ESTIMATION OF STATES AND PARAMETERS OF STOCHASTIC NONLINEAR SYSTEMS WITH MEASUREMENTS CORRUPTED BY CORRELATED NOISE

J. R. RAOL, N. K. SINHA

(Bangolore) (Hamilton)

We consider the problem of estimation of states and parameters of stochastic nonlinear *systems* described **by** difference equations when the measurements are corrupted **by** correlated noise. The solutions to the **filtering** problem are provided for certain general **as** well as a few special cases.

1. Introduction

In many practical systems the measurements are corrupted by coloured noise. Occasionally some of the measurements would be very accurate and hence they can be considered as perfect observations. These two cases represent singular problems in as much as the correlation matrix of the measurements noise can be shown to be singular [1].

We address the problem of state and parameter estimation for discrete-time systems when the noise in measurements is described by a stochastic difference equation. The noise can enter the measurements linearly or nonlinearly. We obtain quite general results based on Pugachev's nonlinear filtering theory [2–7, 9] and its ramifications [10, 11].

Pugachev's approach is based on the criterion of the minimization of the mean-square error [MSE] in a class of measurements that satisfy stochastic differential or difference equations. The theory is based on an arbitrary assignment of the so-called structural functions of the filter. The estimator has variable coefficients as gains which are determined by processing all the a priori information on filtered as well as measured processes.

In [10, 11] the nonlinear filtering problem for stochastic differential systems with measurements containing correlated noise was solved. We obtain similar results for systems described by difference equations.

It must be noted here that linear filtering problems for singular noise in continuous and discrete systems have been solved in [12–16, 26].

2. Statement of the problem

Let us consider the stochastic system represented by the difference equation

$$Y_{j+1} = g(Y_j, j) + p(Y_j, j)V_j \tag{1}$$

and the discrete measurements by the equation

$$W_j = l(Y_j, U_j, j). \tag{2}$$

The measurements noise U is considered to be correlated and is given by the difference equation

$$U_{j+1} = b(U_j, j) + d(U_j, j)V_j. \tag{3}$$

In these equations we have

Y_j — $n \times 1$ stochastic state process
W_j — $m \times 1$ measurement process
U_j — $m \times 1$ measurement noise process
g, l, b — known nonlinear vector-valued functions
p, d — known nonlinear matrix valued functions
V_j — vector-valued sequence of independent random numbers with zero-mean and covariance matrix Q_j.

Given Eqs (1) to (3) and the observations at instants $j = 1, 2, \ldots, N$, it is required to obtain the optimal estimate $\hat{Y}_j$ of the process Y_j in the class of functionals of the measurement process W_j.

We see from Eqs (1) and (3) that the two processes Y and U can be jointly represented as:

$$X_{j+1} = f(X_j, j) + q(X_j, j)V_j \tag{4}$$

$$W_j = h(X_j, j) \tag{5}$$

wherein

$$\left. \begin{aligned} X_j &= [Y_j^T, U_j^T]^T \\ f(X_j, j) &= [g(Y_j, j)^T, b(U_j, j)^T]^T \\ q(X_j, j) &= [p(Y_j, j)^T, d(U_j, j)^T]^T \\ h(X_j, j) &= l(Y_j, U_j, j). \end{aligned} \right\} \tag{6}$$

and

Here f, q and h are known nonlinear functions of appropriate dimensions. The nonlinear functions are such that the expectations, to be defined later on, exist.

Thus the problem of Eqs (1) to (3) reduces to that of Eqs (4) to (6). From Eq.(5) it can be seen that the measurement process depends on a process that is not available. Also it does not contain an independent noise process V_j explicitly.

Оценка состояний и параметров стохастических нелинейных систем при измерениях с коррелированным мешающим шумом

Й. Р. РАОЛ, Н. К. СИНХА

(Бангалоре)   (Хамилтон)

Рассматривается проблема оценки состояний и параметров стохастических нелинейных систем, описываемых разностными уравнениями, когда при измерениях мешающий шум является коррелированным. Даются решения проблемы фильтрации для некоторых общих и частных случаев.

J. R. Raol
Systems Engineering Division
National Aeronautical Laboratory
Bangalore — 560017
India

N. K. Sinha
Department of Electrical and Computer Engineering
McMaster University
Hamilton
Canada
LNS 4L7

The filtering problem posed by **Eqs (4)** to (6) can be considered as a special case of a more general problem [10, 11] described by the following **Eqs:**

$$X_{j+1} = f(X_j, Z_j, j) + q(X_j, Z_j, j) V_j \tag{7}$$

$$Z_{j+1} = h(X_j, Z_j, j) \tag{8}$$

The problem is then to obtain an estimator $\hat{X}$ for the state X utilizing the known samples of process **Z**. We provide the solution to this problem in the next section.

3. Solution of the problem

A. *General Solution*

Let the estimator for the system of Eqs (7) and (8) be given by the difference equation

$$\hat{X}_{j+1} = K \zeta(Z_j, Z_{j+1}, \hat{X}_j, j) + \gamma \tag{9}$$

where

K, y — optimal (j-dependent) filter gains

ζ — preassignable vector-valued nonlinear function of appropriate dimension.

The function ζ determines the structure of the nonlinear estimator (NE). Alternatively it is possible to introduce the concept of permissible estimates in order to make the comparison of accuracy and performance of the filters having similar structures feasible [7, 17].

The time-dependent optimal gains K and γ are determined from the condition of minimum of **MSE**

$$E\{(X_j - \hat{X}_j)^T (X_j - \hat{X}_j)\}$$

and are given by (see the Appendix):

$$KM = L \tag{10}$$

and

$$\gamma = E_0 - KE_1$$

where

$$\left. \begin{array}{l} L = E\{(X_{j+1} - EX_{j+1}) \zeta(Z_j, Z_{j+1}, \hat{X}_j, j)^T \\ M = E\{(\zeta(Z_j, Z_{j+1}, \hat{X}_j, j) - E_1) \zeta(Z_j, Z_{j+1}, \hat{X}_j, j)^T \\ E_0 = E\{X_{j+1}\}; \ E_1 = E\{\zeta(Z_j, Z_{j+1}, \hat{X}_j j)\} \end{array} \right\} . \tag{11}$$

Here 'E stands for the mathematical expectation and 'T' for matrix/vector transposition.

In order to evaluate the indicated expectations it is sufficient to know the joint characteristics function of the processes X_j, Z_j and $\hat{X}_j$, which **is** given as:

$$\Phi_j(\lambda_1, v, \mu) = E \exp \{ i\lambda_1^T [f(X_{j-1}, Z_{j-1}, j-1) + q(X_{j-1}, Z_{j-1}, j-1)V_{j-1}] +$$

$$+ iv^T h(X_{j-1}, Z_{j-1}, j-1) + i\mu^T [K_{j-1}\zeta(Z_{j-1}, Z_j, \hat{X}_{j-1}, j-1) + Y_{j-1}] \}. \quad (12)$$

The evaluation of expectations in Eq. (11) by utilizing the characteristic function Eq. (12) and substitution of the optimal gains in (9) completely solves the general filtering problem as posed in Section 2.

The above results can be used to estimate all the components of the vector X or to obtain estimates of some of the components of X by making the preassignable structural function ζ dependent on the corresponding subset of the vector $\hat{X}$.

B. Correlated measurement noise

As mentioned earlier, the general solution can be utilized to obtain the solution of the filtering problem with singular measurement noise for $Z_{j+1} = W_j$.

Let the system be described by Eqs. (1) to (3). The estimator for the process Y is then given as

$$\hat{Y}_{j+1} = K\zeta(Z_j, W_j, Y_j, j) + \gamma \quad (13)$$

where ζ is independent of $\hat{U}$.

The optimal values of the gains K and y can be obtained from Eqs (10) to (13) by establishing the following relations between functions (see Eqs **(4)** to (8)):

$$\lambda = [\lambda_1^T \lambda_2^T]_T, \quad q(X_j, Z_j, j) = [p(Y_j, j)^T d(U_j, j)^T]^T$$

$$f(X_j, Z_j, j) = [g(Y_j, j)^T, b(U_j, j)^T]^T$$

$$h(X_j, Z_j, j) = l(Y_j, U_j, j)$$

$$\zeta(Z_j, Z_{j+1}, \hat{X}_j, j) = \zeta(Z_j, W_j, \hat{Y}_j, j) \qquad \left. \right\} \cdot \quad (14)$$

Consequently the intermediate gains are given by

$$L = E\{(Y_{j+1} - EY_{j+1})\zeta(Z_j, W_j, \hat{Y}_j, j)^T\}$$

$$M = E\{(\zeta(Z_j, W_j, \hat{Y}_j, j) - E_1)\zeta(Z_j, W_j, \hat{Y}_j, j)^T\} \qquad \left. \right\} \cdot$$

and
$$E_0 = E\{Y_{j+1}\}; \quad E_1 = E\{\zeta(Z_j, W_j, \hat{Y}_j, j)\} \qquad \left. \right\} \quad (15)$$

Using the relationships shown in Eq. (14) we can obtain from **Eq.** (12), the joint characteristic function of the processes Y_j, U_j, Z_j and $\hat{Y}_j$:

$$\Phi_j(\lambda_1, \lambda_2, v, \mu) = E \exp \{ i\lambda_1^T [g(Y_{j-1}, j-1) + p(Y_{j-1}, j-1)V_{j-1}]$$

$$+ i\lambda_2^T [b(U_{j-1}, j-1) + d(U_{j-1}, j-1)V_{j-1}] + iv^T l(Y_{j-1}, U_{j-1}, j-1)$$

$$+ i\mu^T [K_{j-1}\zeta(Z_{j-1}, l(Y_{j-1}, U_{j-1}, j-1), \hat{Y}_{j-1}, j-1) + Y_{j-1}] \}. \quad (16)$$

Thus Eq. (10), with the characteristic function given by Eq. (16) to evaluate the expectations in Eq. (15), completely solves the problem of discrete-time nonlinear filtering with measurements contaminated by coloured noise.

The issues related to selection of appropriate structural function ζ and computation of optimal gains are addressed in Section 7 and [7, 17, 18].

4. Special cases

In order to elucidate the results of the previous section, we consider certain special structural functions and derive filters for linear systems.

A. *Linear Filter from General Solution*

Let the linear system be described by

$$X_{j+1} = FX_j + BV_j \tag{17}$$

where
$$W_j = HX_j$$

F, B, H — known system matrices **of** appropriate dimensions
V_j — zero-mean WGN sequences with correlation matrix Q
X_j, W_j — state and measurement processes.

Let the structural function be selected as:

$$\zeta = [\hat{X}_j^T, W_j^T]^T.$$

Consequently **we** have from Eq. (9)

$$\hat{X}_{j+1} = K'[\hat{X}_j^T, W_j^T]^T + Y$$

$$= G\hat{X}_j + KW_j + Y \tag{18}$$

where $K' = [G|K]$.

The optimal values of the gains are obtained by using Eq. (17) and ζ in **Eqs** (10) and (11):

$$GP_{\hat{x}_j} + KHP_{\hat{x}_j} = FP_{\hat{x}_j}$$

$$GP_{\hat{x}_j}H^T + KHP_{\hat{x}_j}H^T = FP_{x_j}H^T$$

and solving the above equations we obtain
$$G = F - KH$$

$$K = FPH^T(HPH^T)^{-1} \tag{19}$$

where P_x, $P_{\hat{x}}$ are the variance-covariance matrices of vector X and $\hat{X}$ and $P = P_x - P;$.

The vector γ **is** given by

$$\gamma = FE\{X_j\} - GE\{\hat{X}_j\} - KHE\{X_j\}.$$

Substituting the value of G from Eq. (19) and noting that $E\{X_j\} = E\{\hat{X}_j\}$ we see that $\gamma = 0$.

Substituting **Eq.** (19) in (18) we obtain

$$\hat{X}_{j+1} = F\hat{X}_j + K[W_j - H\hat{X}_j].$$ (20)

The covariance matrix equation for the state error is easily obtained and is given by

$$P_{j+1} = FP_jF^T - KHP_jF^T - FP_jH^TK^T + KHP_jH^TK^T + BQB^T$$

$$= (F - KH)P_j(F - KH)^T + BQB^T.$$

After substituting **Eq.** (19) in the above equation and simplifying **we** obtain

$$P_{j+1} = (F - KH)P_jF^T + BQB^T.$$ **(21)**

Equations (19) to (21) describe the linear filtering algorithm.

B. *Correlated* **Noise Case**

Let the linear system be described by

$$Y_{j+1} = FY_j + BV_j$$ (22)

$$W_j = HY_j + DU_j$$ (23)

$$U_{j+1} = AU_j + SV_j$$ **(24)**

where F, B, H, D, **A** and S are known matrices of appropriate dimensions. The correlated measurement noise U_j is described by difference Eq. (24).

We choose the structural function ζ as

$$\zeta(Z_j, W_j, \hat{Y}_j, j) = [\hat{Y}_j^T, W_j^T]^T$$ (25)

and consequently the estimator, Eq. (13), takes the form

$$\hat{Y}_{j+1} = G\hat{Y}_j + KW_j + \gamma.$$ (26)

The intermediate gains L and **M** can be obtained by using Eqs **(22)** to (24) in Eqs (10) and (11):

$$L = [FP_{Y_j\hat{Y}_j} | FP_{Y_j}H^T + FP_{Y_jU_j}D^T]$$ (27)

and

$$M = \begin{bmatrix} P_{\hat{Y}_j} & P_{\hat{Y}_j}H^T + P_{\hat{Y}_jU_j}D^T \\ HP_{\hat{Y}_j} + DP_{\hat{Y}_jU_j} & HP_{\hat{Y}_j}H^T + DP_{U_j}D^T + HP_{Y_jU_j}D^T + DP_{U_jY_j}H^T \end{bmatrix}$$

Then optimal gains G and K can be obtained from the following equations:

$$(F - G - KH)P_{\hat{Y}_j} - KDP_{\hat{Y}_j U_j} = 0$$

$$(F - KH)(P_{Y_j}H^T + P_{Y_j U_j}D^T) - G(P_{\hat{Y}_j}H^T + P_{\hat{Y}_j U_j}D^T) - KD(P_{U_j Y_j}H^T + P_{U_j}D^T) = 0$$

$$\gamma = (F - G - KH)E\{Y_j\} - KDE\{U_j\}. \tag{28}$$

In order to obtain the variance-covariance matrices in **Eq.**(28), we utilize Eqs (22), **(24)** and **(26)** to get

$$\begin{bmatrix} Y_{j+1} \\ U_{j+1} \\ \hat{Y}_{j+1} \end{bmatrix} = \begin{bmatrix} F & 0 & 0 \\ 0 & A & 0 \\ KH & KD & G \end{bmatrix} \begin{bmatrix} Y_j \\ U_j \\ \hat{Y}_j \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \gamma \end{bmatrix} + \begin{bmatrix} B \\ S \\ 0 \end{bmatrix} V_j. \tag{29}$$

From the above **Eq.** (29) we obtain the following equation for the covariance matrix P_j for the random vector $[Y_j^T U_j^T \hat{Y}_j^T]^T$:

$$\mathbf{P}_{j+1} - \mathbf{F}\mathbf{P}_j\mathbf{F}^T + \mathbf{B}Q_j\mathbf{B}^T \tag{30}$$

where $\mathbf{F}$ is the block matrix in Eq. (29), $\mathbf{B} = [B^T S^T 0^T]^T$ and 0^T is a null matrix with appropriate dimensions.

Equations (26),(28) and (30) provide the solution to the filtering problem with correlated measurement noise. It may be noted that this solution does not need differencing of the measurement data as required by certain existing linear filtering results [I, 10, 11].

5. Joint state and parameter estimation

For simplicity let the scalar system be described by

$$X_{j+1} = -\vartheta X_j + b_1 V_1 \tag{31}$$

$$W_j = X_j + U_j \tag{32}$$

$$U_{j+1} = -aU_j + b_2 V_2 \tag{33}$$

where

ϑ — an unknown parameter

b_1, b_2, a — known constants.

We represent the unknown parameter ϑ as a stochastic process.

$$\Theta_{j+1} = \Theta_j. \tag{34}$$

From Eqs (31) and (34) we obtain

$$\begin{bmatrix} X_{j+1} \\ \Theta_{j+1} \end{bmatrix} = \begin{bmatrix} -\Theta_j X_j \\ \Theta_j \end{bmatrix} + \begin{bmatrix} b_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}. \tag{35}$$

Also from Eq. (33) we obtain

$$U_{j+1} = -aU_j + [0 \quad b_2] \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}. \tag{36}$$

We specify the optimal estimator *as:*

$$[\hat{X}_{j+1} \hat{\Theta}_{j+1}]^T = K\zeta(\hat{X}_j, \hat{\Theta}_j, W_j, j) + \gamma. \tag{37}$$

We then select K, ζ and γ as

$$K = \begin{bmatrix} k_{11} & k_{12} & \beta_1 \\ k_{21} & k_{22} & \beta_2 \end{bmatrix}; \ \zeta = [\hat{X}_j, \hat{\Theta}_j, W_j]^T$$

and

$$\gamma = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}. \tag{38}$$

The intermediate gains are obtained as (dropping j for simplicity):

$$L_j = \begin{bmatrix} -P_{\theta x \hat{x}} & -P_{\theta x \hat{\theta}} & -P_{\theta x w} \\ P_{\theta \hat{x}} & P_{\theta \hat{\theta}} & P_{\theta w} \end{bmatrix}_j$$

$$M_j = \begin{bmatrix} P_{\hat{x}} & P_{\hat{x}\hat{\theta}} & P_{\hat{x}w} \\ P_{\hat{\theta}\hat{x}} & P_{\hat{\theta}} & P_{\hat{\theta}w} \\ P_{w\hat{x}} & P_{w\hat{\theta}} & P_w \end{bmatrix}_j$$

and

$$K = L_j M_j^{-1} \tag{41}$$

$$E_0 = \begin{bmatrix} -E\{\Theta_j X_j\} \\ E\{\Theta_j\} \end{bmatrix}$$

$$E_1 = [E\{\hat{X}_j\} E\{\hat{\Theta}_j\} E\{X_j + U_j\}]^T. \tag{42}$$

The entries in Eqs (39) and (40) are the variance and covariance matrices of the random variables $X_j, \hat{X}_j, \Theta_j, \hat{\Theta}_j, W_j$ and $\Theta_j X_j$. In order to simplify these equations we can use the following relations

$$P_{X_j W_j} = P_{\hat{X}_j X_j} + P_{\hat{X}_j U_j}$$

$$P_{\theta_j W_j} = P_{\hat{\theta}_j X_j} + P_{\hat{\theta}_j U_j} \tag{43}$$

$$P_{W_j} = P_{X_j} + P_{U_j}$$

$$P_{X_j \hat{X}_j} = P_{\hat{X}_j}; \ P_{X_j W_j} = P_{\hat{X}_j W_j} \tag{44}$$

and

$$P,_{\hat{\rho}_j} = P'_{\hat{x}_j \hat{\theta}_j}.$$

The relationships in **Eq. (43)** follow from **Eq.** *(32)* and those in **Eq. (44)** follow from $P_{x\zeta} = P'_{\hat{x}\zeta}$ (see **(AS)**).

The joint one-dimensional characteristic function of the processes X_j, Θ_j, U_j, $\hat{X}_j$ and $\hat{\Theta}_j$ is expressed as:

$$\Phi_j(\lambda_1, \lambda_2, v, \mu_1, \mu_2) = \psi_{j-1}(\lambda) E \exp \{ i\lambda_1(-\Theta_{j-1} X_{j-1}) + i(\mu_1\beta_1 + \mu_2\beta_2)\hat{X}_{j-1}$$

$$+ i(\mu_1\beta_1 + \mu_2\beta_2 - va) + i(\mu_1 k_{11} + \mu_2 k_{21})\hat{X}_{j-1}$$

$$+ i(\mu_1 k_{12} + \mu_2 k_{22})\hat{\Theta}_{j-1} + i\mu_1 Y_1 + i\mu_2 Y_2 \} \tag{45}$$

where $\psi_{j-1}(\lambda)$ is the characteristic function of the random process

$$V_{j-1} = [b_1 V_1 | b_2 V_2]^T \quad \text{and} \quad \lambda^T = [\lambda_1 v].$$

We see from **Eq. (45)** that the characteristic function $\Phi_j(\lambda_1, \lambda_2, v, \mu_1, \mu_2)$ is given in terms of the random variables X_{j-1}, Θ_{j-1}, U_{j-1}, $\hat{X}_{j-1}$, $\hat{\Theta}_{j-1}$ and the probability distribution of the variable V_{j-1}. This latter distribution is assumed to be known. Hence Eqs (39) to (42) and (45) provide recursive procedure for determining Φ_j, K_j arid γ_j at each step with the knowledge of Φ_{j-1}, K_{j-1} and γ_{j-1}. The choice of initial conditions is thoroughly discussed in [9].

6. Some generalizations

In this section we obtain the expressions for the optimal gains for the estimators described by [6, 10, 11]:

$$\hat{X}_{j+s} = \delta_j \zeta_j(Z_j, Z_{j+1}, \hat{X}_j, \ldots, \hat{X}_{j+s-1}) + \gamma, \tag{46}$$

for the system described by **Eqs** (7) and (8).

Here s — the order which is specified a priori

   ζ_j — preassignable nonlinear structural function

   δ_j, γ_j — optimal gains.

These gains are given by

$$\delta_j M_j = L_j$$

$$\gamma_j = E_{0j} - \delta_j E_{1j} \tag{47}$$

where

$$L_j = E\{(X_{j+s} - EX_{j+s})\zeta_j(Z_j, Z_{j+1}, \hat{X}_j, \ldots, \hat{X}_{j+s-1})^T\}$$

$$M_j = E\{((\zeta_j(Z_j, Z_{j+1}, \hat{X}_j, \ldots, \hat{X}_{j+s-1}) - E_{1j})\zeta_j(Z_j, Z_{j+1}, \hat{X}_j, \ldots, \hat{X}_{j+s-1})^T\}$$

$$E_{0i} = E\{X_{j+s}\}; \quad E_{1j} = E\{\zeta_j(Z_j, Z_{j+1}, \hat{X}_j, \ldots, \hat{X}_{j+s-1})\}. \tag{48}$$

The one-dimensional joint characteristic function of the processes X_j, Z_j, $\hat{X}_j$, ..., X_{j+s-1} is given by

$$\Phi_j(\lambda, v, \mu) = E \exp\{i\lambda^T[f(X_{j-1}, Z_{j-1}, j-1) + q(X_{j-1}, Z_{j-1}, j-1)V_{j-1}]$$

$$+ iv^T h(X_{j-1}, Z_{j-1}, j-1) + i\sum_{r=1}^{s-1} \mu_r^T \hat{X}_{j+r-1}$$

$$+ i\mu_s^T[\delta_{j-1}\zeta_{j-1}(Z_{j-1}, Z_j, \hat{X}_{j-1}, ..., \hat{X}_{j-s2}) + \gamma_{j-1}]\}. \tag{49}$$

In the above equation μ_1, μ_2, ... μ_s are the columns of the matrix μ. We note here, that since V_{j-1} is independent of other random processes, we have recursive method for solving Eq. (49) and obtaining the optimal gains from **Eqs** (47) and (48) based on the knowledge of the a priori distribution of the variable V_{j-1} and the characteristic function $\Phi_{j-1}(\lambda, v, \mu)$ of the random processes X_{j-1}, Z_{j-1}, $\hat{X}_{j-1}$, ..., $\hat{X}_{j-s-2}$. The estimates $\hat{X}_1$, ..., $\hat{X}_s$ of the variables X_1, ..., X_s should be selected arbitrarily [6]. **Also** the joint characteristic function $\Phi_1(\lambda, v, \mu)$ of the variables X_1, Z_1, $\hat{X}_1$, ..., $\hat{X}_s$ be selected such that proper solution of the above filtering problem is obtained without any ambiguity [6].

We have seen that when the measurements contain correlated noise, the system is described by the set of equations (I) to (3). We can now obtain the solution to this filtering problem for any value of's' from the **above results.** In this case the estimator is given by

$$\hat{Y}_{j+s} = \delta_j\zeta_j(Z_j, W_j, \hat{Y}_j, ..., \hat{Y}_{j+s-1}) + \gamma_j \tag{50}$$

The optimal valuer of the variable gains are obtained from Eqs (47) to (49) by the following replacements:

$$X_j \rightarrow Y_j; \quad h(X_j, Z_j, j) \rightarrow l(Y_j, U_j, j)$$

and

$$\zeta(Z_j, Z_{j+1}, \hat{X}_j, ..., \hat{X}_{j+s-1}) \rightarrow \zeta(Z_j, W_j, \hat{Y}_j, ..., \hat{Y}_{j+s-1}).$$

The required expectations can be evaluated by knowing the joint one-dimensional characteristic function of the random processes Y_j, U_j, Z_j, $\hat{Y}_j$, ..., $\hat{Y}_{j+s-1}$, which is given by

$$\Phi_j(\lambda_1, \lambda_2, v, \mu) = E \exp\{i\lambda_1^T[g(Y_{j-1}, j-1) + p(Y_{j-1}, j-1)V_{j-1}]$$

$$+ i\lambda_2^T[b(U_{j-1}, j-1) + d(U_{j-1}, j-1)V_{j-1}]$$

$$+ iv^T l(Y_{j-1}, U_{j-1}, j-1) + i\sum_{r=1}^{s-1} \mu_r^T \hat{Y}_{j+r-1} \tag{51}$$

$$+ i\mu_s^T[\delta_{j-1}\zeta_{j-1}(Z_{j-1}, l(Y_{j-1}, U_{j-1}, j-1), \hat{Y}_{j-1}, ..., \hat{Y}_{j+s+2}) + \gamma_{j-1}]\}.$$

As per the remarks given after Eq. (49) the recursive procedure is established to solve Eqs (47), (48) and (51), thereby yielding the optimal filter gains δ_j and y_j.

7. Computational aspects

From the results of the previous sections it is apparent that two major issues are involved in the design of the nonlinear estimators: I) the proper selection of the structural function ζ, and 2) computations of the optimal gains.

Due to the nature of the estimator equations we obtain conditionally optimal filters and hence there is an additional degree of freedom for selection of the structural functions. **As** yet no systematic procedure seems to have been established. However based on the studies in [19] some guidelines are available. In most cases the structural function is selected such that it is very similar to the system under consideration.

The problem of selection of structural functions has not been fully resolved. For further details, one may refer to [27].

In [6, **7**, 9, 17, 18] certain special functions were selected in order to obtain some conventional and simple filtering structures for related problems. Such studies may also be extended for the results of this paper. Based on such choices it may be possible to establish interconnections with other familiar results [14, 16, 26]. It must be noted here that, the results of the present paper are general in nature, though linear filtering solutions have been obtained as special cases. A recent study [20, 21] may be useful in selecting the structural functions for the estimators proposed in this paper and, in general, for Pugachev's nonlinear estimators [6, 7, 10, 11]. Further research is required in this direction.

The second aspect is the evaluation of the expectations for computing the optimal gains of the filters. These expectations are functionals of the characteristic function Φ_j. These equations, in general, represent complex functional difference equations. Some approximation methods can be used *for* this purpose. Such methods are discussed in **[9, 17**, 18, **22–25].**

Once a proper choice of the structural function ζ is made, the optimal gains **K,** G, y (or 6) can be obtained by any of the methods described in [9, 19, 23] utilizing only a priori information on the various processes involved. Thus it is possible to design an optimal filter completely before collecting actual measurement data from the system. The subsequent computation of the estimates involves only matrix-vector operations making the actual estimation processor very efficient. The precomputed gains can be stored and fetched from microcomputer's memory when the observation data become available for processing. The estimator retains sequential nature. The results of this paper can be used for wide variety of problems: identification, joint state and parameter estimation, stochastic model building and control [5, 9].

Conclusions

The conditionally optimal solutions for nonlinear filtering problem described by stochastic difference equations with measurements containing correlated noise have been obtained. We have obtained filters for some general and special linear cases. The filter for joint state and parameter estimation for correlated noise case has been derived for a scalar case. The advantages of the present method are: the measurement model can be linear or nonlinear in the correlated noise and the filter's conditionally optimal gains can be precomputed before actually processing the measurement data. This makes estimation procedure very efficient and simple while incorporating the observations sequentially. Some computational aspects have been discussed.

References

1. *Maybeck, P. S.* (1982). Stochastic Models. Estimation and Control. Academic Press, New York.
2. *Pugachev, V. S.* (1978a), Estimation of variables and parameters in stochastic systems described by differential equations. Soviet Math. Doll.. 19, 967.
3. *Pugachev, V. S.* (1978b), Recursive estimation of variables and parameters in stochastic systems described by autoregression equations. Soviet Math. Dokl., 19, 991.
4. *Pugachev, V. S.* (1978c), Recursive estimation o lvariables and parameters in stochastic systems described by difference equations. Soviet Math. Dokl., 19, 1495.
5. *Pugachev, V. S.* (1979a), Estimation of state and parameters of continuous nonlinear systems. Autom. and Remote Control. 40, 831.
6. *Pugachev, V. S.* (1979b), Estimation of variables and parameters in discrete-time nonlinear syrtenis. Autom. and Remote Control. 40, 512.
7. *Pugachev, V. S.* (1982), Conditionally optimal estimation in stochastic differential systems. Automatica, 18, 685.
X. *Pugachev, V. S.* (1965), Theory of Random Functions and its Application lo Control Problems. Pergamon Press, New York.
9. *Pugachev, V. S.* (1984), Probability Theory and Mathematical Statistics fur Engineers. Pergamon Press, New York.
10. *Silujanova, I. D.* (1980), Estimating the variables and parameters in stochastic systems described hy differential equations in a particular case. Soviet Math. Dokl. 21, 152.
11. *Silujanova, I. D.* (1980), State and parameter estimation of nonlinear systems using observations containing colored noise. Autom. and Remote Control. 41, 1380.
12. *Bryson, A. E., Johansen, D. E.* (1965), Linear filtering lor time-varying systems using measurements containing colored noise. IEEE Trans. Automat. Control. AC-10, NO. 1, 4.
13. *Bucy, R. S.* (1967), Optimal filtering for correlated noise. J. of Math. Anal. and App. 20, 1.
14. *Bryson, A. E., Henrikson, L. J.* (1968), Estimation using sampled data containing sequentially correlated noise. J. Spacecraft and Rockets. 6/61, 662.
15. *Stear, E. B., Stubberud, A. R.* (1968), Optimal filtering lor Gauss–Markov noise. Int. J. Control. 8(2), 123.
16. *Deyst, J. J.* (1969), A derivation of the optimum continuous linear estimator lor systems with correlated measurement noise. AIAA Journal. 7(11), 2116.
17. *Raol, J. R., Sinha, N. K.* (1986), On Pugachev's filtering theory lor stochastic nonlinear systems. presented at the Second IFAC Symposium on Stochastic Control (Vilnius. U.S.S.R.).
18. *Raol, J. R., Sinha, N. K.* (1986), A nonlinear filter lor estimation of states of a continuous-time system with discrete measurements, presented at the Second IFAC Symposium on Stochastic Control (Vilnius. U.S.S.R.).

19. *Kazakov, I. E., Malchikov, S. I.* (1981), Approximate design of Pugachev filters of specified complexity. Autom. and **Remote** Control. **42,** 1309.
20. *Ahlbehrendt, N., Madiger, B.* (1981), A new method for synthesis of nonlinear parameter and state estimation for noisely disturbed processes. Proc. of the VIIIth IFAC Congress. Kyoto, Japan, v. 99.
21. *Ahlbehrendt, N., Draeger, U., Madiger, B.* (1985), Synthesis of estimation algorithms for stochastic disturbed systems by the help of Lyapunov functions. Problems of Control and Information Theory. 14(1), 47.
22. *Dashevskii, M. L., Shin. I! I.* (1981), Design of conditionally optimal filters in continuous dynamic systems. Autom. and Remote Control, **42,** 1309.
23. *Demukh, V. I.* (1965), An approximate method for analysing accuracy of nonlinear systems. Autom. and Remote Control. **26,** 1012.
24. *Dashevskii, M. L., Liptser, R. Sh.* (1967), Use of conditional semiinvariants in nonlinear filtering of Markov processes. Autom. and Remote Control, 28. **63.**
25. *Malchikov, S. I!* (1970). Approximate method of determining the phase coordinate distributions of nonlinear automatic systems. Autom. and Remote Control. 31. 717.
26. *Sinha, N. K., Tom,* A. (1978), Adaptive state estimation for systems with colored observation noise. Int. J. Syrt. Sc.. 9. 31.
27. *Raol, J. R., Sinha, N. K.* (1987), Conditionally optimal state estimation for systems governed by difference equations. Can. Clcc. Eng. J.. **12,** 71.

Appendix

For stochastic difference system described by Eqs (7) and (8), we have the following difference equation for the estimator:

$$\hat{X}_{j+1} = K_j\zeta_j(Z_j, Z_{j+1}, \hat{X}_j, j) + \gamma_j. \tag{A1}$$

Then using the mean-square regression theory [6, 8, 9], we obtain

$$E\{[X_{j+1} - E\{X_{j+1}\}]\zeta(Z_j, Z_{j+1}, \hat{X}_j, j)^T\} = K_j E\{[\zeta(Z_j, Z_{j+1}, \hat{X}_j, j) - E\{\zeta(Z_j, Z_{j+1}, \hat{X}_j, j)\}]\zeta(Z_j, Z_{j+1}, \hat{X}_j, j)^T\} \tag{A2}$$

and

$$\gamma_j = E\{X_{j+1}\} - KE\{\zeta(Z_j, Z_{j+1}, \hat{X}_j, j)\}. \tag{A3}$$

Defining L and M as in (11) we obtain for optimal gain the following relation

$$K_j = L_j M^{j-1}. \tag{A4}$$

Also we have the condition

$$E\{(\hat{X}_{j+1} - X_{j+1})\zeta_j(Z_j, Z_{j+1}, \hat{X}_j, j)^T\} = 0. \tag{A5}$$

The characteristic function Φ_j for the random processes X_j, Z_j and $\mathbf{\mathcal{S}}$, can be obtained by substituting Eqs (7) to (9) in the following equation

$$\Phi_j(\lambda_1, v, \mu) = E \exp\{i\lambda_1^T X_j + iv^T Z_j + i\mu^T \hat{X}_j\}. \tag{A6}$$

When the estimator is specified by the following structure

$$\hat{X}_{j+1} = G\xi(Z_j, \hat{X}_j, j) + K\eta(Z_j, \hat{X}_j, j)Z_{j+1} + \gamma \tag{A7}$$

the optimal gains G, K and y can be obtained by representing the structural function as

$$\zeta(Z_j, Z_{j+1}, \hat{X}_j, j) = [\xi(Z_j, \hat{X}_j, j)^T | Z_{j+1}{}^T\eta(Z_j, \hat{X}_j, j)^T]^T$$

in (A?) and (A3).

Similarly the characteristic functional equation (12) can be modified. This representation is useful in obtaining various special cases as discussed in the text.