r.t. some distribution belonging to* $\mathcal{G}$ *and satisfies*

$$\forall \mathcal{D} \in \mathcal{G} \quad r(\mathcal{D}, \widehat{\mathbf{d}}) = \text{const}$$

then it is a $\mathcal{G}$*-minimax estimate.*

2. Description of models. In the sequel we shall consider the following two models.

The model A is described by the plant equations

$$\mathbf{X}_{n+1} = \alpha_n \mathbf{X}_n + \mathbf{w}_n V, \quad n = 0, 1, \ldots, \tag{3}$$

and the observations

$$\mathbf{Y}_n = \beta_n \mathbf{X}_n + \mathbf{Z}_n, \quad n = 0, 1, \ldots \tag{4}$$

The model B is given by

$$\mathbf{X}_{n+1} = \alpha_n \mathbf{X}_n + \mathbf{Z}_n, \quad n = 0, 1, \ldots,$$
$$\mathbf{Y}_n = \beta_n \mathbf{X}_n + \mathbf{w}_n V, \quad n = 0, 1, \ldots$$

In the above equations the vectors $\mathbf{X}_n$, $\mathbf{Y}_n$ have dimensions p, q respectively, the matrices α_n, β_n and vectors $\mathbf{w}_n$, $\mathbf{Z}_n$ have appropriate dimensions, V is a random, real parameter.

We assume that the initial state $\mathbf{X}_0$ has distribution $\mathfrak{N}(\mathbf{m}_0, \Lambda_0)$ and that, for every n, $\mathbf{Z}_n$ has distribution $\mathfrak{N}(\mathbf{o}, \Sigma_n)$. We also assume that V, $\mathbf{X}_0$, $\mathbf{Z}_n$, $n = 0, 1, \ldots$, are independent.

In the model B we assume that $p = q$ and that the matrices β_n, $n = 0, 1, \ldots$, are nonsingular.

3. Statement of problems. The uncertainty classes $\mathcal{G}_1$ and $\mathcal{G}_2$ which will be considered in our paper are defined as follows:

$$\mathcal{D} \in \mathcal{G}_1 \Leftrightarrow E_{\mathcal{D}} V^2 \leq a,$$

$$\mathcal{D} \in \mathcal{G}_2 \Leftrightarrow E_{\mathcal{D}} V = m_1 \wedge E_{\mathcal{D}} V^2 = m_2 ,$$

where the constants a, m_1, m_2 are known.

Problem A.1. Let the nature's choice of the distribution of the parameter V be confined to the class of distributions $\mathcal{G}_1$. For the model A find the minimax estimate for $\mathbf{X}_{n+1}$ based on the observation history $\mathbf{Y}^n = (\mathbf{Y}_0, \mathbf{Y}_1, \ldots, \mathbf{Y}_n)$ against the nature's choice of distribution in the above class.

Problem A.2. Let the nature's choice of the distribution of the parameter V be confined to the class of distributions $\mathcal{G}_2$. For the model A find the minimax estimate for $\mathbf{X}_{n+1}$ based on the observation history $\mathbf{Y}^n = (\mathbf{Y}_0, \mathbf{Y}_1, \ldots, \mathbf{Y}_n)$ against the nature's choice of distribution in the above class.

Problems B.1 and B.2 concern the model B and are similarly formulated.

4. Bayesian estimation. Assume that the parameter V has distribution $\mathfrak{N}(rs^{-1}, s^{-1})$, where the constants r, s are known, $s > 0$. This distribution will be denoted by $\mathcal{D}_{r,s}$. We shall find the Bayes estimates w.r.t. such distributions and their risks.

Under the given assumptions the problem of estimating $\mathbf{X}_{n+1}$ given $\mathbf{Y}^n$ is a standard Bayesian estimating problem. It is well known (see e.g. [2]) that the minimum mean squared error estimate is the expected value of the conditional distribution of $\mathbf{X}_{n+1}$ given $\mathbf{Y}^n$.

Consider the model A. According to the Bayes rule we find that the distributions $\mathcal{P}_{\mathbf{X}_{n+1}|\mathbf{Y}^n=\mathbf{y}^n, V=v}$ and $\mathcal{P}_{V|\mathbf{Y}^n=\mathbf{y}^n}$ are $\mathfrak{N}(\mathbf{m}_{n+1,v}, \Lambda_{n+1,v})$ and $\mathfrak{N}(r_{n+1}s_{n+1}^{-1}, s_{n+1}^{-1})$, respectively, where

$$(5)\quad \begin{cases} \mathbf{m}_{n+1,v} = \mathbf{a}_{n+1} + \mathbf{b}_{n+1} v , & n = 0,1,\ldots, \\ \Lambda_{n+1,v} = \alpha_n (\beta_n^T \Sigma_n^{-1} \beta_n + \Lambda_{n,v}^{-1}) \alpha_n^T , & n = 0,1,\ldots, \\ \mathbf{m}_{0,v} = \mathbf{m}_0 , \quad \Lambda_{0,v} = \Lambda_0 , & \\ s_{n+1} = s_n + \mathbf{b}_n^T \beta_n^T (\beta_n \Lambda_{n,v} \beta_n^T + \Sigma_n)^{-1} \beta_n \mathbf{b}_n , & n = 0,1,\ldots, \\ r_{n+1} = r_n + \mathbf{b}_n^T \beta_n^T (\beta_n \Lambda_{n,v} \beta_n^T + \Sigma_n)^{-1} (\mathbf{y}_n - \beta_n \mathbf{a}_n) , & n = 0,1,\ldots, \\ s_0 = s , \quad r_0 = r , & \end{cases}$$

with $\mathbf{a}_n$, $\mathbf{b}_n$ given as follows:

$$(6)\quad \begin{cases} \mathbf{a}_{n+1} = \alpha_n (\beta_n^T \Sigma_n^{-1} \beta_n + \Lambda_{n,v}^{-1})^{-1} (\Lambda_{n,v}^{-1} \mathbf{a}_n + \beta_n^T \Sigma_n^{-1} \mathbf{y}_n) , \\ \qquad\qquad n = 0,1,\ldots, \\ \mathbf{b}_{n+1} = \alpha_n (\beta_n^T \Sigma_n^{-1} \beta_n + \Lambda_{n,v}^{-1})^{-1} \Lambda_{n,v}^{-1} \mathbf{b}_n + \mathbf{w}_n , \quad n = 0,1,\ldots, \\ \mathbf{a}_0 = \mathbf{m}_0 , \quad \mathbf{b}_0 = 0. \end{cases}$$

Using (5), (6) and the equation

$$f_{\mathbf{X}_{n+1}|\mathbf{Y}^n=\mathbf{y}^n}(\mathbf{x}_{n+1}|\mathbf{Y}^n = \mathbf{y}^n)$$
$$= \int_{\mathbf{R}} f_{\mathbf{X}_{n+1}|\mathbf{Y}^n=\mathbf{y}^n, V=v}(\mathbf{x}_{n+1}|\mathbf{Y}^n = \mathbf{y}^n, V = v) f_{V|\mathbf{Y}^n}(v|\mathbf{Y}^n = \mathbf{y}^n)\, dv$$

we find that $\mathcal{P}_{\mathbf{X}_{n+1}|\mathbf{Y}^n}$ is the distribution $\mathfrak{N}(\mathbf{m}_{n+1}, \Lambda_{n+1})$, where the covariance matrix and the mean take the form

$$(7)\qquad \Lambda_{n+1} = \left[\Lambda_{n+1,v}^{-1} - \frac{\Lambda_{n+1,v}^{-1}\mathbf{b}_{n+1}\mathbf{b}_{n+1}^T\Lambda_{n+1,v}^{-1}}{(s_{n+1} + \mathbf{b}_{n+1}^T\Lambda_{n+1,v}^{-1}\mathbf{b}_{n+1})}\right]^{-1},$$
$$\mathbf{m}_{n+1} = \Lambda_{n+1}\Lambda_{n+1,v}^{-1}\left[\mathbf{a}_{n+1} + \frac{r_{n+1} - \mathbf{b}_{n+1}^T\Lambda_{n+1,v}^{-1}\mathbf{a}_{n+1}}{(s_{n+1} + \mathbf{b}_{n+1}^T\Lambda_{n+1,v}^{-1}\mathbf{b}_{n+1})}\mathbf{b}_{n+1}\right].$$

It is easy to verify (but not so easy to find out) that the inverse matrix which appears on the right-hand side of (7) is equal to $\Lambda_{n+1,v} + s_{n+1}^{-1}\mathbf{b}_{n+1}\mathbf{b}_{n+1}^T$. Now the mean can be expressed in the following obvious way:

$$(8)\qquad \mathbf{m}_{n+1} = \mathbf{a}_{n+1} + \mathbf{b}_{n+1}s_{n+1}^{-1}r_{n+1}.$$

Let $\widehat{\mathbf{x}}_{n+1}(r,s)$ denote the Bayes estimate w.r.t. the distribution $\mathcal{D}_{r,s}$. In view of our previous remarks, the estimate is given by the right-hand side of (8).

5. The risk functions for the Bayes strategies. By using (3)–(6) the state $\mathbf{X}_{n+1}$ and the estimate $\widehat{\mathbf{x}}_{n+1}(r,s)$ can be expressed as follows:

$$(9)\qquad \begin{cases} \mathbf{X}_{n+1} = \alpha_{n,0}\mathbf{X}_0 + \mathbf{h}_n V, & n = 0,1,\ldots, \\ \widehat{\mathbf{x}}_{n+1}(r,s) = \mathbf{c}_1(n) + \mathbf{c}_2(n)s_{n+1}^{-1} + [\mathbf{g}_1(n) + \mathbf{g}_2(n)s_{n+1}^{-1}]V & \\ \qquad + [\sigma_1(n) + \sigma_2(n)s_{n+1}^{-1}]\mathbf{X}_0 + \sum_{i=0}^n [\eta_i(n) + \gamma_i(n)s_{n+1}^{-1}]\mathbf{Z}_i, & \end{cases}$$

where the quantities $\mathbf{h}_n$, $\mathbf{c}_k(n)$, $\mathbf{g}_k(n)$, $\sigma_k(n)$, $\eta_i(n)$, $\gamma_i(n)$, $k = 1,2$, $i = 0,\ldots,n$, are independent of s and for $n = 0,1,\ldots$ they are given by

$$\mathbf{h}_n = \sum_{i=0}^{n} \alpha_{n,i+1}\mathbf{w}_i,$$
$$\mathbf{c}_1(n) = \phi_{n,0}\mathbf{m}_0,$$
$$\mathbf{c}_2(n) = \mathbf{b}_{n+1}\Big(r - \sum_{i=0}^{n} \mathbf{b}_i^T\kappa_i\phi_{i-1,0}\mathbf{m}_0\Big),$$
$$\mathbf{g}_1(n) = \sum_{i=0}^{n-1} \phi_{n,i+2}\psi_{i+1}\beta_{i+1}\mathbf{h}_i,$$

$$\mathbf{g}_2(n) = \mathbf{b}_{n+1} \sum_{i=0}^{n} \mathbf{b}_i^T \kappa_i \Big(\mathbf{h}_{i-1} - \sum_{k=0}^{i-1} \phi_{i-1,k+2} \psi_{k+1} \beta_{k+1} \mathbf{h}_k \Big) ,$$

$$\sigma_1(n) = \sum_{i=0}^{n} \phi_{n,i+1} \psi_i \beta_i \alpha_{i-1,0} ,$$

$$\sigma_2(n) = \mathbf{b}_{n+1} \sum_{i=0}^{n} \mathbf{b}_i^T \kappa_i \Big(\alpha_{i-1,0} - \sum_{k=0}^{i-1} \phi_{i-1,k+1} \psi_k \beta_k \alpha_{k-1,0} \Big) ,$$

$$\eta_i(n) = \phi_{n,i+1} \psi_i ,$$

$$\gamma_i(n) = \mathbf{b}_{n+1} \Big[\mathbf{b}_i^T \beta_i^T (\beta_i \Lambda_{i,v} \beta_i^T + \Sigma_i)^{-1} - \sum_{k=i+1}^{n} \mathbf{b}_k^T \kappa_k \phi_{k-1,i+1} \psi_i \Big]$$

with

$$\kappa_n = \beta_n^T (\beta_n \Lambda_{n,v} \beta_n^T + \Sigma_n)^{-1} \beta_n ,$$
$$\phi_n = \alpha_n (\beta_n^T \Sigma_n^{-1} \beta_n + \Lambda_{n,v}^{-1})^{-1} \Lambda_{n,v}^{-1} ,$$
$$\psi_n = \alpha_n (\beta_n^T \Sigma_n^{-1} \beta_n + \Lambda_{n,v}^{-1})^{-1} \beta_n^T \Sigma_n^{-1} .$$

We have also used the following expression for $\mathbf{b}_n$:

$$\mathbf{b}_n = \sum_{i=0}^{n-1} \phi_{n-1,i+1} \mathbf{w}_i .$$

Note that $\mathbf{c}_2(n)$ is the only quantity which depends on r.

By the definition of the risk function, using (9) we obtain

$$R(v, \widehat{\mathbf{x}}_{n+1}(r,s)) = c_2^{(n)}[s] v^2 + c_1^{(n)}[s,r] v + c_0^{(n)}[s] \tag{10}$$

where the coefficients depend on the variables which are indicated in brackets and take the form:

$$\begin{aligned} c_0^{(n)}[s] = & -2\mathbf{c}_1(n)^T \theta_n \mathbf{m}_0 + \mathbf{m}_0^T \theta_n^T \theta_n \mathbf{m}_0 + \mathbf{c}_1(n)^T \mathbf{c}_1(n) \\ & + \operatorname{tr} \theta_n^T \theta_n \Lambda_0 + \sum_{i=0}^{n} \operatorname{tr} \eta_i(n)^T \eta_i(n) \Sigma_i \\ & + 2 s_{n+1}^{-1} \Big[-\mathbf{c}_2(n)^T \theta_n \mathbf{m}_0 + \mathbf{c}_1(n)^T \sigma_2(n) \mathbf{m}_0 - \mathbf{m}_0^T \theta_0^T \sigma_2(n) \mathbf{m}_0 \\ & + \mathbf{c}_1(n)^T \mathbf{c}_2(n) - \operatorname{tr} \sigma_2(n)^T \theta_n \Lambda_0 + \sum_{i=0}^{n} \operatorname{tr} \gamma_i(n)^T \eta_i(n) \Sigma_i \Big] \\ & + s_{n+1}^{-2} \Big[\mathbf{m}_0^T \sigma_2(n)^T \sigma_2(n) \mathbf{m}_0 + \mathbf{c}_2(n)^T \mathbf{c}_2(n) + 2\mathbf{c}_2(n)^T \sigma_2(n) \mathbf{m}_0 \\ & + \operatorname{tr} \sigma_2(n)^T \sigma_2(n) \Lambda_0 + \sum_{i=0}^{n} \operatorname{tr} \gamma_i(n)^T \gamma_i(n) \Sigma_i \Big] , \end{aligned}$$

$$c_1^{(n)}[s,r] = 2\{[\mathbf{h}_n - \mathbf{g}_1(n)]^T[\theta_n \mathbf{m}_0 - \mathbf{c}_1(n)] - s_{n+1}^{-1}[\mathbf{g}_2(n)^T(\theta_n \mathbf{m}_0 - \mathbf{c}_1(n)) + (\mathbf{h}_n - \mathbf{g}_1(n))^T(\sigma_2(n)\mathbf{m}_0 + \mathbf{c}_2(n))] + s_{n+1}^{-2}\mathbf{g}_2(n)^T[\sigma_2(n)\mathbf{m}_0 + \mathbf{c}_2(n)]\},$$

$$c_2^{(n)}[s] = [\mathbf{h}_n - \mathbf{g}_1(n)]^T[\mathbf{h}_n - \mathbf{g}_1(n)] - 2s_{n+1}^{-1}\mathbf{g}_2(n)^T[\mathbf{h}_n - \mathbf{g}_1(n)] + s_{n+1}^{-2}\mathbf{g}_2(n)^T\mathbf{g}_2(n),$$

with

$$\theta_n = \alpha_{n,0} - \sigma_1(n), \quad n = 0,1,\ldots$$

Note that the above coefficients depend on s through the quantity s_{n+1} and the coefficient $c_1^{(n)}[s,r]$ depends on r through $\mathbf{c}_2(n)$.

For every real m let $\widehat{\mathbf{x}}_{n+1}(m)$ denote the estimate given by

$$\widehat{\mathbf{x}}_{n+1}(m) = \mathbf{a}_{n+1} + \mathbf{b}_{n+1}m$$

where $\mathbf{a}_{n+1}$ and $\mathbf{b}_{n+1}$ are given by (6).

It is easy to verify that for every v

$$R(v, \widehat{\mathbf{x}}_{n+1}(m)) = \lim_{\substack{s\to\infty \\ rs^{-1}\to m}} R(v, \widehat{\mathbf{x}}_{n+1}(r,s))$$

and for every distribution $\mathcal{D}$ of the random parameter V

$$r(\mathcal{D}, \widehat{\mathbf{x}}_{n+1}(m)) = \lim_{\substack{s\to\infty \\ rs^{-1}\to m}} r(\mathcal{D}, \widehat{\mathbf{x}}_{n+1}(r,s)). \tag{11}$$

6. Results for the model B. Considerations similar to those in Sections 4 and 5 lead to the following facts holding for the model B:

FACT 1. *Let $\widetilde{\mathbf{x}}_{n+1}(r,s)$ denote the Bayes estimate w.r.t. the distribution $\mathcal{D}_{r,s}$. Then*

$$\widetilde{\mathbf{x}}_{n+1}(r,s) = \mathbf{p}_{n+1} + \mathbf{q}_{n+1}r_{n+1}s_{n+1}^{-1}$$

where

$$\begin{cases} \mathbf{p}_{n+1} = \alpha_n\beta_n^{-1}\mathbf{Y}_n, & n = 0,1,\ldots, \\ \mathbf{p}_0 = \mathbf{m}_0, & \\ \mathbf{q}_{n+1} = \alpha_n\beta_n^{-1}\mathbf{w}_n, & n = 0,1,\ldots, \\ \mathbf{q}_0 = 0, & \\ r_{n+1} = r_n + (\mathbf{w}_n + \beta_n\mathbf{q}_n)^T(\beta_n^T)^{-1}\Sigma_n^{-1}\beta_n^{-1}(\mathbf{Y}_n - \beta_n\mathbf{p}_n), & n = 0,1,\ldots, \\ r_0 = r, & \\ s_{n+1} = s_n + (\mathbf{w}_n + \beta_n\mathbf{q}_n)^T(\beta_n^T)^{-1}\Sigma_n^{-1}\beta_n^{-1}(\mathbf{w}_n + \beta_n\mathbf{q}_n), & n = 0,1,\ldots, \\ s_0 = s. & \end{cases} \tag{12}$$

FACT 2. *The risk function $R(v, \widetilde{\mathbf{x}}_{n+1}(r,s))$ can be expressed as follows:*

$$R(v, \widetilde{\mathbf{x}}_{n+1}(r,s)) = d_2^{(n)}[s]v^2 + d_1^{(n)}[s,r]v + d_0^{(n)}[s]$$

where

$$\begin{aligned} d_0^{(n)}[s] = {} & \operatorname{tr} \Sigma_n + s_{n+1}^{-2}\Big(\mathbf{q}_{n+1}^T \mathbf{q}_{n+1} r^2 + \operatorname{tr} \mathbf{u}_0 \mathbf{q}_{n+1}^T \mathbf{q}_{n+1} \mathbf{u}_0^T \Lambda_0 \\ & + \sum_{i=0}^{n-1} \operatorname{tr} \mathbf{u}_{i+1} \mathbf{q}_{n+1}^T \mathbf{q}_{n+1} \mathbf{u}_{i+1}^T \Sigma_i\Big), \\ d_1^{(n)}[s,r] = {} & -2\mathbf{q}_{n+1}^T \mathbf{q}_{n+1}\Big(1 - s_{n+1}^{-1} \sum_{i=0}^{n} h_i\Big) s_{n+1}^{-1} r, \\ d_2^{(n)}[s] = {} & \mathbf{q}_{n+1}^T \mathbf{q}_{n+1}\Big(1 - s_{n+1}^{-1} \sum_{i=0}^{n} h_i\Big)^2, \end{aligned}$$

with

$$\begin{aligned} \mathbf{u}_n &= \Lambda_{n,v}^{-1} \mathbf{q}_n, \quad n = 0, 1, \ldots, \\ h_n &= (\beta_n^{-1} \mathbf{w}_n - \mathbf{q}_n)^T \mathbf{u}_n, \quad n = 1, 2, \ldots, \\ h_0 &= (\beta_0^{-1} \mathbf{w}_0)^T \mathbf{u}_0 . \end{aligned}$$

FACT 3. *For every real m let $\widetilde{\mathbf{x}}_{n+1}(m)$ denote the estimate given by*

$$\widetilde{\mathbf{x}}_{n+1}(m) = \mathbf{p}_{n+1} + \mathbf{q}_{n+1} m$$

where $\mathbf{p}_{n+1}$ and $\mathbf{q}_{n+1}$ are given by (12). *Then*

$$R(v, \widetilde{\mathbf{x}}_{n+1}(m)) = \lim_{\substack{s \to \infty \\ rs^{-1} \to m}} R(v, \widetilde{\mathbf{x}}_{n+1}(r,s)),$$

for every v, and

$$r(\mathcal{D}, \widetilde{\mathbf{x}}_{n+1}(m)) = \lim_{\substack{s \to \infty \\ rs^{-1} \to m}} r(\mathcal{D}, \widetilde{\mathbf{x}}_{n+1}(r,s))$$

for every distribution $\mathcal{D}$ of the random parameter V.

7. The minimax estimates. The following propositions provide solutions for Problems A1 and B1.

PROPOSITION 1. *A $\mathcal{G}_1$-minimax estimate for the model A always exists and*

(i) *if $\forall s \geq a^{-1}$, $c_1[s, \sqrt{s^2 a - s}] > 0$ then $\widehat{\mathbf{x}}_{n+1}(\sqrt{a})$ is a $\mathcal{G}_1$-minimax estimate,*

(ii) *if $\forall s \geq a^{-1}$, $c_1[s, -\sqrt{s^2 a - s}] < 0$ then $\widehat{\mathbf{x}}_{n+1}(-\sqrt{a})$ is a $\mathcal{G}_1$-minimax estimate,*

(iii) *if $\exists s^*$, r^* such that $s^* > 0$, $c_1(s^*, r^*) = 0$, $(r^*)^2 (s^*)^{-2} + (s^*)^{-1} = a$ then $\widehat{\mathbf{x}}_{n+1}(r^*, s^*)$ is a $\mathcal{G}_1$-minimax estimate.*

PROPOSITION 2. *The estimate $\widetilde{\mathbf{x}}_{n+1}(0, a^{-1})$ is a $\mathcal{G}_1$-minimax estimate for the model B.*

Proof of Proposition 1. By (1) and (10) the Bayes risk $r(\mathcal{D}, \widehat{\mathbf{x}}_{n+1}(r,s))$ can be expressed as follows:

$$r(\mathcal{D}, \widehat{\mathbf{x}}_{n+1}(r,s)) = c_2^{(n)}[s]E_{\mathcal{D}}V^2 + c_1^{(n)}[s,r]E_{\mathcal{D}}V + c_0^{(n)}[s].$$

Notice that $c_2^{(n)}[s] \geq 0$ for each s.

Now, suppose that the condition in (i) is fulfilled. In view of (11) for each $\mathcal{D} \in \mathcal{G}_1$ we can write

$$\begin{aligned} r(\mathcal{D}, \widehat{\mathbf{x}}_{n+1}(\sqrt{a})) &= \lim_{s\to\infty} \{c_2^{(n)}[s]E_{\mathcal{D}}V^2 + c_1^{(n)}[s, \sqrt{s^2a - s}]E_{\mathcal{D}}V + c_0^{(n)}[s]\} \\ &\leq \lim_{s\to\infty} \{c_2^{(n)}[s]a + c_1^{(n)}[s, \sqrt{s^2a - s}]\sqrt{a} + c_0^{(n)}[s]\} \\ &= \lim_{s\to\infty} \{c_2^{(n)}[s]a + c_1^{(n)}[s, \sqrt{s^2a - s}]s^{-1}\sqrt{s^2a - s} + c_0^{(n)}[s]\} \\ &= \lim_{s\to\infty} r(\mathcal{D}_{\sqrt{s^2a-s},s}, \widehat{\mathbf{x}}_{n+1}(\sqrt{s^2a - s}, s)). \end{aligned}$$

In view of our Lemma (see Section 1) this implies that $\widehat{\mathbf{x}}_{n+1}(\sqrt{a})$ is a $\mathcal{G}_1$-minimax estimate.

The proof of (ii) is similar.

In case (iii), for each $\mathcal{D} \in \mathcal{G}_1$ we obtain

$$\begin{aligned} r(\mathcal{D}, \widehat{\mathbf{x}}_{n+1}(r^*, s^*)) &= c_2^{(n)}[s^*]E_{\mathcal{D}}V^2 + c_1^{(n)}[s^*, r^*]E_{\mathcal{D}}V + c_0^{(n)}[s^*] \\ &\leq c_2^{(n)}[s^*]a + c_0^{(n)}[s^*] = r(\mathcal{D}_{r^*,s^*}, \widehat{\mathbf{x}}_{n+1}(r^*, s^*)). \end{aligned}$$

Setting $\mathcal{D}_k = \mathcal{D}_{r^*,s^*}$ and $\mathbf{d}_k = \widehat{\mathbf{x}}_{n+1}(r^*, s^*)$ for each k in the Lemma we find that the estimate fulfils the condition (2), so it is a $\mathcal{G}_1$-minimax estimate.

It remains to prove that a $\mathcal{G}_1$-minimax estimate always exists; but this follows from the fact that one of the three conditions given in (i)–(iii) must be fulfilled.

Proof of Proposition 2. Notice that the distribution $\mathcal{D}_{0,a^{-1}}$ belongs to $\mathcal{G}_1$ and $d_1^{(n)}[a^{-1}, 0] = 0$. So our assertion can be proved in a similar way to (iii) of the previous proposition.

The next proposition gives $\mathcal{G}_2$-minimax estimates for our models.

PROPOSITION 3. *If $\overline{s} = (m_2 - m_1^2)^{-1}$ and $\overline{r} = m_1(m_2 - m_1^2)^{-1}$ then*

(i) *$\widehat{\mathbf{x}}_{n+1}(\overline{r}, \overline{s})$ is a $\mathcal{G}_2$-minimax estimate for the model A,*
(ii) *$\widetilde{\mathbf{x}}_{n+1}(\overline{r}, \overline{s})$ is a $\mathcal{G}_2$-minimax estimate for the model B.*

Proof. Notice that $\mathcal{D}_{\overline{r},\overline{s}} \in \mathcal{G}_2$. For each $\mathcal{D} \in \mathcal{G}_2$ we have

$$\begin{aligned} r(\mathcal{D}, \widehat{\mathbf{x}}_{n+1}(\overline{r}, \overline{s})) &= c_2^{(n)}[\overline{s}]E_{\mathcal{D}}V^2 + c_1^{(n)}[\overline{s}, \overline{r}]E_{\mathcal{D}}V + c_0^{(n)}[\overline{s}] \\ &= c_2^{(n)}[\overline{s}]m_2 + c_1^{(n)}[\overline{s}, \overline{r}]m_1 + c_0^{(n)}[\overline{s}] = \text{const.} \end{aligned}$$

Hence, in view of the Corollary from Section 1 assertion (i) of the proposition is valid. (ii) can be proved in a similar way.

References

[1] D. L. Alspach, L. L. Scharf and A. Abiri, *A Bayesian solution to the problem of state estimation in an unknown noise environment*, Internat. J. Control 19 (2) (1974), 265–287.

[2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, New Jersey, 1979.

[3] J. C. Darragh and D. P. Looze, *Noncausal minimax linear state estimation for systems with uncertain second-order statistics*, IEEE Trans. Automat. Control AC-29 (1984), 555–557.

[4] G. A. Golubev, *Minimax linear filtering of dynamic discrete time processes*, Automat. Remote Control 45 (1984), 203–211.

[5] M. Mintz, *A Kalman filter as a minimax estimator*, J. Optim. Theory Appl. 9 (2) (1972), 99–111.

[6] J. M. Morris, *The Kalman filter: A robust estimator for some classes of linear quadratic problems*, IEEE Trans. Inform. Theory IT-22 (1976), 526–534.

[7] H. V. Poor and D. P. Looze, *Minimax state estimation for linear stochastic systems with noise uncertainty*, IEEE Trans. Automat. Control AC-26 (1981), 902–906.

[8] H. Rootzen and J. Sternby, *Consistency in least-squares estimation: A Bayesian approach*, Automatica 20 (1984), 471–477.

[9] S. Trybuła, *Optimal control for hypergeometric processes*, Systems Science 11 (1985), 31–57.

[10] S. Verdu and H. V. Poor, *Minimax linear observers and regulators for stochastic systems with uncertain second-order statistics*, IEEE Trans. Automat. Control AC-29 (1984), 499–511.

[11] S. Zacks, *The Theory of Statistical Inference*, Wiley, New York 1971.

ANDRZEJ GRZYBOWSKI
INSTITUTE OF MATHEMATICS
TECHNICAL UNIVERSITY OF CZĘSTOCHOWA
UL. A. DEGLERA 35
42-200 CZĘSTOCHOWA, POLAND

Received on 9.2.1989;
revised version on 10.10.1989