ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Systems & Control Letters ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

SYSTEMS & CONTROL LETTERS

www.elsevier.com/locate/sysconle

Observer matrix gain optimization for stochastic continuous time nonlinear systems

Alex Poznyak*, Alexander Nazin[1], Daishi Murano

Department of Automatic Control, CINVESTAV-IPN, Av. 2508, AP14-740, CP07360 Mexico DF, Mexico

Received 21 August 2002; received in revised form 1 December 2003; accepted 20 February 2004

Abstract

It is shown that for a class of stationary stochastic nonlinear systems (satisfying a global Lipschitz condition) the high-gain observer with a constant gain matrix may guarantee an upper bound for the averaged quadratic error of state estimation. The nonlinearity is assumed to be a priory known. The main contribution of this paper consists in designing of a numerical procedure for the optimal gain matrix minimizing this upper bound. The convergence analysis of this procedure is presented as well as an example illustrating its finite steps workability: it is shown that within a neighborhood of the optimal matrix gain the others provide lower estimation performance.

Keywords: Nonlinear observers; Stochastic systems; Gain matrix optimization

1. Introduction

When the exact and complete knowledge on current states of a dynamic plant is impossible by different reasons (stochastic disturbances, direct sensors absence, etc.), the use of a state estimator (observer) is compulsory to realize a successful closed-loop control. The publications dealing with this concept usually tackle two types of dynamic models: deterministic and stochastic.

Deterministic observers: In the situation when the plant model is incomplete or uncertain, the implementation of high-gain observers seems to be convenient [26,10,14,2]. Most of the known results, dealing with this technique, assume that the system neither contains any external disturbances nor has internal unmodeled dynamics. But if such uncertainties and disturbances take place, then another *min–max approach* is usually applied where the “max”-operation is taken over the set of uncertainty and the “min” corresponds to the construction of the best state estimation. Usually the exact solution of this min–max problem is extremely hard task, that’s why the approach related to the constructions of estimations which minimize an error upper bound is mostly used [3,4,9,17,18,21–25,30].

Stochastic observers: For the class of stochastic nonlinear dynamic models, even the complete

* Corresponding author. Tel.: +52-5557473741; fax: +52-5557477089.

E-mail addresses: apoznyak@ctrl.cinvestav.mx (A. Poznyak), nazine@ipu.rssi.ru (A. Nazin), dmurano@ctrl.cinvestav.mx (D. Murano).

[1] The work of this author has been carried out during his 3 months visit at the end of 2001 to the Department of Control Automatico, CINVESTAV-IPN, Mexico, DF (on leave from the Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia).

0167-6911/$ - see front matter
doi:10.1016/j.sysconle.2004.02.013

information is a priory available, the situation with optimal state estimation turns out to be much more complex since it is related with the solution of the, so-called, Duncan–Mortensen–Zakai partial differential equations [33], which has a finite-dimensional solution only for linear models that corresponds to well-known Kalman filter. Any finite-dimensional observer design demands some sort of approximations (as well as for the deterministic case) related to a *minimization of an upper error bound* [6,13,31,20,19,32,8,12,1,5,15,16,27–29]. None of these papers discusses the problem of the best (from some point of view) selection of the parameters of the suggested observer (filters).

In this paper, we follow the approach developed in [16] and show that for stationary stochastic nonlinear systems (satisfying a globally Lipschitz condition) the high-gain observer with a constant gain matrix may guarantee a tight (sharp) upper bound for the averaged quadratic error of state estimation. The nonlinearity in the right-hand side of the model dynamics are assumed to be a priory known. *The main contribution of this paper consists in the designing of a numerical procedure for the optimal gain matrix construction, which minimizes this upper bound.* The convergence analysis of this procedure is presented as well as an example illustrating its finite steps workability: it is shown that within a neighborhood of the optimal matrix gain the others provide less estimation performance.

2. Nonlinear stochastic system and observer

Let $(\Omega, F, \{F_t\}_{t\geqslant 0}, P)$ be a given filtered probability space with a complete probability space (Ω, F, P); σ-algebra F_0 contains all the P-null sets from F, the filtration $\{F_t\}_{t\geqslant 0}$ is right continuous, that is $F_{t+} := \bigcap_{s>t} F_s = F_t$. Define an m-dimensional standard Brownian motion, that is $\{F_t\}_{t\geqslant 0}$-adapted, R^m-valued random process $(\bar{W}_t, t \geqslant 0)$ (with $\bar{W}_0 = 0$) such that

$$E\{\bar{W}_t - \bar{W}_s | F_s\} = 0, \quad \bar{W}_0 = 0 \quad \text{a.s.},$$

$$E\{[\bar{W}_t - \bar{W}_s][\bar{W}_t - \bar{W}_s]^{\mathsf{T}} | F_s\} = (t-s)I \quad \text{a.s.}, \ s \leqslant t.$$

Consider stochastic nonlinear continuous-time system with the state dynamics x_t and the observable output y_t given by

$$x_t = x_0 + \int_{s=0}^{t} [Ax_s + f(s, x_s)]\,\mathrm{d}s + \int_{s=0}^{t} \sigma^x(s, x_s)\,\mathrm{d}\bar{W}_s^x,$$

$$y_t = y_0 + \int_{s=0}^{t} [Cx_s + h(s, x_s)]\,\mathrm{d}s + \int_{s=0}^{t} \sigma^y(s, x_s)\,\mathrm{d}\bar{W}_s^y, \tag{1}$$

or, in the abstract (symbolic) form

$$\mathrm{d}x_t = [Ax_t + f(t, x_t)]\,\mathrm{d}t + \sigma^x(t, x_t)\,\mathrm{d}\bar{W}_t^x,$$

$$x_0 = x, \quad t \in [0, \infty),$$

$$\mathrm{d}y_t = [Cx_t + h(t, x_t)]\,\mathrm{d}t + \sigma^y(t, x_t)\,\mathrm{d}\bar{W}_t^y. \tag{2}$$

The first integral in (1) is a stochastic ordinary integral and the second one is an Itô integral. In the above, $x_t \in R^n$ is system state at time t, the initial state $x_0 = x$ is assumed to be given (may be, random), $y_t \in R^k$ is the output, $A \in R^{n\times n}$ and $C \in R^{k\times n}$ are given matrices (i.e., a priory known), $f : [0,\infty) \times R^n \to R^n$, $\sigma^x : [0,\infty) \times R^n \to R^{n\times m}$, $\sigma^y : [0,\infty) \times R^n \to R^{k\times m}$ and $h : [0,\infty) \times R^n \to R^k$. Below we use the following notations:

$$\bar{W}_t := \begin{bmatrix} \bar{W}_t^x \\ \bar{W}_t^y \end{bmatrix}.$$

It is assumed that

A1: $\{F_t\}_{t\geqslant 0}$ is the natural filtration generated by $(\bar{W}_t, t \geqslant 0)$ and augmented by the P-null sets from F.

Definition 1. Let L_g be a finite nonnegative constant and Λ_g be a positive definite $n \times n$-matrix. Borel function $g : [0,\infty) \times R^n \to R^n$ is said to belong to the class $\mathbf{F}_2(L_g, \Lambda_g)$ if it is C^2 (twice continuous differentiable) in x (almost everywhere with respect to the Lebesgue measure) for any $t \in [0,\infty)$, $g(t,0) \equiv 0$, and for all $t \in [0,\infty)$ and $x, \hat{x} \in R^n$ the following Lipschitz condition holds:

$$\|g(t,x) - g(t,\hat{x})\|_{\Lambda_g}^2 + \|g_x(t,x) - g_x(t,\hat{x})\|_{\Lambda_g}^2 \leqslant L_g \|x - \hat{x}\|_{\Lambda_g}^2.$$

Here $g_x(t,x)$ stands for the gradient of vector-function $g(t,x)$ by x.

Consider observer given in the following Luenberger-like form:

$$d\hat{x}_t = [A\hat{x}_t + f(t,\hat{x}_t)]\,dt + K_t[dy_t - (C\hat{x}_t + h(t,\hat{x}_t))\,dt] \tag{3}$$

with $\hat{x}_0$ as an arbitrary fixed initial condition. Here K_t stands for a matrix gain which is measured with respect to the σ-algebra

$$G_t := \sigma\{y_\tau, \hat{x}_\tau \mid \tau \in [0,t]\}. \tag{4}$$

Denote the estimation error by

$$\Delta_t := \hat{x}_t - x_t, \tag{5}$$

whose dynamics, in view of (2) and (3), is governed by

$$\begin{aligned} d\Delta_t &= A\Delta_t\,dt + [f(t,\hat{x}_t) - f(t,x_t)]\,dt \\ &\quad + K_t[dy_t - (C\hat{x}_t + h(\hat{x}_t,t))\,dt] - \sigma^x(t,x_t)\,d\bar{W}_t^x \\ &= A\Delta_t\,dt + [f(t,\hat{x}_t) - f(t,x_t)]\,dt + K_t[-(C\Delta_t \\ &\quad + h(\hat{x}_t,t) - h(t,x_t))\,dt] + \sigma(t,x_t)\,d\bar{W}_t, \end{aligned} \tag{6}$$

where

$$\sigma(t,x_t) := [-\sigma^x(t,x_t),\ K_t\sigma^y(t,x_t)]. \tag{7}$$

To analyze the behavior of the state estimate (5), introduce the Lyapunov function given by

$$V_t = \Delta_t^{\mathsf{T}} P_t \Delta_t \tag{8}$$

with $0 \leqslant P_t = P_t^{\mathsf{T}} \in R^{n\times n}$. By the Itô formula [7] and in view of (6), it follows:

$$dV_t = \Delta_t^{\mathsf{T}} \dot{P}_t \Delta_t\,dt + 2\langle P_t\Delta_t, d\Delta_t\rangle + \mathrm{tr}\{\sigma(t,x_t)\sigma(t,x_t)^{\mathsf{T}} P_t\}\,dt \tag{9}$$

The direct substitution of (6), leads to following expression:

$$\begin{aligned} \langle P_t\Delta_t, d\Delta_t\rangle &= \langle P_t\Delta_t, A\Delta_t\,dt + [f(t,\hat{x}_t) - f(t,x_t)]\,dt \\ &\quad + K_t[-(C\Delta_t + h(\hat{x}_t,t) - h(t,x_t))\,dt] \\ &\quad + \sigma(t,x_t)\,d\bar{W}_t\rangle. \end{aligned} \tag{10}$$

Let us analyze each term in (10).

1. The following identity holds:
$$2\langle P_t\Delta_t, A\Delta_t\rangle = \Delta_t^{\mathsf{T}}(P_tA + A^{\mathsf{T}}P_t)\Delta_t.$$
2. For any positive definite matrix Λ_f and for any $f \in \mathbf{F}_2(L_f,\Lambda_f)$, it follows:
$$\begin{aligned} &2\langle P_t\Delta_t, [f(t,\hat{x}_t) - f(t,x_t)]\rangle \\ &\quad\leqslant \Delta_t^{\mathsf{T}} P_t \Lambda_f^{-1} P_t \Delta_t + \|f(t,\hat{x}_t) - f(t,x_t)\|^2_{\Lambda_f} \\ &\quad\leqslant \Delta_t^{\mathsf{T}}[P_t\Lambda_f^{-1}P_t + L_f\Lambda_f]\Delta_t. \end{aligned}$$
3. Furthermore,
$$\begin{aligned} &2\langle P_t\Delta_t, K_t[-(C\Delta_t + h(\hat{x}_t,t) - h(t,x_t))]\rangle \\ &\quad= -\Delta_t^{\mathsf{T}}[P_tK_tC + (K_tC)^{\mathsf{T}}P_t]\Delta_t\,dt \\ &\qquad + 2\langle P_t\Delta_t, [h(t,x_t) - h(\hat{x}_t,t)]\,dt\rangle. \end{aligned}$$
Again, by the same inequality, for any positive definite matrix Λ_h and for any $h \in \mathbf{F}_2(L_h,\Lambda_h)$, we derive
$$\begin{aligned} &2\langle K_t^{\mathsf{T}} P_t\Delta_t, h(t,x_t) - h(\hat{x}_t,t)\rangle \\ &\quad\leqslant \Delta_t^{\mathsf{T}} P_tK_t\Lambda_h^{-1}K_t^{\mathsf{T}}P_t\Delta_t \\ &\qquad + \|h(t,x_t) - h(\hat{x}_t,t)\|^2_{\Lambda_h} \\ &\quad\leqslant \Delta_t^{\mathsf{T}}[P_tK_t\Lambda_h^{-1}K_t^{\mathsf{T}}P_t + L_h\Lambda_h]\Delta_t. \end{aligned}$$

If take P_t verifying the differential Riccati equation

$$\begin{aligned} -\dot{P}_t &= P_t(A - K_tC) + (A - K_tC)^{\mathsf{T}}P_t \\ &\quad + P_t(\Lambda_f^{-1} + K_t\Lambda_h^{-1}K_t^{\mathsf{T}})P_t + (L_f\Lambda_f + L_h\Lambda_h) \\ &\quad + Q_0, \end{aligned} \tag{11}$$

then, the direct application of all inequalities given above to (9) implies

$$dV_t \leqslant -\Delta_t^{\mathsf{T}} Q_0 \Delta_t\,dt + 2\langle P_t\Delta_t, \sigma(t,x_t)\,d\bar{W}_t\rangle + \mathrm{tr}\{\sigma(t,x_t)\sigma(t,x_t)^{\mathsf{T}}P_t\}\,dt.$$

Integrating this inequality over the time interval $[0,t]$ we obtain

$$\begin{aligned} &\frac{1}{t}\int_0^t \Delta_\tau^{\mathsf{T}} Q_0 \Delta_\tau\,d\tau \\ &\quad\leqslant \frac{1}{t}\int_0^t \mathrm{tr}\{\sigma(\tau,x_\tau)\sigma(\tau,x_\tau)^{\mathsf{T}}P_\tau\}\,d\tau \end{aligned}$$

$$+2\,\frac{1}{t}\int_0^t \langle P_\tau \Delta_\tau, \sigma(\tau, x_\tau)\,\mathrm{d}\bar{W}_\tau\rangle + \frac{1}{t}(V_0 - V_t).$$

Since P_t is supposed to be a stable solution of (11), the process is quadratically integrable by Lemma 2 in [16], it follows that:

$$\frac{1}{t}\int_0^t \langle P_\tau \Delta_\tau, \sigma(\tau, x_\tau)\,\mathrm{d}W_\tau\rangle \xrightarrow[n\to\infty]{\text{a.s.}} 0$$

and, as a result, the upper bound to the averaged quadratic state estimation error is as follows:

$$\begin{aligned} J &:= \limsup_{t\to\infty} \frac{1}{t}\int_0^t \Delta_\tau^{\mathsf{T}} Q_0 \Delta_\tau\,\mathrm{d}\tau \\ &\leqslant \limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathrm{tr}\{\sigma(\tau, x_\tau)\sigma(\tau, x_\tau)^{\mathsf{T}} P_\tau\}\,\mathrm{d}\tau \\ &\overset{\text{a.s.}}{\leqslant} \limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathrm{tr}\{[-\sigma^x(t, x_t), K_t\sigma^y(t, x_t)] \\ &\qquad [-\sigma^x(t, x_t),\ K_t\sigma^y(t, x_t)]^{\mathsf{T}} P_\tau\}\,\mathrm{d}\tau. \end{aligned} \tag{12}$$

3. Stationary noise case

3.1. The upper bound for the state estimation error

If $\sigma^x(t, x_t)$ and $\sigma^y(t, x_t)$ are constant matrices, that is,

$$\sigma^x(t, x_t) = \sigma^x = \text{const}, \quad \sigma^y(t, x_t) = \sigma^y = \text{const} \tag{13}$$

and the gain matrix is chosen stationary as well, i.e.,

$$K_t := K = \text{Const}, \tag{14}$$

then the upper bound J_+ in (12) can be expressed as follows:

$$\begin{aligned} J \overset{\text{a.s.}}{\leqslant} J_+ &:= \mathrm{tr}\{[-\sigma^x,\ K\sigma^y][-\sigma^x, K\sigma^y]^{\mathsf{T}} P\} \\ &= \mathrm{tr}\{[\sigma^x\sigma^{x\mathsf{T}} + K\sigma^y\sigma^{y\mathsf{T}}K^{\mathsf{T}}]P\}, \end{aligned} \tag{15}$$

where K is selected in such a way to guarantee the existence of a nonnegative solution P (non obligatory a strictly positive one) the following algebraic Riccati equation:

$$P(A - KC) + (A - KC)^{\mathsf{T}} P + PRP + Q = 0,$$

$$R := \Lambda_f^{-1} + K\Lambda_h^{-1}K^{\mathsf{T}},$$

$$Q := L_f\Lambda_f + L_h\Lambda_h + Q_0. \tag{16}$$

3.2. Gain-matrix optimization

Let us formulate the following optimization problem:

$$J_+ := \mathrm{tr}\{[\sigma^x\sigma^{x\mathsf{T}} + K\sigma^y\sigma^{y\mathsf{T}}K^{\mathsf{T}}]P\} \to \min_K, \tag{17}$$

subjected to constraint (16) together with the condition of P to be nonnegative definite.

To obtain an optimality condition, define the corresponding Lagrange function as

$$\begin{aligned} L(K, P, \Lambda) := J_+ - \mathrm{tr}\{[&P(A - KC) \\ &+ (A - KC)^{\mathsf{T}} P + PRP + Q]\Lambda\}. \end{aligned} \tag{18}$$

Then using the following rules:

$$\frac{\partial}{\partial Y}\mathrm{tr}\{X^{\mathsf{T}} Y\} = X, \quad \frac{\partial}{\partial Y}\mathrm{tr}\{X^{\mathsf{T}}(A + YB)\} = XB^{\mathsf{T}},$$

$$\begin{aligned} \frac{\partial}{\partial K} J_+ &= \frac{\partial}{\partial K}\mathrm{tr}\{[-\sigma^x, K\sigma^y][-\sigma^x, K\sigma^y]^{\mathsf{T}} P\} \\ &= 2\sigma^y(\sigma^y)^{\mathsf{T}} K^{\mathsf{T}} P \end{aligned}$$

and calculating the corresponding derivatives, we obtain

$$\begin{aligned} \frac{\partial}{\partial K} L(K, P, \Lambda) &= 2PK\sigma^y\sigma^{y\mathsf{T}} + 2P\Lambda C^{\mathsf{T}} \\ &\quad - 2P\Lambda PK\Lambda_h^{-1}, \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial P} L(K, P, \Lambda) &= [-\sigma^x,\ K\sigma^y][-\sigma^x, K\sigma^y]^{\mathsf{T}} \\ &\quad - (A - KC)\Lambda^{\mathsf{T}} - \Lambda^{\mathsf{T}}(A - KC)^{\mathsf{T}} \\ &\quad - (RP\Lambda)^{\mathsf{T}} - (\Lambda PR)^{\mathsf{T}}, \end{aligned}$$

that implies the following optimality condition:

$$2PK\sigma^y\sigma^{y\mathsf{T}} + 2P\Lambda C^{\mathsf{T}} - 2P\Lambda PK\Lambda_h^{-1} = 0,$$

$$\begin{aligned} &[-\sigma^x,\ K\sigma^y][-\sigma^x, K\sigma^y]^{\mathsf{T}} - (A - KC)\Lambda^{\mathsf{T}} \\ &\quad - \Lambda^{\mathsf{T}}(A - KC)^{\mathsf{T}} - (RP\Lambda)^{\mathsf{T}} - (\Lambda PR)^{\mathsf{T}} = 0, \end{aligned}$$

$$P(A - KC) + (A - KC)^{\mathsf{T}} P + PRP + Q = 0,$$

$$R := \Lambda_f^{-1} + K\Lambda_h^{-1}K^{\mathsf{T}}.$$

Notice that Λ in (18) may be selected as symmetric matrix since for $S = S^{\mathsf{T}}$

$$\operatorname{tr}\{S\Lambda\} = \operatorname{tr}\{(S\Lambda)^{\mathsf{T}}\} = \operatorname{tr}\{\Lambda^{\mathsf{T}}S\} = \operatorname{tr}\{S\Lambda^{\mathsf{T}}\}$$
$$= \operatorname{tr}\{S(\Lambda^{\mathsf{T}} + \Lambda)/2\} = \operatorname{tr}\{S\tilde{\Lambda}\},$$

$$\tilde{\Lambda} := (\Lambda^{\mathsf{T}} + \Lambda)/2 = \tilde{\Lambda}^{\mathsf{T}},$$

In view of this, the last relations can be transformed to the following:

$$\bar{F}_1(K,P,\Lambda) := K(\sigma^y\sigma^{y\mathsf{T}}\Lambda_h) + \Lambda C^{\mathsf{T}}\Lambda_h - (\Lambda P)K = 0, \quad P > 0,$$

$$\bar{F}_2(K,P,\Lambda) := \sigma^x(\sigma^x)^{\mathsf{T}} + K\sigma^y(\sigma^y)^{\mathsf{T}}K^{\mathsf{T}},$$
$$-(A - KC + RP)\Lambda - \Lambda(A - KC + RP)^{\mathsf{T}} = 0, \tag{19}$$

$$\bar{F}_3(K,P) := P(A - KC) + (A - KC)^{\mathsf{T}}P + PRP + Q = 0,$$

$$R := \Lambda_f^{-1} + K\Lambda_h^{-1}K^{\mathsf{T}}.$$

Remark 2. The necessary optimality conditions (19) are also the sufficient ones. To show this fact let us represent the original problem (17) and (16) in the new variables $P, Y = PK$ that leads to the following problem:

$$J_+ := \operatorname{tr}\{S_xP + YS_yY^{\mathsf{T}}P^{-1}\} \to \min_{Y,P>0},$$

$$S_x := \sigma^x\sigma^{x\mathsf{T}} \geqslant 0, \quad S_y := \sigma^y\sigma^{y\mathsf{T}} \geqslant 0,$$

$$PA + A^{\mathsf{T}}P - YC - C^{\mathsf{T}}Y^{\mathsf{T}} + PRP + Q = 0. \tag{20}$$

Let us show that the aim function in (20) is convex in P, Y. First, notice that to prove that, it is sufficient to prove the convexity of the function

$$f(Y,P) := \operatorname{tr}\{YS_yY^{\mathsf{T}}P^{-1}\}. \tag{21}$$

Second, also notice that the matrix space $M = R^{n\times m}$ is a Hilbert one with the scalar product $\langle X,Y\rangle = \operatorname{tr}\{XY^{\mathsf{T}}\}$ where the linear operator $A : M \to M$, defined by a matrix $A \in R^{n\times n}$, acts as $A(X) = AX$. If $A \geqslant 0$, then

$$\langle AX,X\rangle = \operatorname{tr}\{AXX^{\mathsf{T}}\} = \operatorname{tr}\{X^{\mathsf{T}}AX\}$$
$$= \operatorname{tr}\{(X^{\mathsf{T}}A^{1/2})(A^{1/2}X)\}$$
$$= \operatorname{tr}\{((A^{1/2}X)^{\mathsf{T}}(A^{1/2}X))\} \geqslant 0.$$

It means that the property of positivity (semi-positivity) of operators $A(X)$ in M is induced by the corresponding property for the matrices $A \in R^{n\times n}$. Let us consider the function

$$\varphi(\alpha,\beta) := f(Y + \alpha Z, P + \beta Q)$$
$$= \operatorname{tr}\{(Y + \alpha Z)S_y(Y + \alpha Z)^{\mathsf{T}}(P + \beta Q)^{-1}\}$$

defined for any $Z \in M$ and $Q = Q^{\mathsf{T}}$. Calculating the derivative of $\varphi(\alpha,\beta)$ in the point $\varphi(0,0)$, we get

$$\varphi_\alpha = \operatorname{tr}\{(ZS_yY^{\mathsf{T}} + YS_yZ^{\mathsf{T}})P^{-1}\},$$
$$\varphi_\beta = \operatorname{tr}\{YS_yY^{\mathsf{T}}P^{-1}QP^{-1}\},$$

$$\varphi_{\alpha\alpha} = 2\operatorname{tr}\{ZS_yZ^{\mathsf{T}}P^{-1}\}$$
$$= 2\operatorname{tr}\{P^{1/2}ZS_yZ^{\mathsf{T}}P^{1/2}\} \geqslant 0,$$

$$\varphi_{\beta\beta} = \operatorname{tr}\{YS_yY^{\mathsf{T}}P^{-1}QP^{-1}QP^{-1}\}$$
$$= \operatorname{tr}\{P^{-1/2}[QP^{-1}YS_yY^{\mathsf{T}}P^{-1}Q]P^{-1/2}\} \geqslant 0,$$

$$\varphi_{\alpha\beta} = \operatorname{tr}\{(ZS_yY^{\mathsf{T}} + YS_yZ^{\mathsf{T}})P^{-1}QP^{-1}\}.$$

To show the convexity of $f(Y,P)$ (21) it is sufficient to check if

$$g(Z,Q) := \varphi_{\alpha\alpha}\varphi_{\beta\beta} - \varphi_{\alpha\beta}^2 \geqslant 0$$

for any $Z \in M$ and $Q = Q^{\mathsf{T}}$, that follows from the Cauchy–Bounyakovski inequality

$$\langle P^{-1}Z,Z\rangle\langle PU,U\rangle \geqslant \langle U,Z\rangle^2,$$

since

$$g(Z,Q) = 4\langle ZS_yZ^{\mathsf{T}}P^{-1}\rangle\langle ZS_yY^{\mathsf{T}} + YS_yZ^{\mathsf{T}}, P^{-1}(PP^{-1})QP^{-1}\rangle,$$

$$-\langle ZS_yY^{\mathsf{T}} + YS_yZ^{\mathsf{T}}, P^{-1}QP^{-1}\rangle^2$$
$$=4(\langle P^{-1}\tilde{Z},\tilde{Z}\rangle\langle PU,U\rangle - \langle U,\tilde{Z}\rangle^2) \geqslant 0,$$

$$\tilde{Z} := S_y^{1/2}Z, \quad U := S_y^{1/2}P^{-1}QP^{-1}Y.$$

Taking into account that for $R > 0$ the constraints are also convex, we obtain the convex problem for which the necessary conditions are sufficient ones as well.

Introducing the Lagrange function

$$\tilde{L}(Y,P,\Lambda) := \mathrm{tr}\{S_x P + Y S_y Y^{\mathsf{T}} P^{-1}\},$$

$$-\mathrm{tr}\{[PA + A^{\mathsf{T}}P - YC - C^{\mathsf{T}}Y^{\mathsf{T}} + PRP + Q]\Lambda\}$$

and deriving (analogously to (19) the corresponding conditions of optimality, we get the equations which in the origin variables (K,P) coincide with (19).

3.3. Numerical procedure

Consider the following numerical scheme:

1. Choose an initial K_1.
2. With this value of K_1 use $\bar{F}_3 = 0$ (Riccati equation) to obtain the value for P_1.
3. With K_1 and P_1 obtain Λ_1 using $\bar{F}_2 = 0$.
4. Solving $\bar{F}_1 = 0$ for the obtained values P_n, Λ_n calculate K_2.
5. Repeat the steps 2–4 to obtain the optimal values.
6. We stop the procedure if $\|K_{n+1} - K_n\| \leqslant \varepsilon (\varepsilon > 0)$.

So, for $n = 1, 2, \ldots$, we have

P_n is found from $\bar{F}_3(K_n, P_n) = 0,$

Λ_n is found from $\bar{F}_2(K_n, P_n, \Lambda_n) = 0,$

K_{n+1} is found from $\bar{F}_1(K_{n+1}, P_n, \Lambda_n) = 0.$ (22)

This procedure can be rewritten in the algorithmic recurrent form. To do that let us introduce two following operators:

1. $\mathrm{Syl}(A,B,Q)$: $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \mapsto \mathbb{R}^{n\times n}$. This operator defines the matrix solution $X \in \mathbb{R}^{n\times n}$ to the Sylvester equation

   $$AX + XB^{\mathsf{T}} = -Q,$$

   $$A, X, B, Q \in \mathbb{R}^{n\times n}, \quad \det A \neq 0,$$

   which has the unique solution

   $$X = \mathrm{Syl}(A,B,Q) := -\mathrm{col}^{-1}\{[I + A^{-1} \otimes B^{\mathsf{T}}]^{-1}\mathrm{col}\{A^{-1}B\}\},$$

   ($A^{-1} \otimes B^{\mathsf{T}}$ is the Kronecker matrix product) if and only if $\lambda_i + \mu_j \neq 0$ for any $i, j = 1, \ldots, n$, where λ_i are the eigenvalues of A and μ_j are the eigenvalues of B. If $A = B$ this equation is known as the Lyapunov equation.
2. $\mathrm{Ric}(A,R,Q): \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \mapsto \mathbb{R}^{n\times n}$. This operator defines the unique positive matrix solution $P \in \mathbb{R}^{n\times n}$ to the algebraic matrix Riccati equation

   $$PA + A^{\mathsf{T}}P + Q - PRP = 0,$$

   which provides the stability to the matrix $A_{\mathrm{closed}} := [A - RP]$ if the pair (A,B) is completely controllable and one of two conditions is fulfilled: either $Q > 0$, or Q may be represented as $Q = C^{\mathsf{T}}C$ such that the pair $(A^{\mathsf{T}}, C^{\mathsf{T}})$ is completely observable.

In view of this definition the numerical procedure (22) may be represented as

$$P_n = \mathrm{Ric}([A - K_n C], [-R], Q),$$

$$\Lambda_n = \mathrm{Syl}([-A + K_n C - RP_n], [-A + K_n C - RP_n], [\sigma^x(\sigma^x)^{\mathsf{T}} + K_n \sigma^y (\sigma^y)^{\mathsf{T}} K_n^{\mathsf{T}}]),$$

$$K_{n+1} = \mathrm{Syl}([-\Lambda_n P_n], [\sigma^y \sigma^{y\mathsf{T}} \Lambda_h], [\Lambda_n C^{\mathsf{T}} \Lambda_h]).$$

3.4. Monotonicity property and convergence

By (22), it follows:

$$\begin{aligned} J_+(K_n, P) &\overset{(3)}{=} J_+(K_n, P_n) \overset{(2)}{=} L(K_n, P_n, \Lambda_n) \\ &\overset{(1)}{\geqslant} L(K_{n+1}, P_n, \Lambda_n) \\ &\overset{(3)}{\geqslant} L(K_{n+1}, P_{n+1}, \Lambda_n) \\ &= J_+(K_{n+1}, P_{n+1}). \end{aligned} \quad (23)$$

Therefore, since $J_+(K,P) \geqslant 0$, we obtain

$$J_+(K_n, P_n) \underset{n\to\infty}{\longrightarrow} J_+^* \geqslant 0.$$

Using the strict convexity property of (17) in K and P, the monotonicity implies the convergence $K_{n n\to\infty} \longrightarrow K^*$, $P_{n n\to\infty} \longrightarrow P^*$. The uniqueness of the limit points K^* and P^* also follows from the structure of (17).

3.5. The comparison to the Kalman filter: the scalar case

For a stationary linear system, having the given (A, B, C, D)-realization, the Kalman filter has the same structure as (3) with $f(t, \hat{x}_t) \equiv 0$, $h(t, \hat{x}_t) \equiv 0$ for all $t \geqslant 0$, $\hat{x}_t \in R^n$. Assuming that $\sigma^y(\sigma^y)^\mathsf{T} > 0$, this filter provides the minimal estimation error covariance matrix $P := \lim_{t\to\infty} E\{\Delta_t \Delta_t^\mathsf{T}\}$ satisfying (see, for example, [11]) the following matrix Riccati equation:

$$PA^\mathsf{T} + AP + Q - PRP = 0,$$

$$Q = \sigma^x(\sigma^x)^\mathsf{T}, \quad R = K\sigma^y(\sigma^y)^\mathsf{T}K^\mathsf{T}, \tag{24}$$

where the filter gain matrix K is equal to $K = PC^\mathsf{T}[\sigma^y(\sigma^y)^\mathsf{T}]^{-1}$. Considering, for the simplicity, the scalar case ($n = 1$), one may find the solution of (24)

$$P = \left(\frac{\sigma^y}{C}\right)^2 \left(A + \sqrt{A^2 + \left(C\frac{\sigma^x}{\sigma^y}\right)^2}\right). \tag{25}$$

On the other hand, problem (17) for $n = 1$, which has the form ($L_f = L_h = 0$)

$$J = ((\sigma^x)^2 + (\sigma^y)^2K^2)P \to \min_K,$$

$$2P(A - KC) + Q = 0,$$

$$Q = Q_0 := 1, \quad P > 0, \ (A - KC) \text{ is stable}$$

can be represented as

$$J = J(K) = \frac{((\sigma^x)^2 + (\sigma^y)^2K^2)Q}{2|A - KC|} \to \min_{K<A/C}. \tag{26}$$

The solution of this problem is given by

$$K = K^* = C^{-1}\left(A + \sqrt{A^2 + \left(C\frac{\sigma^x}{\sigma^y}\right)^2}\right).$$

The substitution of K^*into (26) implies

$$J(K^*) = \left(\frac{\sigma^y}{C}\right)^2 \left(A + \sqrt{A^2 + \left(C\frac{\sigma^x}{\sigma^y}\right)^2}\right)$$

that exactly coincides with (25). This simple example shows that the obtained minimal upper bound is "*tight*", that is, there exists a class of systems (in this case, the class of unidimensional linear stationary systems) for which the obtained upper bound is reachable.

3.6. Example

Let the nonlinear stochastic system is given by

$$\mathrm{d}x_{1,t} = [-5x_{1,t} + 3x_{2,t}^{1/2} + 0.05\sin(\omega t)]\,\mathrm{d}t + 0.01\,\mathrm{d}\bar{W}_t^x,$$

$$\mathrm{d}x_{2,t} = [2x_{1,t}^{1/2} - 4x_{2,t} + 0.08\cos(\omega t)]\,\mathrm{d}t + 0.01\,\mathrm{d}\bar{W}_t^x,$$

$$\mathrm{d}y_t = C\,\mathrm{d}x_t + 0.01\,\mathrm{d}\bar{W}_t^y$$

with

$$C = [1 \ 1], \quad \omega = 1, \quad x_{1,0} = 0.1; \quad x_{2,0} = 0.3.$$

Using the methodology analyzed in this paper we obtain the optimal matrix gain K^* as

$$K^* = \begin{bmatrix} 0.7188 \\ 1.1259 \end{bmatrix}.$$

To demonstrate the effectiveness of the methodology proposed we select 2 gain matrices for the comparison

$$K_p = \begin{bmatrix} 4.7188 \\ 5.1259 \end{bmatrix}, \quad K_l = \begin{bmatrix} 0.1188 \\ 0.2125 \end{bmatrix}.$$

The figures below show that both matrices K_p and K_l provides more than 3 times less performance to the estimation process (Figs. 1 and 2).

The performance index (see Fig. 3) after 10 s for each value of K are as follows:

$$I_{t=10} = (t-1)^{-1}\int_{\tau=1}^{t} \|\Delta_\tau\|^2\,\mathrm{d}\tau = 2.209 \times 10^{-7} \quad \text{for } K^*,$$

$$I_{t=10} = 8.21 \times 10^{-6} \quad \text{for } K_p,$$

$$I_{t=10} = 7.168 \times 10^{-6} \quad \text{for } K_1.$$

Fig. 4 shows that the given optimization problem has a unique minimal point that corresponds to the optimal gain.

4. Conclusion

The suggested approach may have a wide spectrum of potential applications dealing with the state estimation (filtering) of nonlinear systems for which an optimal finite-dimensional filter cannot be designed

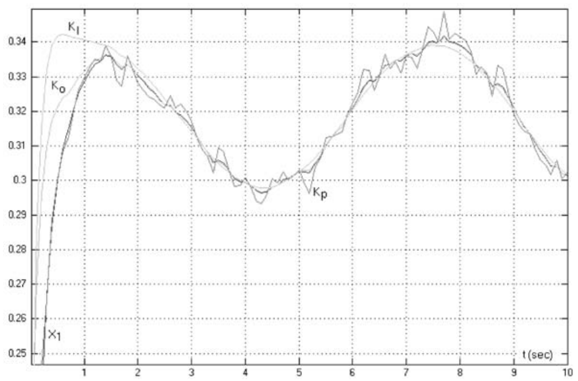

Fig. 1. X1.

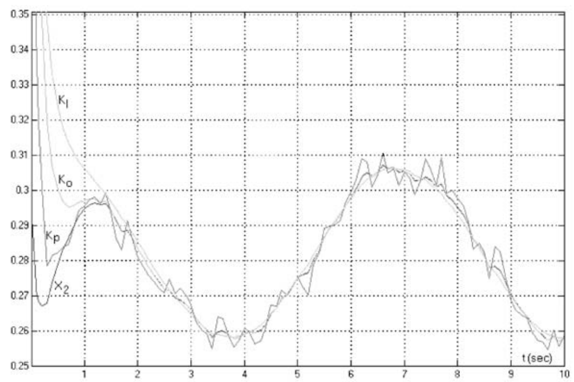

Fig. 2. X2.

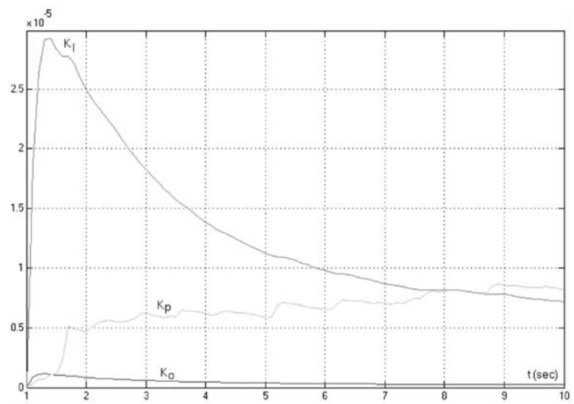

Fig. 3. Performance indexes.

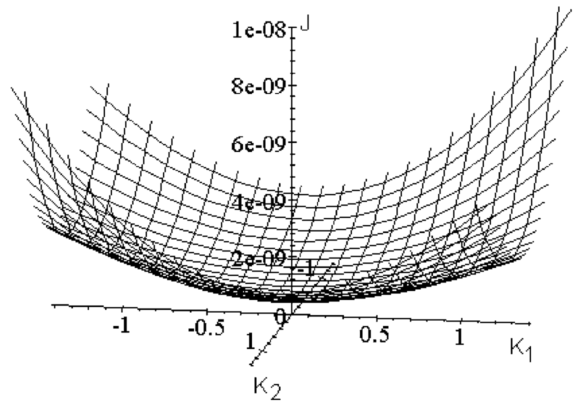

Fig. 4. Convex surface.

in principle. A simple numerical algorithm, suggested here for the optimization of the observer gain matrix, can be realized before the main state estimation process using the standard Matlab procedures based only on a priori available information. Further developments may be related to a design of a robust observer for nonlinear systems which nonlinearity is unknown exactly but belongs to a given class.

Acknowledgements

The authors express their acknowledgments to Prof. B.T. Polyak for helpful discussions concerning to the optimality conditions and the corresponding numerical procedure.

References

[1] D.S. Bayard, Fast observers for spacecraft pointing control, Proceedings of the 37th IEEE Conference on Decision and Control, Vol. 4, NJ, USA, 1998, pp. 4702–4707.
[2] E. Bullinger, F. Allgower, An adaptive high-gain observer for nonlinear systems, Proceedings of the IEEE Conference on Decision and Control, CA, USA, 1997, pp. 4348–4353.
[3] A.I. Bhatti, J.A. Twiddle, S.K. Spurgeon, N.B. Jones, Engine coolant system fault diagnostics with sliding mode observers and fuzzy analyzer, Proceedings of the 18th IASTED International Conference Modelling, Identification and Control, CA, USA, 1999, pp. 273–276.
[4] X. Chen, T. Fukuda, K.D. Young, A new nonlinear robust disturbance observer, Systems Control Lett. 41 (3) (2000) 189–199.

[5] E. Chiricozzi, A. Germani, C. Manes, F. Parasiliti, Deterministic and stochastic observers for electrical drives control, Proceedings of IMACS/IFAC Second Symposium on: Mathematical and Intelligent Models in System Simulation, Vol. 11, Belgium, 1993, pp. 178–194.
[6] H. Deng, M. Krstic, Output-Feedback stochastic nonlinear stabilization, IEEE Trans. Automat. Control 44 (2) (1999) 328–333.
[7] T.C. Gard, Introduction to Stochastic Differential Equations, Marcel Dekker, New York, 1988.
[8] Kwon Hyun Wook, Kim Soo Pyung, Park Poo Gyeon, A receding horizon Kalman FIR filter for linear continuous-time systems, IEEE Trans. Automat. Control 44 (11) (1999) 2115–2120.
[9] Kim Hwan Dong, Kang Yoon Ji, Lee Il Kyo, Robust nonlinear observer for forward kinematics solution of a Stewart platform: an experimental verification, Robotica 18 (6) (2000) 601–610.
[10] A. Dabroom, H. Khalil, Numerical differentiation using high-gain observers, Proceedings of the IEEE Conference on Decision and Control (CDC97), CA, USA, 1997, pp. 4790–4794.
[11] T. Glad, L. Ljung, Control Theory Multivariable and Nonlinear Methods, Taylor & Francis, New York, 2000.
[12] N. Madjarov, L. Mihaylova, Sensitivity analysis of linear optimal stochastic observers, Conference Proceedings 1993, International Conference on Systems, Man and Cybernetics. Systems Engineering in the Service of Humans, Vol. 2, IEEE, NY, USA, 1993, pp. 482–487.
[13] W. NaNacara, E. Yaz, Linear and nonlinear estimation with uncertain observations, Proceedings of the 1994 American Control Conference, Vol. 2, NY, USA, 1994, pp. 1429–1433.
[14] S. Nicosia, P. Tomei, A. Tornambe, A nonlinear observer for elastic robots. IEEE J. Robot. Automat. 4 (1988) 45–52.
[15] A. Osorio-Cordero, A.S. Poznyak, M. Taksar, Robust stochastic filtering for linear continuous uncertain systems with time-varying parameters, Proceedings of the European Control Conference, Brussels, Belgium, July 1997; WE-M F2, CD-ROM.
[16] A.S. Poznyak, R. Martinez-Guerra, A. Osorio-Cordero, Robust high-gain observer for nonlinear closed-loop stochastic systems, Math. Problems Eng. 6 (2000) 31–60.
[17] A.S. Poznyak, A. Osorio-Cordero, Suboptimal robust filtering of states of finite dimensional linear systems with time-varying parameters under nonrandom disturbances, Proceedings of the IEEE Conference on Decision and Control, Lake Buena Vista, FL, 1994, pp. 3931–3936.
[18] G.J. Preston, D.N. Shields, S. Daley, Application of a robust nonlinear fault detection observer to a hydraulic system, International Conference on Control '96, Vol. 2, UK, 1996, pp. 1484–1489.
[19] K. Reif, S. Gunther, E. Yaz, R. Unbehauen, Stochastic stability of the discrete-time extended Kalman filter, IEEE Trans. Automat. Control 44 (4) (1999) 714–728.
[20] K. Reif, S. Gunther, E. Yaz, R. Unbehauen, Stochastic stability of the continuous-time extended Kalman filter, IEE Proc. Control Theory Appl. 147 (1) (2000) 45–52.
[21] K. Reif, F. Sonnemann, E. Yaz, R. Unbehauen, An observer for nonlinear systems based on H^{∞}-filtering techniques, Proceedings of the 1997 American Control Conference, Vol. 4, IL, USA, 1997, pp. 2379–2380.
[22] Zhu Ruijun, Chai Tianyou, Shao Cheng, Robust nonlinear adaptive observer design using dynamic recurrent neural networks, Proceedings of the 1997 American Control Conference, Vol. 2, IL, USA, 1997, pp. 1096–1100.
[23] R. Seliger, P.M. Frank, Robust nonlinear observer-based fault detection for an overhead crane, Automatic Control World Congress 1993, Vol. 5, UK, 1994, pp. 429–432.
[24] R. Seliger, B. Koppen-Seliger, Robust nonlinear observer-based fault detection for a U-tube steam generator, Proceedings of the 1995 American Control Conference, Vol. 2, IL, USA, 1995, pp. 1134–1135.
[25] D.N. Shields, Observer design for nonlinear descriptor systems for control and fault detection, Systems Sci. 22 (1) (1996) 5–20.
[26] A. Tornambe, Use of asymptotic observers having-high-gains in the state and parameter estimation, Proceedings of the IEEE Conference on Decision and Control, Tampa, FL, USA, 1989, pp. 1791–1794.
[27] E. Yaz, A. Azemi, Extensions of deterministic and stochastic variable structure observers with applications to disturbance minimization, Proceedings of the 1992 American Control Conference, Vol. 1, IL, USA, 1992, pp. 881–885.
[28] E. Yaz, A. Azemi, A numerical procedure for discrete-time nonlinear stochastic observer design, Proceedings of the 32nd IEEE Conference on Decision and Control, Vol. 2, NY, USA, 1993, pp. 1237–1238.
[29] E. Yaz, A. Azemi, Observer design for discrete and continuous non-linear stochastic systems, Int. J. Systems Sci. 24 (12) (1993) 2289–2302.
[30] E. Yaz, A. Azemi, Robust/adaptive observers for systems having uncertain functions with unknown bounds, Proceedings of the 1994 American Control Conference, Vol. 1, NY, USA, 1994, pp. 73–74.
[31] E. Yaz, Y. Gao, K. Olejniczak, Comparison of adaptive filter performance in estimating the noise statistics for harmonic models, Proceedings of the 38th IEEE Conference on Decision and Control, Vol. 5, NJ, USA, 1999, pp. 5070–5075.
[32] E. Yaz, A. Ray, Linear unbiased state estimation for random models with sensor delay, Proceedings of the 35th IEEE Conference on Decision and Control, Vol. 1, NY, USA, 1996, pp. 47–52.
[33] M. Zakai, On the optimal filtering of diffusion processes, Z. Wahrschein Verw. Geb. 11 (1969) 230–243.