A SOLUTION OF THE FILTERING AND SMOOTHING PROBLEMS FOR UNCERTAIN-STOCHASTIC DIFFERENTIAL SYSTEMS

ANDREI V. BORISOV
Department of Industrial & Operations Engineering
University of Michigan
Ann Arbor, MI 48109-2117

ALEXEI R. PANKOV
Department of Applied Mathematics
Moscow Aviation Institute
Moscow,127080, Russia

Technical Report 91-33

November, 1991.

A SOLUTION OF THE FILTERING AND SMOOTHING PROBLEMS FOR UNCERTAIN-STOCHASTIC DIFFERENTIAL SYSTEMS

ANDREI V. BORISOV
Department of Industrial Operations Engineering
University of Michigan
Ann Arbor, MI 48109-2117

and

ALEXEI R. PANKOV
Department od Applied Mathematics
Moscow Aviation Institute
Moscow, 127080, Russia

(Revised on November 14, 1991)

In this paper we present a new filtering algorithms for uncertain-stochastic dynamic systems, which are optimal in the sense of some minimax-stochastic criterion. These algorithms give a possibility to estimate the state vector of dynamic systems under lack of a priori information about the system characteristics using observations of the state and input signals. The algorithms allow us to construct an optimal two-filter smoothing algorithm for pure uncertain dynamic systems.

These algorithms are numerically tested, results are compared with the results of Kalman filtering and smoothing in the case of complete information about input signal characteristics.

1 Introduction.

The Kalman filter (see Kalman and Bucy, 1961) is known as the best mean square estimator of state vector of dynamic system of observation. First time using of filtering algorithms was limited by some strict conditions like linearity of system and observations, complete information about system and input signal characteristics, Gaussianity all of noises. After that there were a lot of trials to extend the class of systems of observations for which filtering algorithms also may be used efficiently. It was obtained general equation for nonlinear filtering and filtering of conditionally-gaussian processes (see Liptser and Shiryayev, 1977), filtering of diffusion processes with Poisson-type of observations (Pardoux, 1979) and filtering of Poisson processes in Gaussian noise (Hero, 1991) , filtering of Ito-Volterra processes (Kleptsina and Veretennikov, 1985). The necessary and sufficient conditions for finite dimensionality of recursive filters were also presented in (Brockett and Clark, 1980) and (Tam et al., 1990). All these papers considered a filtering problem assuming complete information about system and input signal characteristics.

In many practical situations a priori information about system and inputs is incomplete. To overcome this serious obstacle one may use adaptive (Goodwin and Sin, 1984),(Fomin, 1984),(Eweda, 1991) , H^{∞} (Nagpal and Khargonekar, 1991), robust (Kassam and Poor, 1985), (Barton and Poor, 1990), minimax (Kurzhansky, 1977),(Malyshev and Kibzun , 1987) approaches. All these algorithms used limited amount of a priori information. In many applied cases it is possible to use both observations of state vectors and observations of some input signal components to obtain a high precision estimates even when we know nothing about input characteristics.

In this paper we consider new recursive filtering algorithms for uncertain-stochastic systems given state and input signal observations. These algorithms were presented in (Borisov and Pankov, 1991) briefly. Here, we prove the optimality of a filtering algorithm for continuous-time systems of observations. This algorithm gives us a possibility to design an optimal fixed-interval smoother for pure uncertain systems. The analog of this smoother for pure stochastic systems was considered in (Mayne, 1966) , (Wall et al., 1981).

We also consider the results of some numerical experiments. Accuracy of the new estimates in the absence of a priori information about inputs is compared with the accuracy of the Kalman filter estimates with completely known input signal characteristics. It turns out these accuracies are close to each other. Hence, we can use a numerically efficient filtering approach for the wide class of uncertain-stochastic systems and obtain acceptable estimates. From these experiments it also follows that the smoothing estimates are significantly more accurate than the corresponding filtering ones. Smoothing estimates are therefore more preferable in the case of postexperimental data processing.

2 Model description.

Consider the following dynamic system

$$dx_t = a(t)x_t dt + dw_t, \qquad t > 0; \qquad x_0 = \nu \tag{1}$$

where x_t is a state vector, ν is an uncertain initial condition, $x_t, \nu \in R^p$, w_t is an input signal, $a(t)$ is known matrix-valued function. Assume that w_t is given by

$$dw_t = b(t)u_t dt + d\xi_t, \tag{2}$$

where u_t is an uncertain vector, $u_t \in R^q$; $\{\xi_t\}$ is a p -dimensional zero-mean Wiener process with differential covariance matrix $C(t)$, $b(t)$ is a known matrix-valued function.

Information concerning $\{x_t\}, \{u_t\}$ and ν is obtained from the observation processes $\{y_t\}, \{z_t\}$ which are given by

$$\begin{cases} y_0 = \phi_0 \nu + \omega_0, \\ dy_t = \phi(t)u_t dt + d\omega_t, \\ dz_t = \psi(t)x_t dt + d\eta_t, \end{cases} \tag{3}$$

where $\{\omega_t\}$ and $\{\eta_t\}$ are zero-mean Wiener processes with the differential covariance matrices $Q(t)$ and $P(t)$ respectively, $\omega_t \in R^m, \eta_t \in R^n$; ω_0 is a zero-mean random vector with covariance matrix Q_0.
We assume for simplicity that $\omega_0, \{\xi_t\}, \{\omega_t\}$ and $\{\eta_t\}$ are independent.
The matrices $a(t), b(t), C(t), \phi(t), \psi(t), Q(t), P(t)$ have piecewise-continuous elements. Let $\mathbf{U_t}$ be a class of Lebesgue integrable functions on $[0, t]$, then for every ν and $\{u_\tau\} \in \mathbf{U_t}$ equation (1) defines the unique second order process $\{x_\tau\}$.

Now let us consider the linear discrete-time dynamic system

$$x_t = a_t x_{t-h} + w_t, \qquad x_0 = \nu, \quad t = h, 2h, ... \tag{4}$$

where x_t is a state vector, ν is an uncertain initial condition, a_t is known matrix , $h > 0$ is a time increment , w_t is an input signal given by

$$w_t = b_t u_t + \xi_t. \tag{5}$$

Here u_t is an uncertain vector, b_t is known matrix, $\{\xi_t\}$ is a zero-mean discrete random process with $cov(\xi_t, \xi_s) = C_t \delta_t^s$.

The corresponding observation model is given by

$$\begin{cases} y_0 = \phi_0 \nu + \omega_0, \\ y_t = \phi_t u_t + \omega_t, \\ z_t = \psi_t x_t + \eta_t, \end{cases} \tag{6}$$

where $\{\omega_t\}, \{\eta_t\}$ are zero-mean discrete random processes with $cov(\omega_t, \omega_s) = Q_t \delta_t^s$ and $cov(\eta_t, \eta_s) = P_t \delta_t^s$; ω_0 is a zero-mean random vector with covariance matrix Q_0 . We assume for simplicity that $\omega_0, \{\xi_t\}, \{\omega_t\}$ and $\{\eta_t\}$ are independent.

3 Filtering problem.

The linear filtering problem for the system (1)-(3) is to calculate the best linear estimate $\hat{x}_t$ given observations $\{y_\tau, z_\tau, 0 \leq \tau \leq t\}$ We consider the best linear estimate in the sense of minimizing the following criterion

$$J_t = \sup_{\nu \in R^p, \{u_\tau\} \in \mathrm{U_t}} \mathbf{E}\{(\mathbf{x_t} - \hat{\mathbf{x}}_\mathbf{t})^\mathbf{T} \boldsymbol{\Sigma}_\mathbf{t} (\mathbf{x_t} - \hat{\mathbf{x}}_\mathbf{t})\} \tag{7}$$

where $\Sigma \geq 0$ is known weight matrix.

Proposition 1. Assume that $\{x_t\}, \{y_t\}$ and $\{z_t\}$ are given by (1)-(3). Let

i) $Q_0 > 0, Q(\tau) > 0, P(\tau) > 0$ for all $\tau \in (0, t]$;
ii) $\phi_0^T \phi_0 > 0, \phi^T(\tau)\phi(\tau) > 0$ for all $\tau \in (0, t]$.

Then the J_t -optimal linear estimate $\hat{x}_t$ is unbiased. $\hat{x}_t$ and its error covariance matrix $k(t)$ are given by the following equations

$$d\hat{x}_t = a(t)\hat{x}_t dt + b(t)[\phi^T(t)Q^{-1}(t)\phi(t)]^{-1}\phi^T(t)Q^{-1}(t)dy_t + k(t)\psi^T(t)P^{-1}(t)[dz_t - \psi(t)\hat{x}_t dt];$$

$$\dot{k}(t) = a(t)k(t) + k(t)a^T(t) - k(t)\psi^T(t)P(t)^{-1}\psi(t)k(t) + b(t)[\phi^T(t)Q^{-1}(t)\phi(t)]^{-1}b^T(t) + C(t) \tag{8}$$

with initial conditions

$$\hat{x}_0 = [\phi_0^T Q_0^{-1}\phi_0]^{-1}\phi_0^T Q_0^{-1} y_0 \quad ; \quad k(0) = [\phi_0^T Q_0^{-1}\phi_0]^{-1} \tag{9}$$

Proof of Proposition 1 see in **Appendix A.**

Now let us consider the discrete-time model (4)-(6). Let $U_t = (u_t^T, ..., u_h^T, \nu^T)^T$ be the m -dimensional block vector,$m = (t/h)q + p$. The linear filtering problem for the system (4) is similar to the filtering problem for the system (1) but the optimality criterion slightly differs from (7)

$$J_t^* = \sup_{U_t \in R^m} \mathbf{E}\{(\mathbf{x_t} - \mathbf{x_t^*})^\mathbf{T}\mathbf{\Sigma_t}(\mathbf{x_t} - \mathbf{x_t^*})\} \tag{10}$$

Proposition 2. Assume that $\{x_t\}, \{y_t\}$ and $\{z_t\}$ are given by (4)-(6). Let

i) $Q_\tau > 0, P_\sigma > 0$ for all $\tau \in [0, t]$, $\sigma \in [h, t]$;
ii)$\phi_\tau^T \phi_\tau > 0$, for all $\tau \in [0, t]$;
iii)a_τ is nonsingular for all $\tau \in [h, t]$.

Then the J_t^*-optimal linear estimate x_t^* is unbiased . x_t^* and its error covariance matrix k_t are given by the following equations

$$\begin{aligned} \bar{x}_t &= a_t x_{t-h}^* + b_t[\phi_t^T Q_t^{-1}\phi_t]^{-1}\phi_t^T Q_t^{-1} y_t, \\ \bar{k}_t &= a_t k_{t-h} a_t^T + b_t[\phi_t^T Q_t^{-1}\phi_t]^{-1} b_t^T + C_t; \end{aligned} \tag{11}$$

$$\begin{array}{l} x_t^* = \bar{x}_t + \bar{k}_t\psi_t^T(\psi_t\bar{k}_t\psi_t^T + P_t)^{-1}(z_t - \psi_t\bar{x}_t) \\ k_t = \bar{k}_t - \bar{k}_t\psi_t^T(\psi_t\bar{k}_t\psi_t^T + P_t)^{-1}\psi_t\bar{k}_t \end{array} \tag{12}$$

with initial conditions

$$x_0^* = [\phi_0^T Q_0^{-1}\phi_0]^{-1}\phi_0^T y_0 \quad ; \quad k_0 = [\phi_0^T Q_0^{-1}\phi_0]^{-1}. \tag{13}$$

Proof of Proposition 2 see (Borisov, Pankov and Sotsky, 1991).

4 Fixed-interval smoothing problem.

The linear fixed-interval smoothing problem for the system (1)-(3) is to calculate for all $t \in [0,T]$ the best linear estimate $\hat{x}_t^s$ given observations $\{y_\tau, z_\tau, 0 \leq \tau \leq T\}$ .We consider the best linear estimate in the sense of minimizing the following criterion

$$\tilde{J}_t = \sup_{\nu \in R^p, \{u_\tau\}, \in \mathbf{U_T}} \mathbf{E}\{(\mathbf{x_t} - \hat{\mathbf{x}}_t^s)^\mathbf{T}\Sigma_t(\mathbf{x_t} - \hat{\mathbf{x}}_t^s)\}, \tag{14}$$

where $\Sigma \geq 0$ is a known weight matrix. The dynamic systems under investigation are purely uncertain, i.e. $C(t) \equiv 0$. The smoothing algorithm is based on the idea of a two-filter smoothing similar to one for pure stochastic systems (Mayne, 1966),(Wall et al., 1981). Let us consider a reversed time system

$$dx_t^r = -a(t)x_t^r dt - dw_t^r \quad , \quad x_T^r = \nu^r, \tag{15}$$

If $\nu^r = x_T$ and $u_t^r = u_t$ for all $t \in [0,T]$ then systems (1) and (15) define pathwise-equal processes. We assume existence of the final moment T state vector observation

$$y_T = \phi_T x_T + \omega_T, \tag{16}$$

where ω_T is a zero-mean random vector with $cov(\omega_T, \omega_T) = Q_T$. Proposition 1 being applied to the model (15) makes possible to obtain the backwards estimate $\hat{x}_t^r$ given observations $\{y_\tau, z_\tau, t < \tau \leq T\}$

i.e. the best linear estimate of x_t given the " future " observations. This estimate is unbiased with the error covariance matrix k_t^r.

Proposition 3. Assume that $\{x_t\}, \{y_t\}$ and $\{z_t\}$ are given by (1),(3),(16). Let a conditions of Proposition 1 hold and

i) the system (1),(2) is pure uncertain;

ii) $\phi_T^T \phi_T > 0, Q_T > 0$;

Then the $\tilde{J}_t$ -optimal unbiased smoothing estimate $\hat{x}_t^s$ and its error covariance matrix k_t^s are given by

$$\hat{x}_t^s = k_t^s[k_t^{-1}\hat{x}_t + (k_t^r)^{-1}\hat{x}_t^r] \quad ; \quad k_t^s = [k_t^{-1} + (k_t^r)^{-1}]^{-1} \tag{17}$$

Proof of the Proposition 3 see in **Appendix B.**

5 Numerical examples.

1. Let us consider the control system given by

$$\ddot{x} - \dot{x} + 0.25x = u_1 + 0.5\dot{u}_1 + u_2 + \dot{\xi}, \quad x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0 \tag{18}$$

where $\{u_1\}, \{u_2\}$, are unknown input signals, $x_0, \dot{x}_0$ are unknown initial conditions, $\{\xi\}$ is the Wiener process with differential covariance $C = 2500$. Equation (18) may be rewritten as the first order differential system

$$\begin{cases} \dot{x}_1 = x_2 + 0.5u_1 \\ \dot{x}_2 = -0.25x_1 + x_2 + 1.5u_1 + u_2 + \dot{\xi} \end{cases} \tag{19}$$

where $x(t) = x_1(t)$.

The observatoins are given by

$$\begin{cases} z_0^* = x_1(0) + \eta_0 \\ z_T^* = x_1(10) + \eta_t \\ \dot{z} = x_1 + \dot{\eta} \\ \dot{y}_1 = u_1 + u_2 + \dot{\omega}_1 \\ \dot{y}_2 = u_2 - u_2 + \dot{\omega}_2 \end{cases} \tag{20}$$

where $\{\eta\}$ and $\{(\omega_1, \omega_2)^T\}$ are the Wiener processes with differential covariances $cov(\eta, \eta) = 900$, $cov(\omega_1, \omega_1) = \mu$, $cov(\omega_2, \omega_2) = \mu, cov(\omega_1, \omega_2) = \mu/2$ respectively. The known parameter μ characterizes the accuracy of the observations; $cov(\eta_0, \eta_0) = cov(\eta_T, \eta_T) = 100000$.

The filtering problem is to calculate recursively the J_t-optimal estimate of the state x given the observations (20). We use the algorithm (8),(9). Table 1 gives the estimation results: $\sigma_1, \sigma_2, \sigma_3$ are the estimation error standard deviations, obtained by $\mu = 100$, $\mu = 1000$, $\mu = 5000$ respectively, σ_K is the estimation error standard deviation obtained by the Kalman filter in the ideal situation when complete information about u_1, u_2 is available. From Table 1, it follows that the estimates given by the algorithm (8),(9) are close enough to those of the Kalman filter and may be used in the case of unknown but observable input signals when the Kalman filter is useless.

time	σ_K	σ_1	σ_2	σ_3
0	316.23	316.23	316.23	316.23
0.2	74.02	74.03	74.15	74.66
1	66.37	66.40	66.63	67.67
2	54.51	54.60	55.45	58.61
3	51.93	52.10	53.48	57.78
5	51.68	51.87	53.37	57.77
6	51.68	51.87	53.37	57.77

Table 1. Estimation error standard deviations $(\sigma_1, \sigma_2, \sigma_3)$ for the filter (8),(9) and corresponding values σ_K for the Kalman filter.

2. Let us consider the system of observations (18)-(20) in the case $C \equiv 0$. The fixed-interval smoothing problem is to calculate $\tilde{J}_t$-optimal estimate for all $t \in [0, 10]$. We use the algorithm given by (17). Table 2 gives the estimation results: $\sigma_1^f, \sigma_2^f, \sigma_3^f$ are the filter estimation error standard deviations obtained by $\mu = 100, \mu = 1000, \mu = 5000$ and $\sigma_1^s, \sigma_2^s, \sigma_3^s$ are the smoothing estimation error standard deviations respectively. From the Table 2 it follows that the smoothing

estimate is significantly more accurate than the corresponding filtering estimate.

time	σ_1^f	σ_1^s	σ_2^f	σ_2^s	σ_3^f	σ_3^s
0	316.23	12.13	316.23	24.93	316.23	36.57
0.2	74.01	11.98	74.13	23.71	74.64	33.02
1	66.17	11.95	66.42	23.14	67.46	32.32
2	53.01	11.89	54.00	22.70	57.57	31.02
3	47.98	11.98	50.28	22.40	56.23	30.80
5	44.75	13.53	49.26	22.32	56.18	30.80
7	44.19	19.82	49.26	23.59	56.18	30.84
8	44.15	25.46	49.26	27.35	56.18	31.43
9	44.13	32.55	49.26	34.57	56.18	37.38
9.8	44.13	37.41	49.26	40.39	56.18	44.01
10.0	44.13	43.71	49.26	48.67	56.18	56.31

Table 2. Standard deviations of the filtering estimate error $(\sigma_1^f, \sigma_2^f, \sigma_3^f)$ and the corresponding values $\sigma_1^s, \sigma_2^s, \sigma_3^s$ for the smoothing.

6 Appendix A.

Proof of the Proposition 1.
An arbitrary linear estimate $\hat{x}_t$ is given by

$$\hat{x}_t = \alpha(t)y_0 + \int_0^t \beta(t,\tau)dy_\tau + \int_0^t \gamma(t,\tau)dz_\tau \qquad (21)$$

where the optimal functions α, β, γ must be calculated. The error $\Delta_t = x_t - \hat{x}_t$ can be decomposed as $\Delta_t = m_t + \Delta_t^\circ$, where $m_t = \mathbf{E}\{\mathbf{x_t} - \hat{\mathbf{x}}_\mathbf{t}\}$ is the bias of the estimate , $\Delta_t^\circ = \Delta_t - m_t$ is a random error. Then $J_t = J_t^1 + J_t^2$, where

$$J_t^1 = \sup_{\nu \in R^p, \{u_\tau\} \in \mathrm{U_t}} [m_t^T \Sigma_t m_t], \quad J_t^2 = \mathbf{E}[\mathbf{\Delta_t^{\circ T} \Sigma_t \Delta_t^\circ}],$$

$$m_t = [\Phi(t,0) - \alpha(t)\phi_0 - \textstyle\int_0^t \gamma(t,\tau)\psi(\tau)\Phi(\tau,0)d\tau]\nu + \qquad (22)$$

$$\int_0^T [\Phi(t,\tau)b(\tau) - \beta(t,\tau)\phi(\tau) - \int_0^t \gamma(t,\sigma)\psi(\sigma)\Phi(\sigma,\tau)d\sigma b(\tau)]u_\tau d\tau,$$

where $\Phi(t,\tau)$ is a solution of the matrix differential equation

$$\begin{cases} \dot{\Phi}(t,\tau) = a(t)\Phi(t,\tau), \\ \Phi(\tau,\tau) = I \end{cases}$$

Since $\nu \in R^p$ and $\{u_t\} \in \mathbf{U}_t$

$$J_t^1 = \begin{cases} 0, & \text{if} \quad \Sigma_t[\Phi(t,0) - \alpha(t)\phi_0 - \int_0^t \gamma(t,\tau)\psi(\tau)\Phi(\tau,0)d\tau] = 0 \\ & \text{and} \quad \Sigma_t[\Phi(t,\tau)b(\tau) - \beta(t,\tau)\phi(\tau) - \int_0^t \gamma(t,\sigma)\psi(\sigma)\Phi(\sigma,\tau)d\sigma b(\tau)] = 0, \\ \infty, & \text{otherwise} \end{cases}$$

By using the matrix Schwarz inequality (Fomin, 1984) it may be shown that the optimal $\alpha(t)$, $\beta(t,\tau)$ are

$$\begin{cases} \alpha(t) = [\Phi(t,0) - \int_0^t \gamma(t,\tau)\psi(\tau)\Phi(\tau,0)d\tau]\phi_0^+, \\ \beta(t,\tau) = [\Phi(t,\tau) - \int_0^t \gamma(t,\sigma)\psi(\sigma)\Phi(\sigma,\tau)d\sigma]b(\tau)\phi^+(\tau) \end{cases} \qquad (23)$$

where $\phi_0^+ = (\phi_0^T Q_0^{-1}\phi_0)^{-1}\phi_0^T Q_0^{-1}$, $\phi^+(\tau) = (\phi^T(\tau)Q^{-1}(\tau)\phi(\tau))^{-1}\phi^T(\tau)Q^{-1}(\tau)$ are pseudoinverse of ϕ_0 and $\phi(\tau)$. From (22) $\hat{x}_t$ is unbiased and hence $\Delta_t = \Delta_t^\circ$. Hence

$$\hat{x}_t = \Phi(t,0)\phi_0^+ y_0 + \int_0^t \Phi(t,\tau)b(\tau)\phi^+(\tau)dy_\tau + \int_0^t \gamma(t,\tau)\{dz_\tau - \psi(\tau)[\Phi(\tau,0)\phi_0^+ y_0 + \int_0^\tau \Phi(\tau,\sigma)b(\sigma)\phi^+(\sigma)dy_\sigma]d\tau\} \qquad (24)$$

Then the error Δ_t satisfies the equation

$$\Delta_t = \int_0^t \gamma(t,\tau)dg_\tau - \lambda_t,$$

where $\{\lambda_t\}, d_t$ are defined by

$$d\lambda_t = a(t)\lambda_t dt + b(t)\phi^+(t)d\omega_t + d\xi_t, \quad \lambda_0 = \phi_0^+\omega_0 \qquad (25)$$

$$dg_t = \psi(t)\lambda_t dt - d\eta_t, \quad g_0 = 0 \qquad (26)$$

The problem is to find a weight function $\gamma(t,\tau)$ that minimizes J_t or equivalently to calculate the J_t^2-optimal estimate $\hat{\lambda}_t$ of λ_t given observations $\{g_\tau\}$.The solution of this problem is given by the Kalman filter

$$\begin{cases} d\hat{\lambda}_t = a(t)\hat{\lambda}_t dt + k(t)\psi^T(t)P^{-1}(t)[dg_t - \psi(t)\hat{\lambda}_t dt], \\ \dot{k}(t) = a(t)k(t) + k(t)a^T(t) - k(t)\psi^T(t)P^{-1}(t)\psi(t)k(t) + \\ b(t)[\phi^T(t)Q^{-1}(t)\phi(t)]^{-1}b^T(t) + C(t), \\ \lambda_0 = 0, \quad k(0) = (\phi_0^T Q_0^{-1}\phi_0)^{-1}. \end{cases} \tag{27}$$

Here $k(t) = cov(\lambda_t - \hat{\lambda}_t, \lambda_t - \hat{\lambda}_t) = cov(x_t - \hat{x}_t, x_t - \hat{x}_t)$.

$$\hat{\lambda}_t = \int_0^t \Theta(t,\tau)k(\tau)\psi^T(\tau)P^{-1}(\tau)dg_\tau \tag{28}$$

where $\Theta(t,\tau)$ is given by

$$\begin{cases} \dot{\Theta}(t,\tau) = [a(t) - k(t)\psi^T(t)P^{-1}(t)\psi(t)]\Theta(t,\tau), \quad t > \tau \\ \Theta(\tau,\tau) = I \end{cases} \tag{29}$$

From (28) it follows that

$$\gamma(t,\tau) = \Theta(t,\tau)k(\tau)\psi^T(\tau)P^{-1}(\tau) \tag{30}$$

Now we can obtain the first equation (8) by substituting (30) into (24) and differentiating both sides of (24) with respect to t.

7 Appendix B.

Proof of Proposition 3.

First we state optimal estimate $\hat{x}_t^s$ to be a linear combination of $\hat{x}_t$ and $\hat{x}_t^r$

$$\hat{x}_t^s = A(t)\hat{x}_t + B(t)\hat{x}_t^r \tag{31}$$

Then from the Gauss-Markov theorem we have

$$\begin{cases} A(t) = k^s(t)k^{-1}(t), \\ B(t) = k^s(t)[k^r(t)]^{-1}, \\ k^s(t) = k^{-1}(t) + [k^r(t)]^{-1}]^{-1}. \end{cases} \tag{32}$$

As it follows from the proof of Proposition 1, $\hat{x}_t^s$ is given by

$$\hat{x}_t = \epsilon(t)\{\Phi(t,0)\phi_0^+ y_0 + \int_0^t \Phi(t,\tau)b(\tau)\phi^+(\tau)dy_\tau\}+$$
$$\int_0^t \gamma(t,\tau)\{dz_\tau - \psi(\tau)[\Phi(\tau,0)\phi_0^+ y_0 + \int_0^\tau \Phi(\tau,\sigma)b(\sigma)\phi^+(\sigma)dy_\sigma]d\tau\}+$$
$$(I-\epsilon(t))\{\Phi^r(t,T)\phi_T^+ y_T + \int_T^t \Phi^r(t,\tau)b(\tau)\phi^+(\tau)dy_\tau+$$
$$\int_T^t \gamma^r(t,\tau)\{dz_\tau - \psi(\tau)[\Phi^r(\tau,T)\phi_T^+ y_T + \int_T^\tau \Phi^r(\tau,\sigma)b(\sigma)\phi^+(\sigma)dy_\sigma]d\tau\} \quad (33)$$

where the optimal coefficients $\gamma(t,\tau), \gamma^r(t,\tau), \epsilon(t)$ must be calculated. So $\mathbf{E}\{\mathbf{\Delta}_t^s\} = \mathbf{E}\{\mathbf{x}_t - \hat{\mathbf{x}}_t^s\} = \mathbf{0}$ i.e. $\hat{x}_t^s$ is an unbiased estimate of x_t . Let us define the stochastic processes $\lambda_t, \lambda_t^r, g_t, g_t^r$ by

$$\begin{cases} d\lambda_t = a(t)\lambda_t dt + b(t)\phi^+(t)d\omega_t, \\ dg_t = \psi(t)\lambda_t dt - d\eta_t, \\ \lambda_0 = \phi_0^+\omega_0, \quad g_0 = 0 \end{cases} \quad (34)$$

$$\begin{cases} -d\lambda_t^r = -a(t)\lambda_t^r dt - b(t)\phi^+(t)d\omega_t, \\ dg_t^r = \psi(t)\lambda_t^r dt - d\eta_t, \\ \lambda_T^r = \phi_T^+\omega_T, \quad g_T^r = 0 \end{cases} \quad (35)$$

The smoothing estimate error Δ_t^s can be decomposed as
$\Delta_t^s = \int_0^t \gamma(t,\tau)dg_\tau - \epsilon(t)\lambda_t + \int_T^t \gamma^r(t,\tau)dg_\tau^r - [1-\epsilon(t)]\lambda_t^r,$
The problem is to find weight functions $\gamma(t,\tau), \gamma^r(t,\tau)$ that minimize $\tilde{J}_t$ or equivalently to calculate the optimal estimate of the linear combination $\epsilon(t)\lambda_t + [1-\epsilon(t)]\lambda_t^r$ given observations $\{g_\tau, 0 \le \tau \le t\}$ and $\{g_\tau^r, t \le \tau \le T\}$. The solution of this problem is given by the Kalman filter for the system (34),(35). As in the proof of Proposition 1 we have

$$\begin{cases} \gamma(t,\tau) = \epsilon(t)\Theta(t,\tau)k(\tau)\psi^T(\tau)P^{-1}(\tau), \\ \gamma^r(t,\tau) = [I-\epsilon(t)]\Theta^r(t,\tau)k^r(\tau)\psi^T(\tau)P^{-1}(\tau) \end{cases} \quad (36)$$

where $\Theta(t,\tau)$ and $\Theta^r(t,\tau)$ are given by

$$\begin{cases} \dot{\Theta}(t,\tau) = [a(t) - k(t)\psi^T(t)P^{-1}(t)\psi(t)]\Theta(t,\tau), \ t > \tau, \ \Theta(\tau,\tau) = I, \\ \dot{\Theta}^r(t,\tau) = [-a(t) - k^r(t)\psi^T(t)P^{-1}(t)\psi(t)]\Theta^r(t,\tau), \ t < \tau, \ \Theta^r(\tau,\tau) = I. \end{cases}$$

Now we can substitute (36) into (33) and obtain (31) in the form
$\hat{x}_t^s = \epsilon(t)\hat{x}_t + [I-\epsilon(t)]\hat{x}_t^r.$

References

Barton, R.J., and Poor, H.V., 1990 Robust estimation and signal detection . *IEEE Transactions on Information Theory*,**36**,485-501.

Borisov, A.V., Pankov, A.R., and Sotsky, N.M.,1991 Filtering and smoothing in uncertain-stochastic systems with partially observable input signals. *Automation and Remote Control*, **3**,85-95 (in Russian).

Borisov, A.V.,and Pankov, A.R., 1991 Optimal signal processing for uncertain-stochastic systems. *Proceedings of 30th Conference of Decision and Control.*

Brocett, R.W, and Clark, J.M.C., 1980 The geometry of the conditional density functions. *Analysis and Optimization of Stochastic Systems*, O.L.R. Jacobs et al., eds, Academic Press, 299-309.

Eweda,E.,1991 Convergence of the sign algorithm for adaptive filtering with correlated data. *IEEE Transactions on Information Theory*,**37**, 1450-1457.

Fomin, V.N., 1984 Recursive estimation and adaptive filtering. Moscow, Nauka (in Russian).

Goodwin , G.C.,and Sin, K.S., 1984 Adaptive filtering, prediction and control. New Jersey, Prentice Hall Inc.

Hero, A.O., 1991 Timing estimation for filtered Poisson processes in Gaussian noise.*IEEE Transactions on Information Theory*,**37**, 92-106.

Kalmàn R., and Bucy, R.S., 1961 New results in linear filtering and prediction theory. *Transactions of ASME*,**83**, 95-108.

Kassam, S.A., and Poor, H.V., 1985 Robust techniques for signal processing: a survey . *Proceedings of IEEE*,**73**,433-481.

Kleptsina, M.L., and Veretennikov, A.Yu., 1985 On filtering and properties of conditional laws of Ito-Volterra processes. *Statistics and Control of Stochastic Processes, Steklov Seminar*, N.V. Krylov et al.,eds., Optimization Software Inc.

Kurzhansky, A.B., 1977 Control and observation under uncertainty. Moscow, Nauka (in Russian).

Liptser, R.S., and Shiryayev, A.N., 1977 Statistics of random processes. New York, Springer-Verlag.

Malyshev, V.V., and Kibzun, A.I., 1987 Analysis and sinthysis of the aircraft high accuracy control. Moscow, Mashinostroenie (in Russian).

Mayne, D.Q., 1966 A solution of the smoothing problem for the linear dynamic systems. *Automatica*,**4**,73-92.

Nagpal, K.M., and Khargonekar, P.P., 1991 Filtering and smoothing in an H^{∞} Setting. *IEEE Transactions on Automatic Control*,**36**,152-166.

Pardoux, E., 1979 Filtering of a diffusion processes with Poisson-type observation. *Stochastic Control Theory and Stochastic Differential Systems*, M. Kohlman and W. Vogel eds., *Lecture Notes in Control and Information Sciences*,**16**, New York, Springer-Verlag.

Tam, L.-M., Wong, W.S., and Yau, S.S.-T., 1990 On necessary and sufficient condition for finite dimensionality of estimation algebras.*SIAM Journal of Control and Optimization*,**28**, 173-185.

Wall, J.E., Willsky, A.S., and Sandell, N.R., 1981 On the fixed-interval smoothing problem. *Stochastics*,**5**,1-42.