15th Triennial World Congress, Barcelona, Spain

FIXED-ORDER DISSIPATIVE ESTIMATOR DESIGN FOR UNCERTAIN STOCHASTIC SYSTEMS

Edwin Engin Yaz and Yvonne Ilke Yaz*

Department of EECE, Marquette University, Milwaukee, WI 53201 USA
**Department of Mathematics, Carthage College, Kenosha, WI 53140 USA*

Abstract: A general class of discrete-time uncertain nonlinear stochastic systems with quadratic sum constraints is considered. A linear fixed order state estimator for state estimation is presented for various estimation error performance criteria in a unified framework. The observer is of order equal to the difference between the state and output vector dimensions. The performance criteria considered in this paper include guaranteed-cost suboptimal versions of estimation objectives like H_2, H_∞, stochastic passivity, etc. The design of fixed-order linear state estimators that satisfy these criteria are given using a common matrix inequality formulation. *Copyright © 2002 IFAC*

Keywords: Estimation theory, reduced-order, design, discrete-time, uncertain dynamics.

1. INTRODUCTION

The problem of fixed (reduced)-order linear state estimator design is formulated using linear matrix inequalities (LMI) for a general class of uncertain discrete-time nonlinear stochastic systems. Various criteria are used in designing discrete-time estimators leading to a common LMI framework. The reason for choosing this approach is the possible utilization of efficient numerical schemes for solving LMI (Boyd, et al 1994). In the model used, the system and measurement vectors are assumed to be driven by possibly state-dependent deterministic or stochastic disturbances defined by quadratic sum constraints. These disturbances may be generated e.g., by white noises for which covariance upper bounds are determined by unknown nonlinear functions of the state vector. Such nonlinear models with pointwise quadratic constraints are introduced in (Jacobson, 1974), and their system theoretic properties are investigated using an LMI approach in (Yaz and Yaz, 1999a). Mean-square stabilizability by state feedback and mean-square detectability concepts for this class of systems are developed and based on these, the infinite horizon linear quadratic regulator is presented in (Yaz, 1989a). Covariance assignment formulation is used to characterize static output feedback stabilizability in the mean-square sense and parameterize all stabilizing gains in (Yaz and Yaz, 1999b). Full-order minimum variance state estimator design is introduced in (Yaz, 1988). In the present work, various estimation problems including guaranteed-cost suboptimal version of H_2, H_∞, stochastic passivity, etc. are tackled with a common framework. In that sense, the present work can be viewed as an extension of minimum-variance results of (Yaz, 1988) to the case of reduced-order estimation for generalized performance criteria with quadratic sum rather than pointwise quadratic constraints in time.

The following notation is used. $\mathbf{Z}^+$ is the set of nonnegative integers. For an n-dimensional vector of real elements $x \in \boldsymbol{R}^n$, $||x||$ denotes the 2-norm, $(x^T x)^{1/2}$. For an n x n symmetric matrix A, $A > 0$ ($A < 0$) and $A \geq 0$ ($A \leq 0$) denote positive (negative) definite and positive (negative) semidefinite A. The Schur complement formula which states the equivalence of $\begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \geq 0$ and $C > 0$, $A \geq BC^{-1}B^T$ and $A > 0$, $C \geq B^T A^{-1} B$ is used in this work. $\lambda_{\max}(A), \lambda_{\min}(A)$ (for symmetric A)

respectively mean the maximum and minimum eigenvalue of A. $A^{\dagger}$ denotes the unique Moore-Penrose pseudo-inverse of the A matrix. The form of Rayleigh's inequality $\lambda_{\min}(A)I \le A \le \lambda_{\max}(A)I$ for symmetric A is used. $E\{x\}$ and $E_y\{x\}$ respectively denote the expectation of x and its expectation conditional on y. The interlacing of expectations $E\{E_y\{x\}\} = E\{x\}$ is also used.

2. SYSTEM AND MEASUREMENT MODELS

We assume that the signal is generated by the following system and the measurement equations

$$x_{k+1} = Ax_k + Bp_k \qquad (1)$$

$$y_k = C_1 x_k + Dp_k \qquad (2)$$

where the state $x_k \in \mathbf{R}^n$ and the measurement $y_k \in \mathbf{R}^p$. C_1 is of full rank without loss of generality. The initial state x_0 is assumed to have the known mean $E\{x_0\} = \bar{x}_0$, covariance $E\{x_0 x_0^T\} = X_0$, and to be uncorrelated with other noise sources. The unknown disturbance p_k, which may be a nonlinear stochastic function of x_k in general, is defined by stochastic dissipation inequality as follows:

$$\delta_f \sum_{k=0}^{N} E\{\|p_k\|^2\} + \varepsilon_f \sum_{k=0}^{N} E\{\|q_k\|^2\} \le \beta_f \sum_{k=0}^{N} E\{p_k^T q_k\} \qquad (3)$$

for any $N \in \mathbf{Z}^+$, where

$$q_k = C_q x_k + D_q p_k \qquad (4)$$

This description is motivated by the breadth of realistic feedback models it encompasses, for example:

(a) Linear systems with state-multiplicative noises $C_q x_k v_k$ where v_k is a zero mean uncorrelated noise sequence with an uncertain second moment having a known bound.

(b) Nonlinear systems with random sequences whose powers depend on sector-bound nonlinear function of the state $\|\psi(x_k)\| \le \alpha \|x_k\|$ where the form of the nonlinearity may not be known but the bound α is known.

(c) Nonlinear systems with a random sequence whose power depends on a bounded nonlinear function of the state $\phi(x_k) C_q x_k v_k$ where $\phi(x_k)$ satisfies $\|\phi(x)\| \le \alpha$ for all $x \in \mathbf{R}^n$.

(d) Linear systems with state multiplicative disturbances $C_q x_k w_k$ where $w_k \in l_2$ (i.e. $\sum_{k=0}^{\infty} \|w_k\|^2 < \infty$) or $w_k \in l_\infty$ (i.e. $sup_k \|w_k\|) < \infty$).

(e) Linear systems having nonlinear stochastic systems in the feedback loop, where the feedback system has a stochastic dissipativity property. For example, if δ_f, ε_f, and β_f are all positive real numbers, the feedback system possesses stochastic version of very strict passivity. There are several special cases where for example taking $\delta_f = \varepsilon_f = 0$ would result in simple (mean-square stochastic) passivity:

$$\sum_{k=0}^{N-1} E\{p_k^T q_k\} \ge 0 .$$

Another special case is e.g. $\delta_f = 1$, $\varepsilon_f = -1$, and $\beta_f = 0$ to yield an H_∞ norm less than or equal to 1:

$$\sum_{k=0}^{N} E\{\|p_k\|^2\} \le \sum_{k=0}^{N} E\{\|q_k\|^2\}.$$

Other models that fit this description can be found in (Jacobson, 1974), (Yaz and Yaz, 1999a), and (Yaz, 1989b). One can see that some of the most important deterministic and stochastic nonlinear models in use today can be treated in this general framework.

3. ESTIMATOR DESIGN

We define the linear state transformation

$$z_k = C_2 x_k \qquad (5)$$

where $z_k \in \mathbf{R}^{n-p}$ and C_2 is a full rank matrix such that $\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$ is nonsingular with

$$\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}^{-1} = \begin{pmatrix} \Omega_1 & \Omega_2 \end{pmatrix} \qquad (6)$$

It is always possible to find a C_2 for a given full rank C_1 such that (6) is true. In fact, based on the full rank properties of C_1 and C_2, we can find explicit expressions for Ω_1 and Ω_2 using the results in (Ben-Israel and Greville, 1974) as follows:

$$\begin{pmatrix} \Omega_1 & \Omega_2 \end{pmatrix} = (C_1^{\dagger} C_1 + C_2^{\dagger} C_2)^{-1} (C_1^{\dagger} \quad C_2^{\dagger})$$

$=(C_1^T(C_1C_1^T)^{-1}C_1+C_2^T(C_2C_2^T)^{-1}C_2)^{-1}$

$(C_1^T(C_1C_1^T)^{-1} \quad C_2^T(C_2C_2^T)^{-1})$ (7)

The vector z_k defined by (5) is one of the intermediate variables in this estimator design process. Note that it follows from (2), (5), and (6) that if we knew what z_k and p_k were, we would be able to find x_k as follows:

$$x_k = \Omega_1(y_k - Dp_k) + \Omega_2 z_k \qquad (8)$$

However, p_k is an unknown disturbance and z_k can not be directly measured, so the estimate of x_k denoted by $\hat{x}_k$ can be formed by

$$\hat{x}_k = \Omega_1 y_k + \Omega_2 \hat{z}_k \qquad (9)$$

Therefore, to be able to estimate x_k, an estimate of z_k (which is (n-p)-dimensional) is sufficient since the measurement y_k is available. This implies that we need to look at the dynamics of y_k and z_k:

$$\begin{pmatrix} y_{k+1} \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}(Ax_k + Bp_k) + \begin{pmatrix} D \\ 0 \end{pmatrix} p_{k+1}$$
$$= \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}[A(\Omega_1 \quad \Omega_2)\begin{pmatrix} y_k - Dp_k \\ z_k \end{pmatrix} + Bp_k]$$
$$+ \begin{pmatrix} D \\ 0 \end{pmatrix} p_{k+1} \qquad (10)$$

where we substituted from (1) and (8). We define

$$v_k \doteq z_k - Ky_k \qquad (11)$$

which is the variable that is directly used in forming the estimator. K is called the observer gain. Note that v_k has the same dimension as z_k and estimating v_k is equivalent to estimating z_k since y_k is the known measurement vector, or

$$\hat{v}_k = \hat{z}_k - Ky_k \qquad (12)$$

The dynamics of v_k can be found from the dynamics of z_k and y_k as follows:

$$v_{k+1} = (C_2 - KC_1)A(\Omega_1 \quad \Omega_2)\begin{pmatrix} y_k - Dp_k \\ z_k \end{pmatrix}$$
$$+ (C_2 - KC_1)Bp_k - KDp_{k+1}$$
(13)

The estimate of v_k can be constructed (based on the same assumptions as made before) as

$$\hat{v}_{k+1} = (C_2 - KC_1)A(\Omega_1 \quad \Omega_2)\begin{pmatrix} y_k \\ \hat{z}_k \end{pmatrix}$$

$$= (C_2 - KC_1)A(\Omega_1 \quad \Omega_2)\begin{pmatrix} y_k \\ \hat{v}_k + Ky_k \end{pmatrix}$$
(14)

by using (12) and (13). It can be inferred from (5) and (11) that this iteration is to be initialized with $\hat{v}_0 = C_2\bar{x}_0 - Ky_0$. The resulting estimation error $e_k \doteq v_k - \hat{v}_k$ has the dynamics

$$e_{k+1} = (C_2 - KC_1)A\Omega_2 e_k + (C_2 - KC_1)$$
$$(B - A\Omega_1 D)p_k - KDp_{k+1}$$
$$\doteq A_0 e_k + B_0 p_k - KDp_{k+1} \qquad (15)$$

The following result summarizes the main contribution of this paper:

Theorem 1. Consider the stochastic nonlinear model (1) – (4), the performance output

$$r_k = C_r e_k + D_r p_k \qquad (16)$$

and the reduced-order linear state estimator given by (9), (12) and (14). If the following LMI hold

$$\Gamma = \begin{pmatrix} g_{11} & g_{12} & g_{13} & g_{14} & g_{15} \\ * & g_{22} & g_{23} & g_{24} & g_{25} \\ * & * & g_{33} & g_{34} & g_{35} \\ * & * & * & g_{44} & g_{45} \\ * & * & * & * & g_{55} \end{pmatrix} \geq 0$$
(17)

where

$$g_{11} = P_1 - A^T P_1 A + \tau_1 \varepsilon_f C_q^T C_q$$
$$+ \tau_2 \varepsilon_f A^T C_q^T C_q A, g_{12} = 0$$
$$g_{13} = -A^T P_1 B + \tau_1 C_q^T (\varepsilon_f D_q - 0.5 \beta_f I)$$
$$+ \tau_2 \varepsilon_f A^T C_q^T C_q B$$
$$g_{14} = A^T C_q^T (\tau_2 \varepsilon_f D_q + 0.5 \beta_f I)$$
$$g_{15} = 0$$
$$g_{22} = P_2 - \delta C_r^T C_r$$
$$g_{23} = -C_r^T (\delta D_r - \beta I)$$
$$g_{24} = 0$$
$$g_{25} = \Omega_2^T A^T (C_2^T P_2 - C_1^T Y^T)$$
$$g_{33} = \tau_1 (\delta_f I + \varepsilon_f D_q^T D_q)$$
$$- 0.5 \tau_1 \beta_f (D_q + D_q^T) + \tau_2 \varepsilon_f B^T C_q^T C_q B$$
$$- \delta D_r^T D_r + 0.5 \beta (D_r + D_r^T) - \varepsilon I$$
$$g_{34} = 0$$
$$g_{35} = (B - A \Omega_1 D)^T (C_2^T P_2 - C_1^T Y^T)$$
$$g_{44} = \tau_2 (\delta_f I + \varepsilon_f D_q^T D_q) + 0.5 \beta_f (D_q + D_q^T)$$
$$g_{45} = -D^T Y^T$$
$$g_{55} = P_2$$

for some Y, $P_1 > 0$, $P_2 > 0$, and τ_1, $\tau_2 \geq 0$, then the state and estimation error dynamics satisfy:

$$E\{x_N^T P_1 x_N\} + E\{e_N^T P_2 e_N\} \leq E\{x_0^T P_1 x_0\}$$
$$+ E\{e_0^T P_2 e_0\}$$
$$+ \tau_1 \sum_{k=0}^{N-1} E\{ \delta_f ||p_k||^2 + \varepsilon_f ||q_f||^2$$
$$-\beta_f\ p_k^T q_k \} + \tau_2 \sum_{k=0}^{N-1} E\{ \delta_f ||p_{k+1}||^2 + \varepsilon_f ||q_{k+1}||^2$$
$$- \beta_f\ p_{k+1}^T q_{k+1} \}$$
$$- \sum_{k=0}^{N-1} E\{\delta ||r_k||^2 + \varepsilon\ ||p_k||^2 - \beta r_k^T p_k \}$$
(19)

for all $N \in \mathbf{Z}^+$, where the estimator gain is found by $K = P_2^{-1} Y$.

Sketch of Proof. Letting $V_k = x_k^T P_1 x_k + e_k^T P_2 e_k > 0$, consider

$$E_{x_k, e_k} \{V_{k+1} - V_k - \tau_1 (\delta_f ||p_k||^2 + \varepsilon_f ||q_k||^2$$
$$- \beta_f\, p_k^T q_k) - \tau_2 (\delta_f ||p_{k+1}||^2$$
$$+ \varepsilon_f ||q_{k+1}||^2 - \beta_f\, p_{k+1}^T q_{k+1}) + \delta ||r_k||^2$$
$$+ \varepsilon ||p_k||^2 - \beta r_k^T p_k \} \doteq \Delta V_k \qquad (20)$$

Upon substitution from (1), (15) and (16), taking expectations of both sides, using the interlacing property of expectations and rearranging, leads to

$$E\{\Delta V_k\} \leq -E\{\begin{pmatrix} x_k^T & e_k^T & p_k^T & p_{k+1}^T \end{pmatrix} \mathrm{H} \begin{pmatrix} x_k \\ e_k \\ p_k \\ p_{k+1} \end{pmatrix}\} \qquad (21)$$

where

$$\mathrm{H} = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ * & h_{22} & h_{23} & h_{24} \\ * & * & h_{33} & h_{34} \\ * & * & * & h_{44} \end{pmatrix}$$

with

$$h_{11} = P_1 - A^T P_1 A + \tau_1 \varepsilon_f C_q^T C_q$$
$$+ \tau_2 \varepsilon_f A^T C_q^T C_q A, h_{12} = 0$$
$$h_{13} = -A^T P_1 B + \tau_1 C_q^T (\varepsilon_f D_q - 0.5 \beta_f I)$$
$$+ \tau_2 \varepsilon_f A^T C_q^T C_q B$$
$$h_{14} = A^T C_q^T (\tau_2 \varepsilon_f D_q + 0.5 \beta_f I)$$
$$h_{22} = P_2 - A_0^T P_2 A_0 - \delta C_r^T C_r$$
$$h_{23} = -A_0^T P_2 B_0 - C_r^T (\delta D_r - \beta I)$$
$$h_{24} = A_0^T P_2 K D$$

$$h_{33} = \tau_1(\delta_f I + \varepsilon_f D_q^T D_q) - B_0^T P_2 B_0$$
$$-0.5\tau_1\beta_f(D_q + D_q^T) + \tau_2\varepsilon_f B^T C_q^T C_q B$$
$$-\delta D_r^T D_r + 0.5\beta(D_r + D_r^T) - \varepsilon I$$
$$h_{34} = B_0^T P_2 KD + \tau_2\varepsilon_f B^T C_q^T C_q D_q$$
$$+0.5\beta_f B^T C_q^T$$
$$h_{44} = \tau_2(\delta_f I + \varepsilon_f D_q^T D_q) - D^T K^T P_2 KD$$
$$+0.5\beta_f(D_q + D_q^T)$$

Since LMI (17) hold, then by Schur's complement result, H ≥ 0. Therefore,

$$-E\{\begin{pmatrix} x_k^T & e_k^T & p_k^T & p_{k+1}^T \end{pmatrix} \mathrm{H} \begin{pmatrix} x_k \\ e_k \\ p_k \\ p_{k+1} \end{pmatrix}\} \leq 0$$

(22)

Summing over k, leads to (19) which completes the proof.

The above proof of the main result provides a procedure for designing state estimators with improved response. By applying Rayleigh's inequality in (19), one can see that

$$\lambda_{\min}(P_2)E\{||e_N||^2\} \leq \lambda_{\min}(P_1)\mathrm{E}\{||x_N||^2\} + \lambda_{\min}(P_2)\mathrm{E}\{||e_N||^2\} \leq \lambda_{\max}(P_1)\mathrm{E}\{||x_0||^2\} + \lambda_{\max}(P_2)\mathrm{E}\{||e_0||^2\} + \tau_1\sum_{k=0}^{N-1} E\{\delta_f ||p_k||^2 + \varepsilon_f ||q_k||^2 - \beta_f\ p_k^T q_k\} + \tau_2\sum_{k=0}^{N-1} E\{\delta_f ||p_{k+1}||^2 + \varepsilon_f ||q_{k+1}||^2 - \beta_f p_{k+1}{}^T q_{k+1}\} - \sum_{k=0}^{N-1} E\{\delta ||r_k||^2 + \varepsilon ||p_k||^2 - \beta r_k^T p_k\} \qquad (23)$$

So, maximizing $\lambda_{\min}(P_2)$, τ_1 and τ_2, minimizing $\lambda_{\max}(P_1)$ and $\lambda_{\max}(P_2)$ subject to LMI (17) will give a smaller mean-square error. This is a generalized eigenvalue problem solvable by available LMI software as explained in (Boyd, et al, 1994).

4. APPLICATION TO VARIOUS ESTIMATION PROBLEMS

Theorem 1 given above allows one to design different estimators for a variety of performance criteria for this class of systems. For example, taking $\delta = 0$, $\beta = 0$ and $\varepsilon < 0$ yields

$$E\{e_N^T P_2 e_N\} \leq E\{x_0^T P_1 x_0\} + E\{e_0^T P_2 e_0\} - \varepsilon\sum_{k=0}^{N-1} E\{\|p_k\|^2\}$$

This means that by employing the optimization procedure described above, we can obtain a tight bound on the mean-square estimation error.

By taking $\delta > 0$, $\beta = 0$, $\varepsilon = 0$, $B = 0$, $D = 0$, $D_q = 0$ and $D_r = 0$, we obtain

$$\delta\sum_{k=0}^{N-1} E\{||r_k||^2\} \leq E\{x_0^T P_1 x_0\} + E\{e_0^T P_2 e_0\}$$

which yields a bound on the energy of the performance output in terms of the initial estimation error (suboptimal H_2 result).

If we set $\delta = 1$, $\beta = 0$, and $\varepsilon < 0$, for $x_0 = 0$ and $e_0 = 0$, this produces the result

$$\sum_{k=0}^{N-1} E\{||r_k||^2\} \leq -\varepsilon\sum_{k=0}^{N-1} E\{\ ||p_k||^2\}$$

which is a bound on the stochastic (mean-square) l_2 to l_2 gain of the estimator (suboptimal H_∞ result.

Several dissipative estimator designs are also possible using this formulation. For example, taking $x_0 = 0$, $e_0 = 0$, $\delta = 0$, $\beta = 1$ and $\varepsilon > 0$ will yield the stochastic (mean-square) version of the input strict passivity result:

$$\sum_{k=0}^{N-1} E\{\ r_k^T p_k\} \geq \varepsilon\sum_{k=0}^{N-1} E\{||p_k||^2\}$$

Other similar dissipativity results are also possible. For example, setting $\delta = 0$, $\beta = 1$, and $\varepsilon = 0$ will give stochastic passivity:

$$\sum_{k=0}^{N-1} E\{r_k^T p_k\} \geq 0$$

Setting $\delta > 0$, $\beta = 1$, and $\varepsilon = 0$ will yield output strict passivity:

$$\sum_{k=0}^{N-1} E\{r_k^T p_k\} \geq \delta \sum_{k=0}^{N-1} E\{||r_k||^2\}$$

Also, setting $\delta > 0$, $\beta = 1$, and $\varepsilon > 0$ will give strict passivity both in terms of the input and the output (very strict passivity in the mean-square sense):

$$\sum_{k=0}^{N-1} E\{r_k^T p_k\} \geq \delta \sum_{k=0}^{N-1} E\{||r_k||^2\} + \varepsilon \sum_{k=0}^{N-1} E\{||p_k||^2\}$$

So, one can see that this LMI formulation allows one to consider a variety of performance criteria in a common framework.

5. CONCLUSIONS

We have considered reduced-order linear state estimator designs for a class of discrete-time uncertain stochastic systems with quadratic sum constraints and general dissipative performance criteria. We have shown that a common framework using linear matrix inequality formulations can be provided to solve diverse estimator design problems. The future work will involve the robustness study of these estimators in the presence of uncertain parameters and the performance study to inquire whether degradation in performance results due to the use of this reduced-order formulation.

REFERENCES

Ben-Israel and T. N. E. Greville (1974) Generalized Inverses: Theory and Applications, Wiley: New York.

S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan (1994) Linear Matrix Inequalities in System and Control Theory, SIAM: Philadelphia, PA.

D. H. Jacobson (1974) "A general result in stochastic optimal control of nonlinear discrete-time systems with quadratic performance criteria," J. Math. Anal. Applications, **47**, 153.

E. Yaz (1988) "Linear state estimators for nonlinear stochastic systems with noisy nonlinear observations," Int. J. Control, **48**, 2465.

E. Yaz (1989a) "Infinite horizon quadratic optimal control of a class of nonlinear stochastic systems," IEEE Trans. Autom. Control, **34**, 1176.

E. Yaz (1989b) "Robust design of stochastic controllers for nonlinear systems," IEEE Trans. Automatic Control, **34**, 349.

Y. I. Yaz and E. E. Yaz (1999a) "On LMI formulations of some problems arising in nonlinear stochastic system analysis," IEEE Trans. Automatic Control, **44**, 813.

E. E. Yaz and Y. I. Yaz (1999b) "Stochastic stabilizability of a class of discrete-time nonlinear stochastic systems," Proc. of 1999 IFAC World Congress, Beijing, China, **J**, 253.