ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2023, Vol. 84, No. 2, pp. 121–134.

On the Lagrange Duality of Stochastic and Deterministic Minimax Control and Filtering Problems

M. M. Kogan

Nizhny Novgorod State University of Architecture and Civil Engineering, Nizhny Novgorod, Russia
e-mail: mkogan@nngasu.ru

Received August 24, 2022
Revised November 18, 2022
Accepted November 30, 2022

Abstract—As shown below, the linear operator norms in the deterministic and stochastic cases are optimal values of the Lagrange-dual problems. For linear time-varying systems on a finite horizon, the duality principle leads to stochastic interpretations of the generalized H_2 and H_∞ norms of the system. Stochastic minimax filtering and control problems with unknown covariance matrices of random factors are considered. Equations of generalized H_∞-suboptimal controllers, filters, and identifiers are derived to achieve a trade-off between the error variance at the end of the observation interval and the sum of the error variances on the entire interval.

Keywords: stochastic minimax control, Kalman filter, Lagrange duality, generalized H_∞-optimal control and filtering, generalized H_2-optimal control, linear matrix inequalities

DOI: 10.25728/arcRAS.2023.12.47.001

1. INTRODUCTION

As is known, there exist two major concepts in the theory of control under uncertainty: stochastic, when uncertain factors (initial conditions, disturbances, and perturbations) are assumed to be random and are assigned some probabilistic characteristics (mean and covariance), and deterministic, when uncertain factors are assumed to be deterministic and take values in some sets. In the former case, the goal is to optimize a performance criterion of the system in a probabilistic sense under given probabilistic characteristics or their bounds. given. In the latter case, the goal is to minimize the maximum value of a performance criterion on the value set of uncertain factors satisfying given constraints. Within both concepts, effective methods were developed for solving a variety of linear-quadratic estimation, filtering, and control problems, including problems with the H_∞ norm as a performance criterion for deterministic and stochastic systems; for example, see [1–5]. However, both concepts suffer from drawbacks: in the stochastic paradigm, it is difficult to determine the probabilistic characteristics of uncertain factors; in the deterministic one, it is difficult to find adequate constraints on the values of uncertain factors. To overcome these disadvantages, broaden the range of problems resolved, and interpret the results obtained in one paradigm in terms of the other, it is crucial to identify the relationship between these paradigms.

Note that this topic was addressed repeatedly: a classical example is the relationship between the recurrence least-squares method and the Kalman filter established in [6]. In addition, as shown in [7], the static Kalman filter is optimal for random (white) noise and, moreover, minimizes over time the maximum Euclidean norm of the error under deterministic disturbances with bounded energy. In the widely known paper [8], J.C. Willems gave a purely deterministic interpretation of the results obtained in linear-quadratic optimal stochastic filtering and control. (Also, see [9].) Namely, he proved the following fact: among all deterministic disturbances under which the observed signal can be realized, let us choose the one with the smallest norm and substitute it into the system equations; then the resulting equations will coincide with the equations of the optimal filter or controller under random Gaussian disturbances.

In this paper, we demonstrate the Lagrange duality of two optimization problems: maximizing the squared Euclidean norm of the output of a linear operator (transformation) that maps deterministic vectors satisfying an ellipsoidal constraint and maximizing the output variance of this operator that maps random vectors satisfying an averaged ellipsoidal constraint. For operators generated by linear dynamic systems, this leads to the duality of stochastic and deterministic minimax estimation and control problems with performance criteria in the form of the generalized H_2 and H_∞ norms. With the duality principle, we formulate and solve new problems of optimal and robust control and filtering in the stochastic statement with unknown covariance matrices of random factors. Also, we derive equations of generalized H_∞-suboptimal controllers, filters, and identifiers to achieve a trade-off between the error variance at the end of the observation interval and the sum of the error variances on the entire time interval.

2. THE LAGRANGE DUALITY OF STOCHASTIC AND DETERMINISTIC PARADIGMS

Consider two vectors $\xi \in \mathrm{R}^{n_\xi}$ and $\eta \in \mathrm{R}^{n_\eta}$, further called the input and output, respectively, with the linear relationship

$$\eta = \Psi\xi, \tag{2.1}$$

where Ψ is a deterministic matrix of dimensions $(n_\eta \times n_\xi)$. We introduce the notations

$$|a|_R^2 = a^{\mathrm{T}} R^{-1} a, \quad \|b\|^2_{G\,[t_0,t]} = \sum_{i=t_0}^{t-1} |b(i)|^2_{G(i)},$$

where $R = R^{\mathrm{T}} > 0$ and $G(t) = G^{\mathrm{T}} > 0$ are weight matrices. Let the generalized norm of the operator Ψ with a weight matrix $K = K^{\mathrm{T}} > 0$ be defined as

$$\|\Psi\|_K^2 = \sup_{\xi\neq 0} \frac{|\eta|^2}{|\xi|_K^2} = \sup_{\xi \neq 0} \frac{\xi^{\mathrm{T}}\Psi^{\mathrm{T}}\Psi\xi}{\xi^{\mathrm{T}} K^{-1}\xi} = \lambda_{\max}(\Psi K \Psi^{\mathrm{T}}). \tag{2.2}$$

This induced generalized norm of the operator Ψ, further called the damping rate of deterministic disturbances, equals the maximum value of the Euclidean norm of the output η under all inputs ξ belonging to the ellipsoid $\mathcal{E}_\xi(K) = \left\{\xi : \xi^{\mathrm{T}} K^{-1}\xi \leqslant 1\right\}$. It can be found by solving the following optimization problem.

Problem D.

$$\gamma_d^2(\Psi) = \max_\xi |\eta|^2 : \ \eta = \Psi\xi, \quad \xi^{\mathrm{T}} K^{-1} \xi \leqslant 1. \tag{2.3}$$

If $\xi = \xi_s$ is a random vector with zero mean and the covariance matrix $E\xi_s\xi_s^{\mathrm{T}} = K_\xi$, then the output covariance matrix is $K_\eta = \Psi K_\xi \Psi^{\mathrm{T}}$ and the expectation of the squared Euclidean norm of the output is the trace of this matrix, i.e., $E|\eta|^2 = \mathrm{tr}\,(\Psi K_\xi \Psi^{\mathrm{T}})$. Given an unknown covariance matrix of the vector ξ, we define the damping rate of the random disturbance of the operator Ψ as the square root of the maximum value of the ratio of the output variance to the expectation of the quadratic form with the input matrix K^{-1} under all nonzero covariance input matrices K_ξ [10]:

$$\gamma_s^2(\Psi) = \sup_{K_\xi \geqslant 0} \frac{E|\eta|^2}{E|\xi|_K^2} = \sup_{K_\xi \geqslant 0} \frac{\mathrm{tr}\,\Psi K_\xi \Psi^{\mathrm{T}}}{\mathrm{tr}\,K^{-1} K_\xi}.$$

This value represents an induced norm of the operator Ψ with random vectors ξ and η equipped with the norms $|\xi|_s = (E\xi^{\mathrm{T}}K^{-1}\xi)^{1/2}$ and $|\eta|_s = (E|\eta|^2)^{1/2}$, respectively. It can be found by solving the following optimization problem.

Problem S.

$$\gamma_s^2(\Psi) = \max_{K_\xi \geqslant 0} \operatorname{tr} \Psi K_\xi \Psi^{\mathrm{T}} : \eta = \Psi\xi, \quad \operatorname{tr} K^{-1}K_\xi \leqslant 1. \tag{2.4}$$

Theorem 2.1. *Problems* S *and* D *are dual, and the damping rates of the random and deterministic disturbances with the weight matrix K of the operator Ψ coincide with the spectral radius of the output covariance matrix under the input covariance matrix equal to the weight matrix:*

$$\gamma_s^2(\Psi) = \gamma_d^2(\Psi) = \lambda_{\max}(\Psi K \Psi^{\mathrm{T}}).$$

Proof of Theorem 2.1. We write the Lagrange function for problem S and express the optimal value of its dual function as

$$\min_{\lambda \geqslant 0} \max_{K_\xi \geqslant 0} \left[\operatorname{tr} \Psi K_\xi \Psi^{\mathrm{T}} + \lambda(1 - \operatorname{tr} K^{-1}K_\xi)\right] = \min_{\lambda \geqslant 0} \max_{K_\xi \geqslant 0} \left[\lambda + \operatorname{tr} K_\xi(\Psi^{\mathrm{T}}\Psi - \lambda K^{-1})\right].$$

This value is finite if $\Psi^{\mathrm{T}}\Psi - \lambda K^{-1} \leqslant 0$, and then the maximum is reached at $K_\xi = 0$. In this case, the optimal value of the dual problem coincides with (2.2). Since the function is convex and there exists an interior point satisfying the constraint, the optimal values of the primal and dual problems coincide [11]. The proof of this theorem is complete.

In addition, the image of the operator Ψ (in other words, the reachability set of the vector η in (2.1) under all deterministic vectors ξ belonging to the ellipsoid $\mathcal{E}_\xi(K)$) can be characterized in terms of the covariance matrix of the random vector ξ, as formulated in the following theorem.

Theorem 2.2. *If the deterministic input ξ takes values in the ellipsoid $\mathcal{E}_\xi(K)$ and the matrix $K_\eta = \Psi K \Psi^{\mathrm{T}}$ is nonsingular, the reachability set of the vector η in* (2.1) *is an ellipsoid $\mathcal{E}_\eta(K_\eta)$ with the matrix K_η coinciding with the covariance matrix of the random vector η provided that the covariance matrix of the random vector ξ is K.*

Proof of Theorem 2.2. We find the support function of the reachability set η in (2.1), i.e.,

$$\varrho(x) = \sup_{\xi \in \mathcal{E}_\xi(K)} x^{\mathrm{T}}\eta.$$

If the vector x has unit length, the support function is the upper bound of the projections of x onto this set. For the resulting constrained optimization problem, the Lagrange function takes the form

$$L(\xi, \lambda) = x^{\mathrm{T}}\Psi\xi + \lambda(1 - \xi^{\mathrm{T}}K^{-1}\xi).$$

Equating to zero the gradient of this function with respect to ξ, we find $\xi = (2\lambda)^{-1}K\Psi^{\mathrm{T}}x$. Substituting this expression into the constraint yields $2\lambda = (x^{\mathrm{T}}\Psi K \Psi^{\mathrm{T}}x)^{1/2}$. Finally, we obtain $\varrho(x) = (x^{\mathrm{T}}\Psi K \Psi^{\mathrm{T}}x)^{1/2}$, which corresponds to the support function of the ellipsoid $\mathcal{E}_\eta(\Psi K \Psi^{\mathrm{T}})$ [12]. The proof of this theorem is complete.

The results above remain valid if the random vector ξ has a nonzero mean $E\xi = \xi_*$. In this case, the damping rate of the random disturbances is given by

$$\gamma_s^2(\Psi) = \sup_{K_\xi \geqslant 0} \frac{E|\eta - \Psi\xi_*|^2}{E|\xi - \xi_*|_K^2} = \sup_{K_\xi \geqslant 0} \frac{\operatorname{tr} \Psi K_\xi \Psi^{\mathrm{T}}}{\operatorname{tr} K^{-1}K_\xi},$$

where $K_\xi = E(\xi - \xi_*)(\xi - \xi_*)^{\mathrm{T}}$; the damping rate of the deterministic disturbances is given by

$$\gamma_d^2(\Psi) = \sup_{\xi \neq \xi_*} \frac{|\eta - \Psi\xi_*|^2}{|\xi - \xi_*|_K^2} = \sup_{\xi \neq \xi_*} \frac{(\xi - \xi_*)^{\mathrm{T}}\Psi^{\mathrm{T}}\Psi(\xi - \xi_*)}{(\xi - \xi_*)^{\mathrm{T}}K^{-1}(\xi - \xi_*)} = \lambda_{\max}(\Psi K \Psi^{\mathrm{T}}).$$

Here, the reachability set of the vector η is an ellipsoid centered at the point $\eta_* = \Psi\xi_*$ with the matrix $K_\eta = \Psi K \Psi^{\mathrm{T}}$.

Next, we apply these results to linear operators induced by linear dynamic systems.

3. THE MAXIMUM DEVIATION AND THE GENERALIZED H_2 NORM

We define a linear operator $\Psi_2(t)$ mapping the input $\xi(t) = \mathrm{col}\,(x_0, v(t_0), \ldots, v(t-1))$ into the output $\eta = z(t)$ of a linear dynamic system described by the equations

$$\begin{gathered} x(t+1) = A(t)x(t) + B(t)v(t), \quad x(t_0) = x_0, \\ z(t) = C(t)x(t), \quad t \in [t_0, t_f]. \end{gathered} \tag{3.1}$$

The solution of this system can be represented as

$$z(t) = C(t)\Phi(t, t_0)x(t_0) + \sum_{i=t_0}^{t-1} F(t, i+1)B(i)v(i), \tag{3.2}$$

where $F(t, i) = C(t)\Phi(t, i+1)B(i)$, $i = t_0, \ldots, t-1$, is the unit impulse response matrix of the system and $\Phi(t, t_0)$ is the transition matrix of the system, i.e.,

$$\Phi(t+1, t_0) = A(t)\Phi(t, t_0), \quad t \geqslant t_0, \quad \Phi(t_0, t_0) = I. \tag{3.3}$$

Therefore,

$$\Psi_2(t) = (C(t)\Phi(t, t_0)\ F(t, t_0)\ \ldots\ C(t)B(t-1)). \tag{3.4}$$

With a block-diagonal weight matrix

$$K(t) = \mathrm{diag}\,(R, G(t_0), \ldots, G(t-1)),$$

where $R = R^{\mathrm{T}} > 0$ and $G(i) = G^{\mathrm{T}}(i) > 0$, $i = 0, \ldots, t-1$, the generalized norm of this operator,

$$\|\Psi_2(t)\|_{K(t)}^2 = \sup_{\xi(t) \neq 0} \frac{|\Psi_2(t)\xi(t)|^2}{|\xi(t)|_{K(t)}^2} = \sup_{x_0,\, v(\tau), \tau \in [t_0, t-1]} \frac{|z(t)|^2}{|x_0|_R^2 + \|v\|_{G\,[t_0, t]}^2},$$

is the so-called maximum deviation of the system output at the time instant t under arbitrary deterministic initial state and disturbance satisfying the constraint

$$|\xi(t)|_{K(t)}^2 = |x_0|_R^2 + \|v\|_{G\,[t_0, t]}^2 \leqslant 1. \tag{3.5}$$

It will be called the ellipsoidal constraint. According to [13, 14], the generalized H_2 norm of system (3.1) on a finite horizon $[t_0, t_f]$ with weight matrices R and $G(t)$ is defined as

$$\|H\|_{g2}^2 = \sup_{x_0,\, v(t), t \in [t_0, t_f - 1]} \frac{\max\limits_{\tau \in [t_0, t]} |z(\tau)|^2}{|x_0|_R^2 + \|v\|_{G\,[t_0, t]}^2}.$$

In other words, this norm is the maximum value over time among the maximum deviations of the system output.

We define the damping rate of random disturbances in system (3.1) at a time instant t as the damping rate of the random disturbances of the operator $\Psi_2(t)$:

$$\gamma_s^2(\Psi_2(t)) = \sup_{K_{\xi(t)}\geqslant 0} \frac{E|z(t)|^2}{E\left(|x_0|_R^2 + \|v\|_{G\,[t_0,t]}^2\right)},$$

where the random initial state and disturbances are independent. (In the general case, they are colored noises.)

By Theorem 2.1, the damping rates of the random and deterministic disturbances are $\lambda_{\max}[K_z(t)]$, where $K_z(t)$ denotes the covariance matrix of the system output provided that its input $\xi(t) = \mathrm{col}\,(x_0, v(t_0), \ldots, v(t-1))$ has the covariance matrix $K(t) = \mathrm{diag}\,(R, G(t_0), \ldots, G(t-1))$. (In other words, its initial state and disturbances are independent and have the covariance matrices R and $G(t)$, respectively.) From the system equations we obtain $K_z(t) = C(t)P(t)C^{\mathrm{T}}(t)$, where $P(t)$ is the solution of the equation

$$P(t+1) = A(t)P(t)A^{\mathrm{T}}(t) + B(t)G(t)B^{\mathrm{T}}(t) \tag{3.6}$$

with the initial condition $P(t_0) = R$. Thus, we arrive at the following result.

Theorem 3.1. *The damping rate of random disturbances in system* (3.1) *coincides with the maximum deviation of its output, i.e.,*

$$\begin{aligned}&\sup_{K_{\xi(t)}\geqslant 0} \frac{E|z(t)|^2}{E\left(|x_0|_R^2 + \|v\|_{G\,[t_0,t]}^2\right)}\\ &= \sup_{x_0,\, v(\tau), \tau\in[t_0,t-1]} \frac{|z(t)|^2}{|x_0|_R^2 + \|v\|_{G\,[t_0,t]}^2} = \lambda_{\max}[C(t)P(t)C^{\mathrm{T}}(t)],\end{aligned} \tag{3.7}$$

where $P(t)$ is the solution of equation (3.6).

Corollary 3.1. *The generalized H_2 norm of system* (3.1) *is characterized as*

$$\|H\|_{g2}^2 = \max_{t\in[t_0,t_f]} \sup_{K_{\xi(t)}\geqslant 0} \frac{E|z(t)|^2}{E\left(|x_0|_R^2 + \|v\|_{G\,[t_0,t]}^2\right)} = \max_{t\in[t_0,t_f]} \lambda_{\max}[C(t)P(t)C^{\mathrm{T}}(t)].$$

Remark 1. From (3.6) and (3.7) it follows that the damping rate of random disturbances at a time instant t can be found by solving the semidefinite programming problem

$$\begin{gathered}\min \lambda : Y(i+1) - A(i)Y(i)A^{\mathrm{T}}(i) - B(i)G(i)B^{\mathrm{T}}(i) \geqslant 0, \quad i = t_0, \ldots, t-1,\\ Y(t_0) = R, \quad C(t)Y(t)C^{\mathrm{T}}(t) \leqslant \lambda I;\end{gathered} \tag{3.8}$$

the generalized H_2 norm is the minimum value $\lambda > 0$ under which inequalities (3.8) and $C(i)Y(i)C^{\mathrm{T}}(i) \leqslant \lambda I$ hold for $t = t_f$ and $i = t_0, \ldots, t_f$, respectively.

4. THE GENERALIZED H_∞ NORM

Along the trajectories of the system

$$\begin{gathered}x(t+1) = A(t)x(t) + B(t)v(t), \quad x(t_0) = x_0,\\ z(t) = C(t)x(t) + D(t)v(t), \quad t \in [t_0, t_f],\end{gathered} \tag{4.1}$$

we define a linear operator Ψ_∞ such that $\eta = \Psi_\infty \xi$, where

$$\eta = \mathrm{col}\,(z(t_0), \ldots, z(t_f-1), S^{1/2}x(t_f)), \quad \xi = \mathrm{col}\,(x(t_0), v(t_0), \ldots, v(t_f-1)),$$

and $S = S^{\mathrm{T}} \geqslant 0$ is a given matrix. Considering (3.2) and (3.3), we find

$$\Psi_\infty = \begin{pmatrix} C(t_0) & D(t_0) & \cdot & 0 \\ C(t_0+1)\Phi(t_0+1,t_0) & F(t_0+1,t_0) & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ C(t_f-1)\Phi(t_f-1,t_0) & F(t_f-1,t_0) & \cdot & D(t_f-1) \\ S^{1/2}\Phi(t_f,t_0) & S^{1/2}\Phi(t_f,t_0+1)B(t_0) & \cdot & S^{1/2}B(t_f-1) \end{pmatrix}.$$

With the weight matrix $K = \mathrm{diag}\,(R, G(t_0), \ldots, G(t_f-1))$, the generalized norm of this operator coincides with its generalized H_∞ norm:

$$\|\Psi_\infty\|_K^2 = \sup_{\xi \neq 0} \frac{|\Psi_\infty \xi|^2}{|\xi|_K^2} = \sup_{x(t_0),\, v(\tau), \tau \in [t_0, t_f-1]} \frac{\|z\|^2_{[t_0,t_f]} + x(t_f)^{\mathrm{T}} S x(t_f)}{|x(t_0)|_R^2 + \|v\|^2_{G\,[t_0,t_f]}} = \|H\|^2_{g\infty},$$

where $S = S^{\mathrm{T}} \geqslant 0$ is a given weight matrix. Thus, the damping rate of deterministic disturbances in system (4.1) can be obtained by solving the following optimization problem.

Problem D_∞.

$$\gamma_d^2(\Psi_\infty) = \max_{x_0,\, v(\tau), \tau \in [t_0, t_f-1]} \sum_{t=t_0}^{t_f-1} |z(t)|^2 + x^{\mathrm{T}}(t_f) S x(t_f) : |x_0|_R^2 + \sum_{t=t_0}^{t_f-1} |v(t)|^2_{G(t)} \leqslant 1 \tag{4.2}$$

under the assumption that the initial and exogenous disturbances in system (4.1) form a sequence of deterministic vectors $\xi = \mathrm{col}\,(x(t_0), v(t_0), \ldots, v(t_f-1))$.

In turn, we define the damping rate of random disturbances as

$$\gamma_s^2(\Psi_\infty) = \sup_{K_\xi \geqslant 0} \frac{E\left[\|z\|^2_{[t_0,t_f]} + x^{\mathrm{T}}(t_f) S x(t_f)\right]}{E\left[|x(t_0)|_R^2 + \|v\|^2_{G[t_0,t_f]}\right]}, \tag{4.3}$$

where K_ξ is the covariance matrix of the vector $\xi = \mathrm{col}\,(x(t_0), v(t_0), \ldots, v(t_f-1))$. According to Theorem 2.1, the damping rates of the random and deterministic disturbances are equal to each other, i.e., $\gamma_s^2(\Psi_\infty) = \|H\|^2_{g\infty}$.

We show that the damping rate of random disturbances can be found by solving a semidefinite programming problem with certain covariance matrices as the variables. Denoting

$$E \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} \begin{pmatrix} x(t) \\ v(t) \end{pmatrix}^{\mathrm{T}} = W(t), \quad M = (I\ 0), \quad H = (0\ I) \tag{4.4}$$

and omitting the argument t wherever no confusion occurs, from (4.1) we obtain the following equations for the matrices $W(t)$:

$$MW(t+1)M^{\mathrm{T}} = (A\ B)W(t)(A\ B)^{\mathrm{T}}, \quad t = t_0, \ldots, t_f-1. \tag{4.5}$$

We express the means in (4.3) in terms of the matrices $W(t)$ and formulate another problem.

Problem S_∞.

$$\gamma_s^2(\Psi_\infty) = \max_{W(t)\geqslant 0,\, t\in[t_0,t_f]} \sum_{t=t_0}^{t_f-1} \operatorname{tr}(C\ D)W(t)(C\ D)^{\mathrm{T}} + \operatorname{tr} SMW(t_f)M^{\mathrm{T}}\ : \tag{4.6}$$
$$\operatorname{tr} R^{-1}MW(t_0)M^{\mathrm{T}} + \sum_{t=t_0}^{t_f-1} \operatorname{tr} G^{-1}HW(t)H^{\mathrm{T}} \leqslant 1,$$

where the matrices $W(t)$ satisfy equations (4.5).

The next result is proved in the Appendix.

Theorem 4.1. *Problems* S_∞ *and* D_∞ *are Lagrange dual and their optimal values coincide with the generalized* H_∞ *norm of system* (4.1), *i.e.,*

$$\|H\|_{g\infty}^2 = \sup_{W(t)\geqslant 0\, t\in[t_0,t_f]} \frac{\sum\limits_{t=t_0}^{t_f-1} \operatorname{tr}(C\ D)W(t)(C\ D)^{\mathrm{T}} + \operatorname{tr} SMW(t_f)M^{\mathrm{T}}}{\operatorname{tr} R^{-1}MW(t_0)M^{\mathrm{T}} + \sum\limits_{t=t_0}^{t_f-1} \operatorname{tr} G^{-1}HW(t)H^{\mathrm{T}}}.$$

This norm is calculated as

$$\|H\|_{g\infty}^2 = \min_{\lambda\geqslant 0, X(t)} \lambda :$$
$$\begin{pmatrix} A^{\mathrm{T}}X(t+1)A - X(t) & * & * \\ B^{\mathrm{T}}X(t+1)A & B^{\mathrm{T}}X(t+1)B - G^{-1} & * \\ C & D & -\lambda I \end{pmatrix} \leqslant 0, \tag{4.7}$$
$$X(t_0) = R^{-1}, \quad \begin{pmatrix} X(t_f) & * \\ S^{1/2} & \lambda I \end{pmatrix} \geqslant 0, \quad t = t_0, \ldots, t_f - 1.$$

Now we proceed to filtering and control problems.

5. FILTERING

5.1. The Generalized H_2-optimal Filtering

Consider the filtering problem for a linear discrete object described by the difference equations

$$\begin{aligned} x(t+1) &= A(t)x(t) + B(t)v(t), \quad x(0) = x_0, \\ y(t) &= C(t)x(t) + D(t)v(t), \\ z(t) &= C_z(t)x(t), \quad t = t_0, \ldots, t_f, \end{aligned} \tag{5.1}$$

with the following notations: $x(t) \in \mathrm{R}^{n_x}$, $y(t) \in \mathrm{R}^{n_y}$, and $z(t) \in \mathrm{R}^{n_z}$ are the object's state, the measured output, and the target output, respectively; x_0 and $v(t) \in R^{n_v}$ are the initial state and the exogenous disturbance, respectively; finally, $A(t)$, $B(t)$, $C(t)$, and $D(t)$ are given matrices of compatible dimensions. To estimate the object's state by available output measurements, we construct the filter

$$\begin{aligned} x_f(t+1) &= A(t)x_f(t) + \Theta(t)[y(t) - C(t)x_f(t)], \quad x_f(0) = 0, \\ z_f(t) &= C_z(t)x_f(t), \end{aligned} \tag{5.2}$$

where $x_f(t) \in \mathrm{R}^{n_x}$ is the filter's state, $z_f(t) \in \mathrm{R}^{n_z}$ is the target output estimate, and $\Theta(t)$ is the filter parameter matrix. Introducing the state estimation error $e(t) = x(t) - x_f(t)$ and the output estimation error $e_z(t) = z(t) - z_f(t)$, from (5.1) and (5.2) we obtain the filtering error equations

$$\begin{aligned} e(t+1) &= A_c(t)e(t) + B_c(t)v(t), \quad e(0) = x_0, \\ e_z(t) &= C_z(t)e(t), \end{aligned} \tag{5.3}$$

where $A_c(t) = A(t) - \Theta(t)C(t)$ and $B_c(t) = B(t) - \Theta(t)D(t)$. Let the covariance matrix $K_{\xi(t)}$ of the random vector $\xi(t) = \mathrm{col}\,(x_0, v(t_0), \ldots, v(t-1))$ (the initial state and disturbance) be unknown. We find the parameters $\Theta_*(t)$ of the filter (5.2) that minimize the damping rate of the random disturbances with weight matrices $R > 0$ and $G(i) > 0$, $i = 1, \ldots, t-1$:

$$J_s(\Theta_{t_0}^{t-1}) = \sup_{K_{\xi(t)} \geqslant 0} \frac{E|e_z(t)|^2}{E\left(|x_0|_R^2 + \|v\|_{G\,[t_0,t]}^2\right)}, \tag{5.4}$$

where Θ_i^j denotes the set $\Theta(i), \ldots, \Theta(j)$.

According to Theorem 3.1, such a filter minimizes the maximum deviation of the error under deterministic factors, i.e.,

$$J_d(\Theta_{t_0}^{t-1}) = \sup_{x_0,\, v(\tau), \tau \in [t_0, t-1]} \frac{|e_z(t)|^2}{|x_0|_R^2 + \|v\|_{G\,[t_0,t]}^2},$$

and coincides with the filter minimizing the spectral radius of the covariance matrix of the output $e_z(t)$ in equation (5.3) provided that the initial state and disturbances form a sequence of random independent vectors with the covariance matrices R and $G(t)$. This means that its parameters are obtained from $\min_\Theta \lambda_{\max}(K_z(\Theta))$, where $K_z(\Theta) = C_z(t)P(t)C_z^{\mathrm{T}}(t)$ and $P(t)$ satisfies the equation

$$P(t+1) = [A(t) - \Theta(t)C(t)]P(t)[A(t) - \Theta(t)C(t)]^{\mathrm{T}} + [B(t) - \Theta(t)D(t)]G(t)[B(t) - \Theta(t)D(t)]^{\mathrm{T}}$$

with the initial condition $P(t_0) = R$. To simplify the formulas, we assume that the disturbances in the object and output measurements are independent and the matrices $D(t)$ have full rank, i.e.,

$$\begin{pmatrix} B(t) \\ D(t) \end{pmatrix} G(t) \begin{pmatrix} B(t) \\ D(t) \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} G_B(t) & 0 \\ 0 & G_D(t) \end{pmatrix}, \quad G_D(t) > 0. \tag{5.5}$$

Completing the square in $\Theta(t)$ on the right-hand side of the latter equation yields

$$\Theta_*(t) = A(t)P_2(t)C^{\mathrm{T}}(t)\left[C(t)P_2(t)C^{\mathrm{T}}(t) + G_D(t)\right]^{-1}, \tag{5.6}$$

where the matrix $P_2(t)$ satisfies the equation

$$P_2(t+1) = A(t)P_2(t)A^{\mathrm{T}}(t) - A(t)P_2(t)C^{\mathrm{T}}(t)\left[C(t)P_2(t)C^{\mathrm{T}} + G_D(t)\right]^{-1} C(t)P_2(t)A^{\mathrm{T}}(t) + G_B(t).$$

By the well-known matrix inversion formula (e.g., see [15, p. 254]), it can be written as

$$P_2(t+1) = A(t)\left[P_2^{-1}(t) + C^{\mathrm{T}}(t)G_D^{-1}(t)C(t)\right]^{-1} A^{\mathrm{T}}(t) + G_B(t) \tag{5.7}$$

with the initial condition $P_2(t_0) = R$.

Note that the resulting filter coincides with the Kalman filter [2] for estimating the state of system (5.1) under the random and independent initial state and disturbance with the covariance matrices R and $G(t)$ (the weight matrices). In other words, the Kalman filter constructed for the system with the random and independent initial state and disturbance with the covariance matrices R and $G(t)$ minimizes the maximum deviation of the error under deterministic factors satisfying the constraint

$$|x_0|_R^2 + \|v\|_{G\,[t_0,t]}^2 \leqslant 1 \tag{5.8}$$

and, moreover, minimizes the damping rate of the random disturbances (5.4) under the weight matrices R and $G(t)$. According to Theorem 2.2, the reachability set of the error $e_z(t)$ under (5.8) is the ellipsoid $\mathcal{E}_z[C_z(t)P_2(t)C_z^{\mathrm{T}}(t)]$; hence, the unknown vector $z(t)$ lies in the same ellipsoid centered at $z_f(t)$.

The parameters $\Theta_{g2}(t)$ of the generalized H_2-optimal filter minimizing the maximum variance of the estimation errors on the entire interval, i.e.,

$$J_{g2}(\Theta_{t_0}^{t_f-1}) = \max_{t\in[t_0,t_f]} \sup_{K_{\xi(t)}\geqslant 0} \frac{E|e_z(t)|^2}{E\left(|x_0|_R^2 + \|v\|_{G\,[t_0,t]}^2\right)},$$

are calculated in a standard way: it suffices to solve the linear matrix inequalities (3.8) (e.g., see [15]), replacing A by $A - \Theta(t)C$ and B by $B - \Theta(t)D$ and denoting $X(t+1)\Theta(t) = Z(t)$. Then the filter parameters are given by $\Theta_{g2}(t) = X_*^{-1}(t+1)Z_*(t)$, where the asterisks indicate the solutions of these inequalities.

5.2. The Generalized H_∞-optimal Filtering

Now consider the generalized H_∞-optimal filtering problem in the stochastic statement: under an unknown covariance matrix K_ξ of the initial state vector and disturbances $\xi = \mathrm{col}\,(x(t_0), v(t_0), \ldots, v(t_f-1))$ in (5.1), it is required to find the filter parameters (5.2) minimizing the performance criterion

$$J_{g\infty}(\Theta_{t_0}^{t_f-1}) = \sup_{K_\xi\geqslant 0} \frac{E\left[\|e_z\|_{[t_0,t_f]}^2 + e_x^{\mathrm{T}}(t_f)Se_x(t_f)\right]}{E\left[|x(t_0)|_R^2 + \|v\|_{G[t_0,t_f1]}^2\right]}$$

along the trajectories of system (5.3).

Theorem 5.1. *The generalized H_∞-norm of system* (5.3) *describing the estimation error dynamics for the state of* (5.1) *is smaller than λ, $J_{g\infty} < \lambda$, if the filter* (5.2) *has the parameters*

$$\Theta_\infty(t) = A(t)\left[P_\infty^{-1}(t) + C^{\mathrm{T}}(t)G_D^{-1}(t)C(t) - \lambda^{-1}C_z^{\mathrm{T}}(t)C_z(t)\right]^{-1} C^{\mathrm{T}}(t)G_D^{-1}(t), \tag{5.9}$$

where

$$P_\infty(t+1) = A(t)\left[P_\infty^{-1}(t) + C^{\mathrm{T}}(t)G_D^{-1}(t)C(t) - \lambda^{-1}C_z^{\mathrm{T}}(t)C_z(t)\right]^{-1} A^{\mathrm{T}}(t) + G_B(t), \tag{5.10}$$

$P_\infty(t_0) = R$, and $C_z(t)P_\infty(t)C_z^{\mathrm{T}}(t) < \lambda I$ and $S^{1/2}P_\infty(t_f)S^{1/2} < \lambda I$ for all t.

Equations (5.9) and (5.10) determine the so-called central generalized H_∞-suboptimal solutions. These equations have compact form and can be written without the inverse matrices $P_\infty^{-1}(t)$ by applying the matrix inversion formula. The parameters of the generalized H_∞-suboptimal filter ensuring $J_{g\infty} < \lambda$, particularly with the minimum value λ, can also be obtained in the standard way by solving the linear matrix inequalities (4.7).

A direct comparison of (5.10) and (5.7) shows that $P_\infty(t) \to P_2(t)$ as $\lambda \to \infty$. Passing the limit as $\lambda \to \infty$ in (5.9) and (5.10) and applying the matrix inversion formula, we easily establish the following result: the limit $\Theta_\infty(t)$ coincides with $\Theta_*(t)$ given by (5.6), and the generalized H_∞-suboptimal filter turns into the Kalman filter constructed for system (5.1) in which the independent initial state and disturbances have the covariance matrices R and $G(t)$, respectively. By tuning $0 < \lambda < \infty$, we can achieve a tradeoff between the variance of the state estimation error at a certain time instant and the sum of the variances of the target output estimation errors on the entire (preceding) time interval.

5.3. Optimal Estimation of Linear Regression Parameters

As one application we consider the optimal estimation of unknown parameters of the linear regression

$$\chi(t) = \Phi(t)\zeta_0 + v(t), \quad t = t_0, \dots, t_f, \tag{5.11}$$

with the following notations: $\chi(t)$ is the measurement vector, $\Phi(t)$ is the regressor matrix, ζ_0 is the unknown parameter vector, and $v(t)$ is the measurement noise vector. We represent this problem as designing an optimal observer for the system state

$$\zeta(t+1) = \zeta(t), \quad \zeta(t_0) = \zeta_0 \tag{5.12}$$

of the form

$$\widehat{\zeta}(t+1) = \widehat{\zeta}(t) + \Theta(t)[\chi(t) - \Phi(t)\widehat{\zeta}(t)], \quad \widehat{\zeta}(t_0) = \zeta_*.$$

The error $\widetilde{\zeta}(i) = \zeta_0 - \widehat{\zeta}(i)$ satisfies the equation

$$\widetilde{\zeta}(i+1) = [I - \Theta(i)\Phi(i)]\widetilde{\zeta}(i) - \Theta(i)v(i), \quad \widetilde{\zeta}(t_0) = \zeta_0 - \zeta_*, \quad i = t_0, \dots, t-1.$$

Under an unknown covariance matrix of the vector composed of the initial error $\zeta_0 - \zeta_*$ and the measurement noises on the entire time interval, the matrix $\Theta_*(t) = P_2(t)\Phi(t)\big[\Phi(t)P_2(t)\Phi^{\mathrm{T}}(t) + G(t)\big]^{-1}$ yields the parameters of the optimal filter (5.2), (5.6) minimizing the damping rate of random disturbances at a time instant t for the error equation. (This fact can be easily verified.) Therefore, the optimal estimates are given by the recurrence equations

$$\begin{gathered}\widehat{\zeta}(t+1) = \widehat{\zeta}(t) + P_2(t+1)\Phi^{\mathrm{T}}(t)G^{-1}(t)[\chi(t) - \Phi(t)\widehat{\zeta}(t)], \quad \widehat{\zeta}(t_0) = \zeta_*, \\ P_2^{-1}(t+1) = P_2^{-1}(t) + \Phi^{\mathrm{T}}(t)G^{-1}(t)\Phi(t), \quad P_2(t_0) = R.\end{gathered} \tag{5.13}$$

As is well known, these equations describe the recurrent modification of the weighted least squares method and the Kalman filter for estimating the state of system (5.12) under the covariances $E\zeta_0\zeta_0^{\mathrm{T}} = R$ and $Ev(t)v^{\mathrm{T}}(t) = G(t)$; moreover, the estimate $\widehat{\zeta}(t)$ minimizes the performance criterion

$$J_t(\zeta) = (\zeta - \zeta_*)^{\mathrm{T}}R^{-1}(\zeta - \zeta_*) + \sum_{i=0}^{t-1}\big(\chi(i) - \Phi(i)\zeta\big)^{\mathrm{T}}G^{-1}(i)\big(\chi(i) - \Phi(i)\zeta\big).$$

Thus, the weighted least squares method defines an estimate minimizing the damping rate of random disturbances with appropriate weight matrices. According to the duality principle, this estimate minimizes the maximum deviation of the error under unknown deterministic parameters and disturbances satisfying the constraint

$$|\zeta_0 - \zeta_*|_R^2 + \|v\|_{G\,[t_0,t]}^2 \leqslant 1, \quad t \in [t_0, t_f]. \tag{5.14}$$

Finally, we construct a generalized H_∞-suboptimal filter to estimate the unknown parameters in the linear regression (5.11) under which the sum of squared errors on the entire time interval will not exceed, with a multiplier λ, the weighted sum of the squared initial deviation of the estimate and squared measurement noises. Due to (5.9) and (5.10), the parameters of this filter are given by

$$\Theta_\infty(t) = \Big[P_\infty(t) + \Phi^{\mathrm{T}}(t)G^{-1}(t)\Phi(t) - \lambda^{-1}I\Big]^{-1} \Phi^{\mathrm{T}}(t)G^{-1}(t);$$

under the conditions $P_\infty(t) < \lambda I$ and $S^{1/2}P_\infty(t_f)S^{1/2} < \lambda I$, the filter satisfies the recurrent equations

$$\begin{aligned}\widehat{\zeta}(t+1) &= \widehat{\zeta}(t) + P_\infty(t+1)\Phi^{\mathrm{T}}(t)G^{-1}(t)[\chi(t) - \Phi(t)\widehat{\zeta}(t)], \quad \widehat{\zeta}(t_0) = \zeta_*,\\ P_\infty^{-1}(t+1) &= P_\infty^{-1}(t) + \Phi^{\mathrm{T}}(t)G^{-1}(t)\Phi(t) - \lambda^{-1}I, \quad P_\infty(t_0) = R.\end{aligned} \tag{5.15}$$

As $\lambda \to \infty$ the equations of the generalized H_∞-optimal filter turn into the recurrent equations of the weighted least squares method. By tuning the value λ, we can achieve a trade-off between the error variance at the end of the time interval and the sum of the error variances on the entire time interval.

6. CONTROL

Consider the control problem for a linear discrete-time object described by the difference equations

$$\begin{aligned}x(t+1) &= A(t)x(t) + B_u(t)u(t) + B(t)v(t), \quad x(0) = x_0,\\ z(t) &= C_z(t)x(t) + D_z(t)u(t), \quad t = t_0, \ldots, t_f.\end{aligned} \tag{6.1}$$

Let the initial state and exogenous disturbances form a random vector $\xi(t) = \mathrm{col}\,(x_0, v(t_0), \ldots, v(t-1))$ with zero mean and an unknown covariance matrix $K_{\xi(t)}$. A controller of the form $u(t) = \Theta(t)x(t)$ is required to minimize the performance criterion

$$J_s(\Theta_{t_0}^{t-1}) = \sup_{K_{\xi(t)} \geqslant 0} \frac{E|z(t)|^2}{E\left(|x_0|_R^2 + \|v\|_{G\,[t_0,t]}^2\right)}.$$

By Theorem 3.1, such a controller minimizes the maximum output deviation under deterministic factors:

$$J_d(\Theta_{t_0}^{t-1}) = \sup_{x_0,\, v(\tau), \tau\in[t_0,t-1]} \frac{|z(t)|^2}{|x_0|_R^2 + \|v\|_{G\,[t_0,t]}^2};$$

moreover, it coincides with the controller minimizing the spectral radius of the output covariance matrix of system (6.1) at a time instant t provided that the initial state and disturbances form a sequence of random independent vectors with the covariance matrices R and $G(i)$, $i = t_0, \ldots, t-1$. The parameters of such a controller can be found by solving a semidefinite programming problem under constraints defined by linear matrix inequalities obtained in the standard way from (3.8). Note that control problems with a performance criterion formulated in terms of the covariance matrix of the steady-state output on an infinite horizon were considered in [16].

Now let the initial state and disturbances on the entire time interval $[t_0, t_f]$ form a random vector $\xi = \mathrm{col}\,(x_0, v(t_0), \ldots, v(t_f-1))$ with zero mean and an unknown covariance matrix K_ξ. A controller of the form $u(t) = \Theta(t)x(t)$ is required to minimize the total output "energy" on the entire time interval considering the terminal state provided that the total "energy" of the initial state and the exogenous disturbance are bounded on the entire time interval. In this case, the "energy" is measured by the expected value of the quadratic form of the corresponding vector with a given weight matrix. Therefore, the performance criterion has the form

$$J_s(\Theta_{t_0}^{t_f-1}) = \sup_{K_\xi \geqslant 0} \frac{E\left[\|z\|^2_{[t_0,t_f]} + x^{\mathrm{T}}(t_f)Sx(t_f)\right]}{E\left[|x_0|^2_R + \|v\|^2_{G[t_0,t_f]}\right]}, \tag{6.2}$$

where $S = S^{\mathrm{T}} \geqslant 0$, $R = R^{\mathrm{T}} > 0$, and $G(t) = G^{\mathrm{T}}(t) > 0$ are given weight matrices. Note that in the classical stochastic linear-quadratic control problem, the initial state vectors and disturbances are assumed to form a sequence of independent random vectors with given covariance matrices.

According to Theorem 4.1, this problem is dual to the control problem of the object (6.1) in which the initial state and external disturbances form an arbitrary deterministic vector sequence and the performance criterion is given by the generalized H_∞ norm with the corresponding weight matrices:

$$J_d(\Theta_{t_0}^{t_f-1}) = \sup_{x_0, v(\tau), \tau\in[t_0,t_f-1]} \frac{\|z\|^2_{[t_0,t_f]} + x^{\mathrm{T}}(t_f)Sx(t_f)}{|x_0|^2_R + \|v\|^2_{G[t_0,t_f]}}. \tag{6.3}$$

The parameters of the desired controller are found by solving a semidefinite programming problem obtained from (4.7) in the standard way.

7. CONCLUSIONS

We have established the duality principle for linear operators in the deterministic and stochastic cases. This result is useful due to relating the stochastic and deterministic paradigms in control and filtering problems. In particular, let the deterministic initial state and disturbance in a linear time-varying system on a finite horizon satisfy an ellipsoidal constraint with given weight matrices; in this case, the maximum value (over time) among the maximum output deviations, i.e., the generalized H_2-norm of the system, and the maximum value of the integral quadratic performance criterion, i.e., the generalized H_∞-norm of the system, will coincide with the maximum value (over time) among the maximum output variances and the maximum value of the averaged integral quadratic performance criterion, respectively, under the random initial state and disturbance with unknown covariance matrices satisfying the averaged ellipsoidal constraint. Both of these norms are also characterized as spectral norms of the covariance matrices of the outputs of linear operators under random and independent initial states and disturbances of the system in which the covariance matrices coincide with the corresponding weight matrices of the ellipsoidal constraint.

We have formulated new minimax problems for linear dynamic systems in the stochastic statement with unknown covariance matrices of random factors. As demonstrated above, their solutions coincide with the solutions of dual deterministic minimax problems. For example, the minimax stochastic controller under an unknown joint covariance matrix of the initial state and disturbances with a bounded trace coincides with the deterministic generalized H_∞-optimal controller. The optimal filter minimizing the damping rate of the random initial state and disturbances with an unknown joint covariance matrix coincides with the optimal filter minimizing the maximum deviation of the filtering error under the deterministic initial state and disturbances satisfying an ellipsoidal constraint with given weight matrices. This filter turns out to be a Kalman filter constructed for this system under the random and independent initial state and disturbances whose covariance matrices are equal to the corresponding weight matrices of the ellipsoidal constraint.

APPENDIX

Proof of Theorem 4.1. We write the Lagrange function for problem S_∞ and find its dual function:

$$\min_{\lambda\geqslant 0, X(t)} \max_{W(t)\geqslant 0} \sum_{t=t_0}^{t_f-1} \operatorname{tr}(C(t)\ D(t))W(t)(C(t)\ D(t))^{\mathrm{T}} + \operatorname{tr} SMW(t_f)M^{\mathrm{T}}$$

$$-\lambda\left[\operatorname{tr} R^{-1}MW(t_0)M^{\mathrm{T}} + \sum_{t=t_0}^{t_f-1}\operatorname{tr} G^{-1}(t)HW(t)H^{\mathrm{T}} - 1\right]$$

$$+\sum_{t=t_0}^{t_f-1}\operatorname{tr}\Big[(A(t)\ B(t))W(t)(A(t)\ B(t))^{\mathrm{T}} - MW(t+1)M^{\mathrm{T}}\Big]X(t+1)$$

$$= \min_{\lambda\geqslant 0, X(t)} \max_{W(t)\geqslant 0}\left\{\lambda + \sum_{t=t_0}^{t_f-1}\operatorname{tr} W(t)\Big[(C(t)\ D(t))^{\mathrm{T}}(C(t)\ D(t)) + (A(t)\ B(t))^{\mathrm{T}}X(t+1)(A(t)\ B(t))\right.$$

$$\left. - M^{\mathrm{T}}X(t)M - \lambda H^{\mathrm{T}}G^{-1}(t)H\Big] + \operatorname{tr} W(t_f)M^{\mathrm{T}}[S - X(t_f)]M\right\},$$

where $X(t_0) = \lambda R^{-1}$. The dual function is finite under the following inequalities:

$$\begin{gathered}(C(t)\ D(t))^{\mathrm{T}}(C(t)\ D(t)) + (A(t)\ B(t))^{\mathrm{T}}X(t+1)(A(t)\ B(t)) \\ - M^{\mathrm{T}}X(t)M - \lambda H^{\mathrm{T}}G^{-1}(t)H \leqslant 0,\ t = t_0,\dots,t_f-1,\ S - X(t_f) \leqslant 0.\end{gathered} \tag{A.1}$$

(Otherwise, $W(t)$ can be chosen so that the corresponding term will become infinite.) Thus, inequalities (A.1) must hold, but in this case, the minimum in the minimax problem is reached at $W(t) = 0$, $t = t_0,\dots,t_f$. As a result, we arrive at the dual problem: $\min\lambda$ subject to the constraints (A.1). With the introduced notations and the variable $X(t)$ replaced by $\lambda X(t)$, these constraints are reduced to inequalities (4.7). Since the function is convex and there exists an interior point satisfying the constraints, the values of the primal and dual problems coincide.

We define the function $V(t) = x^{\mathrm{T}}(t)X(t)x(t)$, where $X(t)$ satisfies inequalities (4.7). The increment of this function along the trajectories of system (4.1) satisfies the conditions

$$\begin{gathered}\Delta V(t) + \lambda^{-1}|z(t)|^2 - v^{\mathrm{T}}(t)G^{-1}v(t) \leqslant 0, \\ V(t_0) = x^{\mathrm{T}}(t_0)R^{-1}x(t_0), \quad V(t_f) \geqslant \lambda^{-1}x^{\mathrm{T}}(t_f)Sx(t_f).\end{gathered} \tag{A.2}$$

Hence,

$$\sum_{t=t_0}^{t_f-1}|z(t)|^2 + x^{\mathrm{T}}(t_f)Sx(t_f) \leqslant \lambda + \lambda\left[x(t_0)R^{-1}x(t_0) + \sum_{t=t_0}^{t_f-1} v^{\mathrm{T}}(t)G^{-1}(t)v(t) - 1\right],$$

i.e., the minimum value λ making inequalities (4.7) solvable is the optimal value in problem D_∞ and coincides with the generalized H_∞ norm of system (4.1).

Proof of Theorem 5.1. Let us apply Theorem 4.1 to system (5.3): if inequalities (4.7) hold with the matrix A replaced by $A - \Theta C$ and the matrix B replaced by $B - \Theta C$, then the generalized H_∞ norm of this system is smaller than λ. Using Schur's complement lemma, we transform these inequalities to

$$\begin{gathered}Y(t+1) - (A-\Theta C)Y(t)(A-\Theta C)^{\mathrm{T}} - (B-\Theta D)G(B-\Theta D)^{\mathrm{T}} \\ - (A-\Theta C)Y(t)C_z^{\mathrm{T}}(\lambda I - C_zY(t)C_z^{\mathrm{T}})^{-1}C_zY(t)(A-\Theta C)^{\mathrm{T}} \geqslant 0\end{gathered}$$

provided that $C_zY(t)C_z^{\mathrm{T}} < \lambda I$. Completing the square in $\Theta(t)$ on the left-hand side of the latter inequality yields

$$Y(t+1) - A\left[Y^{-1}(t) + C^{\mathrm{T}}G_D^{-1}C - \lambda^{-1}C_z^{\mathrm{T}}C_z\right]^{-1}A^{\mathrm{T}} - G_B$$
$$-(\Theta - \Theta_\infty)\left[Y^{-1}(t) + C^{\mathrm{T}}G_D^{-1}C - \lambda^{-1}C_z^{\mathrm{T}}C_z\right](\Theta - \Theta_\infty)^{\mathrm{T}} \geqslant 0,$$

where Θ_∞ is given by (5.9) for $P(t) = Y(t)$. (Here, we have involved the notations (5.5) and some manipulations.) Hence, if the filter parameters are given by (5.9), where the matrix $P(t)$ satisfies equation (5.10), then $\gamma_s^2 < \lambda$.

FUNDING

This work was supported by the Scientific and Educational Mathematical Center "Mathematics of Future Technologies," agreement no. 075-02-2021-1394.

REFERENCES

1. Kurzhanski, A.B., *Upravlenie i nablyudenie v usloviyakh neopredelennosti* (Control and Observation under Uncertainty), Moscow: Nauka, 1977.
2. Kwakernaak, H. and Sivan, R., *Linear Optimal Control Systems*, 1st ed., Wiley-Interscience, 1972.
3. Kailath, T., Sayed, A.H., and Hassibi, B., *Linear Estimation*, New Jersey: Prentice Hall, 2000.
4. Hinrichsen, D. and Pritchard, A.J., Stochastic H_∞, *SIAM J. Control Optim.*, 1998, vol. 36, no. 5, pp. 1504–1538.
5. Petersen, I.R., Ugrinovskii, V.A., and Savkin, A.V., *Robust Control Design Using H_∞-Methods*, London: Springer-Verlag, 2000.
6. Schweppe, F.C., Recursive State Estimation: Unknown but Bounded Errors and System Inputs, *IEEE Trans. Autom. Control*, 1968, vol. 13, no. 1, pp. 22–28.
7. Wilson, D., Extended Optimality Properties of the Linear Quadratic Regulator and Stationary Kalman Filter, *IEEE Trans. Autom. Control*, 1990, vol. 35, no. 5, pp. 583–585.
8. Willems, J.C., Deterministic Least Squares Filtering, *Journal of Econometrics*, 2004, vol. 118, pp. 341–373.
9. Buchstaller, D., Liu, J., and French, M., The Deterministic Interpretation of the Kalman Filter, *Int. J. Control*, 2021, vol. 94, no. 11, pp. 3226–3236.
10. Kogan, M.M., Optimal Estimation and Filtration under Unknown Covariances of Random Factors, *Autom. Remote Control*, 2014, vol. 75, no. 11, pp. 1964–1981.
11. Boyd, S. and Vandenberghe, L., *Convex Optimization*, Cambridge: University Press, 2004.
12. Chernousko, F.L., *State Estimation for Dynamic Systems*, CRC Press, 1993.
13. Wilson, D.A., Convolution and Hankel Operator Norms for Linear Systems, *IEEE Trans. Autom. Control*, 1989, vol. 34, no. 1, pp. 94–97.
14. Balandin, D.V., Biryukov, R.S., and Kogan, M.M., Minimax Control of Deviations for the Outputs of a Linear Discrete Time-Varying System, *Autom. Remote Control*, 2019, vol. 80, no. 12, pp. 345–359.
15. Balandin, D.V. and Kogan, M.M., *Sintez zakonov upravleniya na osnove lineinykh matrichnykh neravenstv* (Control Law Design Based on Linear Matrix Inequalities), Moscow: Fizmatlit, 2007.
16. Hsieh, C. and Skelton, R., All Covariance Controllers for Linear Discrete-Time Systems, *IEEE Trans. Autom. Control*, 1990, vol. 35, no. 8, pp. 908–915.

This paper was recommended for publication by B.M. Miller, a member of the Editorial Board