FINITE-DIMENSIONAL ALGORITHMS OF NONLINEAR SYSTEM STATE ESTIMATION

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

ALEXEI V. BOSOV
Department of Applied Mathematics
Moscow Aviation Institute
Moscow,127080, Russia

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

Technical Report 92-13

January, 1992.

FINITE-DIMENSIONAL ALGORITHMS OF NONLINEAR SYSTEM STATE ESTIMATION

ALEXEI R. PANKOV, [1] ALEXEI V. BOSOV, [2] ANDREI V. BORISOV[3]

(Revised on January 28, 1992)

Absract. In this paper we consider the method of conditionally-minimax nonlinear filtering (CMNF) of processes in nonlinear stochastic dynamic discrete-time systems. In order to compare CMNF with the optimal nonlinear filter (ONF), we also derive a special numerical algorithm which is based on the theory of spline approximation. This algorithm allows us to calculate the best mean-square estimate of the system state-vector. The results of numerical experiments with CMNF, ONF and the extended Kalman filter (EKF) show the favorable properties of the derived CMNF algorithm.

Key Words. Optimal nonlinear filtering, conditionally-minimax filtering, extended Kalman filter, spline approximation.

1 Introduction.

The problem of nonlinear state estimation is a very important part of the general control problem for systems with incomplete data. It is well-known that the problem of nonlinear filtering may be reformulated as the problem of solving the stochastic functional equations for the conditional distribution of the system state given all observations (Liptser, Shiryayev,1977) , (Davis,Marcus, 1981). So the estimation problem in the general case is infinite-dimensional. Its complete solution is practically impossible.

Recently various authors tried to determine nonlinear system properties that provide provide the finite-dimensionality of the optimal estimator (Sawitzki, 1981) , (Benes, 1981), (Wong, 1983) , (Daum, 1986) , (Tam, Wong and Yau, 1990). The investigations show that a finite-dimensional optimal nonlinear filter is not typical, and even in very simple cases may not take place (Hazewinkel,Markus and Sussan, 1983). That is why we support the idea of obtaining the finite-dimensional estimators by means of appropriate approximations of the optimal nonlinear infinite-dimensional estimation algorithms. Some of these approximations are well-known and widely

[1]Department of Applied Mathematics, Moscow Aviation Institute, Volokolamskoye shosse, 4, Moscow, 127080, Russia

[2]Department of Applied Mathematics, Moscow Aviation Institute, Volokolamskoye shosse, 4, Moscow, 127080, Russia

[3]Department of Industrial & Operations Engineering, University of Michigan, Ann Arbor MI 48109-2117

used: the linearized and the extended Kalman filters, the second order filter and others (Sage,Melsa, 1972). Their practical utilization shows that they often obtain very unpleasant properties, i.e. give a seriously biased estimate which usually diverges. So it is important to continue efforts to obtain the finite-dimensional nonlinear filters which provide an estimate with improved properties.

In this paper we consider two different approaches for finite-dimensional nonlinear filtering. The first approach is based on the idea of a local conditional optimization of the filter structure given the class of admissible filters. This idea was first realized by V.S. Pugachev (Pugachev, 1979) , (Pugachev and Sinitsyn, 1990), who derived the conditionally-optimal nonlinear filter. To utilize this method one needs to know the joint characteristic function of the system state and the estimate vectors. This function may be obtained by solving the special functional equation, which is nonrandom but rather complicated. It takes substantional effort to obtain the desired solution (Raol,Sinha 1987). Using the idea of the conditionally-optimal filter we derive the conditionally-minimax filter (CMNF) which may be determined without information about the above-mentioned characteristic function by some a priori information (Pankov, 1990).

The second method is based on spline approximation of the conditional system state density given an observation process. We call this method the optimal nonlinear filter (ONF) because it may be shown that the ONF provides the optimal (in a mean square sense) estimate approximation with the desired accuracy.

In order to test properties of the obtained algorithms we compare the CMNF and the ONF with the extended Kalman filter (EKF). The corresponding computer simulation results are also considered in this paper.

2 CMNF algorithm.

We shall use the following notations: $\mathbf{E}\{x\}$ - the mathematical expectation of x ; $cov\{x,y\}$ - the covariance of x and y ; $P(m,S)$ - the set of random vectors with $\mathbf{E}\{x\} = m$, $cov\{x,x\} \leq S$; A^+ - the pseudoinverse matrix with respect to A ; $y = col(x_1,...,x_n)$ if y is a vector-column of the form $y = col(x_1^t,...,x_n^T)^T$; $||x|| = (x^TWx)^{1/2}$ for some weight matrix W : $W = W^T$, $W \geq 0$.

Let us consider the following discrete-time stochastic dynamic observation model:

$$\begin{cases} y_n = \phi(y_{n-1}, w_n), \quad n = 1,2,...; \quad y_0 = \eta, \\ z_n = \psi(y_n, v_n) \end{cases} \quad (1)$$

where $y_n \in R^p$ is a state vector; $w_n \in R^q$ is a vector of random disturbances; $\eta \in R^p$ is a vector of initial conditions; $z_n \in R^m$ is a vector of observations ; $v_n \in R^r$ is a vector of random observation errors, and $\phi_n(y,w)$ and $\psi_n(y,v)$ are known nonlinear functions.

Let $Z^n = col(z_n,...,z_1)$, i.e. Z^n is a vector of all observations up to the moment n.

Consider the structure of the CMNF for the processes $\{y_n\}$ given Z^n. Let $\xi_n(y)$ and $\pi_n(y,z)$ be some fixed nonlinear vector functions (the choice of these functions will be considered in a sequel) and $\hat{y}_{n-1}$ be the CMNF-estimate of y_{n-1} given Z^{n-1}. Let $\bar{y}_n = \xi_n(\hat{y}_{n-1})$. Then the conditionally-minimax prediction $\tilde{y}_n$ for y_n takes the form:

$$\tilde{y}_n = \phi^*(\bar{y}_n), \tag{2}$$

where

$$\phi_n^*(\bar{y}_n) = arg \min_{\phi\in\Phi} \max_{P(m_1,S_1)} \mathbf{E}\{||y_n - \phi(\bar{y}_n)||^2\} \tag{3}$$

and $x_1 = col(y_n, \bar{y}_n) \in P(m_1, S_1)$, Φ is a set of measurable functions, $\mathbf{E}||\phi(\bar{y}_n)||^2 < \infty$. The estimate $\hat{y}_n$ is a conditionally-minimax correction of $\tilde{y}_n$ given the new observation z_n:

$$\hat{y}_n = \tilde{y}_n + \psi^*(\pi_n(\tilde{y}_n, z_n)), \tag{4}$$

where

$$\psi_n^*(\nu_n) = arg \min_{\psi\in\Phi} \max_{P(m_2,S_2)} \mathbf{E}\{||y_n - \tilde{y}_n - \psi(\nu_n)||^2\} \tag{5}$$

and $x_2 = col(y_n - \tilde{y}_n), \nu_n) \in P(m_2, S_2)$, $\nu_n = \pi_n(\tilde{y}_n, z_n)$.
If we solve (3),(5) and find the functions $\phi_n^*(.)$ and $\psi_n^*(.)$ then equations (2),(4) will define the recursive CMNF. It may be shown that under some rather general conditions there exists a solution of (3),(5) where the corresponding optimal functions $\phi_n^*(.)$ and $\psi_n^*(.)$ are linear.
Consider the following conditions for the model (1), $\xi_n(y)$ and $\pi_n(y,z)$.
1) There exists $C_n^{\phi}(w) < \infty$, $w \in R^q$ such, that
$||\phi(y,w)||^2 \le C_n^{\phi}(w)(1 + ||y||^2)$ and $\mathbf{E}\{C_n^{\phi}(w_n)\} < \infty$, $y \in R^p$;
2) There exists $C_n^{\xi} < \infty$, $||\xi_n(y)||^2 \le C_n^{\xi}(1 + ||y||^2)$, $y \in R^p$;
3.a) There exists $C_n^{\pi} < \infty$: for all $y \in R^p$, $z \in R^m$
$||\pi_n(y,z)||^2 \le C_n^{\pi}(1 + ||y||^2)$;
3.b) The processes $\{v_n\}$ is independent of $\{w_n\}$ and η:
There exists $C_n^{\pi} < \infty$: $||\pi_n(y,z)||^2 \le C_n^{\pi}(1 + ||y||^2 + ||z||^2)$;
There exists $C_n^{\psi}(v) < \infty$: $||\psi_n(y,v)||^2 \le C_n^{\psi}(1 + ||y||^2)$ and $\mathbf{E}\{C_n^{\psi}(v_n)\} < \infty$

Proposition 1.*Let 1),2) and 3.a) or 3.b) hold. Let also* $\mathbf{E}\{||\eta||^2\} < \infty$ *and* $\bar{y}_0 = \mathbf{E}\{\eta\}$, *then*

$$\phi_n^*(y) = F_n y + f_n; \qquad \psi_n^*(\nu) = H_n \nu + h_n, \tag{6}$$

where

$$\begin{cases} F_n = cov(y_n, \bar{y}_n) cov^+(\bar{y}_n, \bar{y}_n); & f_n = \mathbf{E}\{y_n - F_n \bar{y}_n\}; \\ H_n = \mathbf{E}\{y_n - \tilde{y}_n \nu_n^T\} cov^+(\nu_n, \nu_n); & h_n = -H_n \mathbf{E}\{\nu_n\}; \end{cases} \tag{7}$$

The estimate $\hat{y}_n$ is unbiased and has the error covariance matrix

$$\hat{K}_n = K_n - F_n cov(\bar{y}_n, y_n) - H_n cov(\nu_n, y_n - \tilde{y}_n), \tag{8}$$

where $\hat{K}_n = cov(y_n - \hat{y}_n, y_n - \hat{y}_n)$, and $K_n = cov(y_n, y_n)$ is the covariance matrix of the process $\{y_n\}$.

Proof of Proposition 1: see Appendix 1.

From Proposition 1 it follows that CMNF is defined by (2),(3),(6)-(8).
If we consider a dynamic system which is not general as (1), then restrictions 1)-3) may be weakened. Let the dynamic observation system be described by the discrete-time diffusion equation

$$\begin{cases} y_n = a_n(y_{n-1}) + b(y_{n-1})w_n, \quad n = 1, 2, ...; \quad y_0 = \eta; \\ z_n = \psi_n(y_n) + v_n, \end{cases} \tag{9}$$

where $\{w_n\}$, $\{v_n\}$ are independent gaussian white noises; η is a gaussian vector independent of $\{w_n\}$, $\{v_n\}$. Model (9) is widely used in applications (Sage,Melsa 1972).

Proposition 2.*Let $\{y_n, z_n\}$ be described by (9) and:*
1)there exists $C_n^1, \alpha_n < \infty$: $||a_n(y)|| + ||b_n(y)|| \leq C_n^1(1 + ||y||^{\alpha_n})$;
2)there exists $C_n^2 < \infty$: $||\xi_n(y)|| \leq C_n^2(1 + ||y||^{\alpha_n})$;
$||\psi_n(y)|| \leq C_n^2(1 + ||y||^{\alpha_n})$
$||\pi_n(y, z)|| \leq C_n^2(1 + ||y||^{\alpha_n} + ||z||^{\alpha_n})$;
3)$\hat{y}_0 = \mathbf{E}\{\eta\}$.
Then the result of Proposition 1 holds.

Proof of Proposition 2: see Appendix B.

It should be mentioned that the conditions of Proposition 2 allow the polynomial growth of all functions. However in Proposition 1 the growth is only linear.
Consider some obvious types of the structural functions $\xi_n(y)$ and $\pi_n(y, z)$.
a) $\xi_n(\hat{y}_{n-1}) = \hat{y}_{n-1}$ -linear prediction;
b) $\xi_n(\hat{y}_{n-1}) = \phi_n(\hat{y}_{n-1}, \mathbf{E}\{w_n\})$-prediction via dynamic system equations (1);
c) $\pi_n(\tilde{y}_n, z_n) = z_n - \psi_n(\tilde{y}_n, v_n^0)$ - residual, where $v_n^0 = \mathbf{E}\{v_n\}$;
d) $\pi_n(\tilde{y}_n, z_n) = \tilde{K}_n \Phi_n^T(\Phi_n \tilde{K}_n \Phi_n^T + R_n)^+(z_n - \psi(\tilde{y}_n, v_n^0)$,
where $\tilde{K}_n = cov(y_n - \tilde{y}_n, y_n - \tilde{y}_n)$, $\Phi_n = \partial\psi(\tilde{y}_n, v_n^0)/\partial\tilde{y}_n$,
$R_n = [\partial\psi(\tilde{y}_n, v_n^0)/\partial v_n^0]cov(v_n.v_n)[\partial\psi(\tilde{y}_n, v_n^0)/\partial v_n^0]$- transformed residual. $\pi_n(\tilde{y}_n, z_n)$ in the given above form is used in the EKF as a correction term. The only difference is that in the EKF instead of $\tilde{K}_n$ is used some random approximation (given by the Riccati-type equation) of the real prediction error covariance matrix $\tilde{K}_n$.
There are some other types of $\xi_n(y)$ and $\pi_n(y, z)$. In the general case we choose

$\xi_n(\hat{y}_{n-1})$ and $\pi_n(\tilde{y}_n, z_n)$ to approximate (may be in the minimax sense) the conditional expectations $\mathbf{E}\{y_n|\hat{y}_{n-1}\}$ and $\mathbf{E}\{y_n - \tilde{y}_n|\tilde{y}_n, z_n\}$ respectively.
All unknown parameters (7) of the CMNF algorithm may be determined a priori by a computer simulation. The corresponding algorithm for this is considered in detail in (Pankov, 1990).

3 ONF spline algorithm.

It is well-known that if $\{y_n\}$ is the second order process, then the conditional expectation $y_n^* = \mathbf{E}\{y_n|Z^n\}$ minimizes the mean-square criterion $J_n = \mathbf{E}\{||y_n - y_n^*||^2\}$ (Davis,Marcus 1981). The estimate y_n^* may be calculated recursively if $\{y_n\}$ is the Markov process. It may be shown that if $\eta, w_1, w_2, ...$ are independent random vectors, $\mathbf{E}\{||\eta||^2\} < \infty$ and $\mathbf{E}\{||\phi_n(y, w_n)||^2\} \leq c_n(1 + ||y||^2) < \infty$, $n \geq 1$, then the process $\{y_n\}$ has finite second-order moments and obtain the Markov property.
We assume existence of the following probability characteristics: the conditional density $g_{n-1}(y)$ of y_{n-1} given Z^{n-1} ; the transition density $\mu(n-1, x; n, y)$ of the process $\{y_n\}$; the conditional density $\rho_n(y, z)$ of z_n given $y_n = y$. Denote $\omega_n(y)$ the unnormalized conditional density of y_n given Z^n (Davis,Marcus 1981). then as in (Takeuchi,Akashi 1981)

$$\left\{ \begin{array}{l} \omega_n(y) = \rho_n(z_n, y) \int_{R^p} \mu(n-1, x; n, y)\omega_{n-1}(x)dx, \\ \omega_0(y) = p_\eta(y), \end{array} \right. \quad (10)$$

where $p_\eta(y)$ is the η probability density. Conditional densities $g_n(y)$ and $\omega_n(y)$ are connected by the formula

$$g_n(y) = \omega_n(y)[\int_{R^p} \omega_n(x)dx]^{-1} \quad (11)$$

From (11) it follows that y_n^* is obtained by

$$y_n^* = \int_{R^p} x\omega_n(x)dx[\int_{R^p} \omega_n(x)dx]^{-1} \quad (12)$$

We are going to use $\omega_n(y)$ spline approximation for the numerical treatment of (10),(12). The spline approximation methods are widely used for solving the statistical estimation problems (Andrade Netto,Gimeno and Mendes 1978) , (Wegman,Wright 1983) , (Villalobos,Wahba 1987).
We shall consider the case $y_n \in R^1$ for simplicity. Let us consider the sequence of partitions $\Delta_N^{(n)} = \{x_0^{(n)}, .., x_N^{(n)}\}$ of the intervals $[a^{(n)}, b^{(n)}]$: $x_k^{(n)} = x_{k-1}^{(n)} + h_k^{(n)}$, $x_0^{(n)} = a^{(n)}$, $x_N^{(n)} = b^{(n)}$. $\Delta_N^{(n)}$ may be extended in the following way: $x_{-i}^{(n)} = x_0^{(n)} - h_1^{(n)} i$, $x_{N+1}^{(n)} = x_N^{(n)} + h_N^{(n)} i$, $i = 1, 2, ...$.
To approximate $\omega_{n-1}(y)$, $\omega_n(y)$ we shall use polynomial splines of degree $2m-1$. The

general representation of the spline is defined in (Grebennikov, 1983) by the formula

$$S_{2m,p}(x) = \sum_{l=-p}^{N+p}[\sum_{j=-p}^{p} a_{ij} f_{l+j}] s_{l-m,2m}(x) = \sum_{l=-p}^{N+p} C_l s_{l-m,2m}(x), \quad x \in [a,b], \tag{13}$$

where $\{a_{ij}\}$ are unknown paremeters; $f_k = f(x_k)$: $f(x)$ is functional to be approximated on $[a,b]$; P is an arbitrary integer number from the set $\{0,1,...,[(N-1)/2]\}$; $\{s_{i,2m}(x)\}$ is a system of B-splines.
The recursive procedure of the numerically stable determination of B-splines at the point $x \in [a,b]$ is defined by the formula (Cox 1978):
$s_{i,l}(x) = [(x-x_i)/(x_{i+l-1}-x_i)]s_{i,l-1}(x) + [(x_{i+1}-x)/(x_{i+l}-x_{i+1})]s_{i+1,l-1}(x),$

$$\begin{cases} s_{i,1}(x) = 1, & if \quad x \in [x_i, x_{i+1}], \quad i = 2,...,2m \\ s_{i,1}(x) = 0, & otherwise \end{cases}$$

Let us describe the procedure of y_n^* determination given the spline-approximation of $\omega_{n-1}(y)$ on $\Delta_N^{(n-1)}$:

$$\hat{\omega}_{n-1}(y) = S_{2m,p}^{(n-1)}(y) = \sum_{l=-p}^{N+p} C_l^{n-1} s_{l-m,2m}^{n-1}(y), \tag{14}$$

where $\{s_{i,2m}^{n-1}\}$ is a system of B-splines on $\Delta_N^{(n-1)}$.
Let us introduce the following notations

$$\begin{cases} \eta_{k,l}^{(n)} = \int_{-\infty}^{+\infty} \mu(n-1,x;n,x_k^{(n)}) s_{l-m,2m}^{(n-1)}(x)dx; \\ \theta_l^{(n)} = \int_{-\infty}^{+\infty} x s_{l-m,2m}^{(n-1)}(x)dx; \\ \kappa_l^{(n)} = \int_{-\infty}^{+\infty} s_{l-m,2m}^{(n-1)}(x)dx. \end{cases} \tag{15}$$

Then from (11) , (12) , (14) , (15) it follows that, at the knots $\{x_k^{(n)}\}$ and $\Delta_N^{(n)}$, the following conditions are fulfiled:

$$\hat{\omega}_n(x_k^{(n)}) = \rho_n(z_n, x_k^{(n)}) \sum_{l=-p}^{N+P} C_l^{(n-1)} \eta_{k,l}^{(n)}. \tag{16}$$

Using the values of $\hat{\omega}_n(x_k^{(n)})$given by (16), we can define unknown parameters $\{C_l^{(n)}\}$ of the spline $S_{2m,p}^{(n)}(y)$, which is an approximation of the $\omega_n(y)$. The corresponding estimate $\hat{y}_n^*$ of the conditional expectation y_n^* is given by

$\hat{y}_n^* = \sum_{l=-p}^{N+P} C_l^{(n)} \theta_l^{(n)} [\sum_{l=-p}^{N+P} C_l^{(n)} \kappa_l^{(n)}]^{-1},$

To obtain $\{C_l^{(n)}\}$ we shall use a special method, which was called the method of direct approximation (Grebennikov, 1983). At the points $\{x_k^{(n)}\}$ we use the approximation conditions of the following form

$$S_{2m,p}^{(n)}(x_k^{(n)}) = \hat{\omega}_n(x_k^{(n)}) + \sum_{l=1}^{N+P-1} d_{kl}^{(n)}\delta^{2l}[\hat{\omega}_n(x_k^{(n)})], \qquad (17)$$

where $\delta^2[.]$ is a central divided difference of degree $2l$ of the function $\hat{\omega}(x)$ at the knot $x_k^{(n)}$. The coefficients $\{d_{kl}^{(n)}\}$ together with $\{C_l^{(n)}\}$ (see formula (13)) can be found from (17) (Grebennikov, 1983).

It should be mentioned that this method of spline-approximation differs from the standard spline-interpolation method (Ahlberg,Nilson and Walsh, 1967). The coefficients $\{C_l^{(n)}\}$ satsfy some system of linear equations, which may be solved analytically in some particular but important cases. For example, for the cubic spline defined on the uniform partition $\Delta_N^{(n)}$ (i.e. $h_k^{(n)} = h^{(n)}, k = 1, ..., N$), it may be shown that for $l = -P, ..., N+P$

$C_l^{(n)} = -1/6\hat{\omega}_{l-1} + 4/3\hat{\omega}_l - 1/6\hat{\omega}_{l+1}$ if $P = 1$;

$C_l^{(n)} = 1/36(\hat{\omega}_{l-2} + \hat{\omega}_{l+2}) - 5/18(\hat{\omega}_{l-1} + \hat{\omega}_{l+1}) + 3/2\hat{\omega}_l$ if $P = 2$,

where $\hat{\omega}_k = \hat{\omega}_n(x_k^{(n)})$.

The similar formulas are valid for the five-degree spline-approximation.

Using the B-spline representation of the approximating spline allows us to organize the effective numerical integration in (15) because $\{s_{l,2m}(x)\}$ is a system of functions with compact support. It should be mentioned that the obtained method may be extended to the case of $p > 1$ by using the B-splines of the vector argument and the corresponding direct spline-approximation (Grebennikov, 1983).

4 Numerical example.

Consider the following nonlinear dynamic observation system

$$\begin{cases} y_n = y_{n-1}(1 + (y_{n-1})^2)^{-1} + 0.7w_n, & y_0 = \eta, \\ z_n = 0.6y_n + 0.4y_n^2 + 0.2v_n, \end{cases} \qquad (18)$$

where $\{w_n\}$ and $\{v_n\}$ are independent standard gaussian white noises and η is a standard gaussian random variable. The model (18) satisfies the conditions of Proposition 2 with $\alpha_n = 2$.

For $n = 1, 2, ..., 10$ this example was simulated on a computer. $M = 2000$ realizations $\{\hat{y}_n^i\}_{i=1}^M$ were obtained of the CMNF-estimate $\hat{y}_n$, $\{y_n^{*i}\}_{i=1}^M$ of the ONF-estimate y_n^* and $\{\tilde{y}_n^i\}_{i=1}^M$ of the EKF-estimate y_n^E (the EKF is described in (Sage,Melsa, 1972)).

The ONF-estimate y_n^* was calculated by the spline-filtering algorithm with $N = 20$, $a = -0.3$, $b = 3.0$, $p = 1$. The structural functions for the CMNF were taken as follows:

$\xi_n(\hat{y}_{n-1}) = \hat{y}_{n-1}(1 + (\hat{y}_{n-1})^2)^{-1}$,

$\pi_n(\tilde{y}_n, z_n) = \tilde{K}_n\Phi_n(\tilde{K}_n(\Phi_n)^2 + 0.04)^{-1}(z_n - 0.6\tilde{y}_n - 0.4(\tilde{y}_n)^2)$,

$\Phi_n = 0.6 + 0.8\tilde{y}_n$.

We have compared the following values of the mean-square errors of the above described estimates:

$$\hat{J}_n = M^{-1}\sum_{i=1}^{M}(y_n^i - \hat{y}_n^i)^2;$$
$$J_n^* = M^{-1}\sum_{i=1}^{M}(y_n^i - y_n^{*i})^2;$$
$$J_n^E = M^{-1}\sum_{i=1}^{M}(y_n^i - \tilde{y}_n^i)^2;$$

where $\{y_n^i\}_{i=1}^M$ is a set of realizations of the process $\{y_n\}$ simulated in accordance with (18). The following table shows the numerical results.

n	J_n^*	$\hat{J}_n$	J_n^E	$\hat{J}_n/J_n^*$	J_n^E/J_n^*
1	0.261	0.278	0.563	1.07	2.16
2	0.238	0.257	0.450	1.08	1.89
3	0.236	0.253	0.419	1.07	1.77
4	0.224	0.237	0.392	1.06	1.75
5	0.269	0.279	0.487	1.04	1.81
6	0.224	0.238	0.387	1.06	1.73
7	0.215	0.233	0.373	1.08	1.73
8	0.227	0.244	0.387	1.07	1.70
9	0.236	0.256	0.446	1.08	1.89
10	0.239	0.258	0.421	1.08	1.76

Table 1. Values of the mean-square criterion for the estimates of ONF , CMNF and EKF.

The obtained results show that the CMNF-estimate and the ONF-estimate have practically the same accuracy, and the EKF gives a much less accurate estimate. It should be mentioned that the algorithms CMNF and EKF require nearly the same computation time which is significantly less than the computation time of the ONF. This feature of the CMNF is that the optimal values F_n , H_n , f_n and h_n can be calculated before the process of filtering itself as in the case of the linear Kalman filter or the EKF. However, in the ONF it is necessary to integrate a complicate enough system each time with a new coming measurement. It takes substantional computational efforts in real time in comparison with the CMNF.

5 Appendix A.

Lemma. Let $z = col(x,y)$, $m = E\{z\} = col(m_x, m_y)$, $s = cov(z,z)$ is an unknown matrix which is upper bounded by the known matrix S

$$s \leq S = \begin{bmatrix} S_x & S_{xy} \\ S_{yx} & S_y \end{bmatrix}$$

Then

1)

$$\phi^*(y) = S_{xy}S_y^+ + (m_x - S_{xy}S_y^+ m_y) = m_x + S_{xy}S_y^+(y - m_y) = arg \min_{\phi \in \Phi} \max_{P(m,S)} \mathbf{E}\{||x - \phi(y)||^2\};$$

2) $\mathbf{E}\{x - \hat{x}\} = 0$, $cov(x - \hat{x}, x - \hat{x}) \leq S_x - S_{xy}S_y^+ S_{yx}$,
where $\hat{x} = \phi^*(y)$ is the estimate of x given y and the upper inequality holds as an equality in the case of $s = S$.

Proof. Let us denote
$f(\phi(.), p_z) = \mathbf{E}_{p_z}\{||x - \phi(y)||^2$, p_z^* is a gaussian distribution with expectation m and covariance matrix S , $\phi^*(y) = m_x + S_{xy}S_y^+(y - m_y)$. Then it is sufficiently to prove the following inequality

$$f(\phi^*(.), p_z) \leq f(\phi^*(.), p_z^*) \leq f(\phi(.), p_z^*) \tag{19}$$

for all $\phi(.) \in \Phi$, $p_z \in P(m, S)$.
$\phi^*(y) = \mathbf{E}_{p_z^*}\{x|y\}$ is known to be the best mean-square estimate of x given y and from the normal correlation theorem (Liptser,Shiryayev, 1977)it follows that the right inequality (19) is true.
Let $p_z \in P(m, S)$ is an arbitrary distribution with the expectation m and the covariance matrix $s \leq S$. It should be mentioned that

$$0 \leq c = S - s = \begin{bmatrix} c_x & c_{xy} \\ c_{yx} & c_y \end{bmatrix}$$

$\mathbf{E}_{p_z}\{(x - \phi^*(y))(x - \phi^*(y))^T\} = s_x - S_{xy}S_{yy}^+ s_{yy} S_{yy}^+ S_{yx}$
$- S_{xy}S_{yy}^+ s_{yx} - s_{xy}S_{yy}^+ S_{yx}$,

$\mathbf{E}_{p_z^*}\{(x - \phi^*(y))(x - \phi^*(y))^T\} = S_x - S_{xy}S_{yy}^+ S_{yx} =$
$s_x + c_x - S_{xy}S_{yy}^+(s_{yx} + c_{yx}) - (s_{xy} + c_{xy})S_{yy}^+ S_{yx} + S_{xy}S_{yy}^+(s_y + c_y)S_{yy}^+ S_{yx}$.
Then $\mathbf{E}_{p_z^*}\{(x - \phi^*(y))(x - \phi^*(y))^T\} - \mathbf{E}_{p_z}\{(x - \phi^*(y))(x - \phi^*(y))^T\} = qcq^T \geq 0$, where $q = [S_{xy}S_{yy}^+ \;\; I]$. It follows that
$f(\phi^*(.), p_z^*) - f(\phi^*(.), p_z) = trWqcq^T \geq 0$,
i.e. the left inequality (19) is true for all $p_z \in P(m, S)$. Lemma is proved.

Proof of Proposition 1. Suppose that $\mathbf{E}||y_{n-1}||^2 < \infty$. Let us check that $\mathbf{E}||y_n||^2 < \infty$.
$\mathbf{E}\{||\phi_n(y, w_n)||^2\} \leq \mathbf{E}\{C_n^\phi(w_n)(1 + ||y||^2)\} \leq \mathbf{E}\{C_n^\phi(w_n)\}(1 + ||y||^2)$ for any $y \in R^p$.

Hence $\mathbf{E}\{||y_n||^2\} = \mathbf{E}\{\mathbf{E}\{||\phi_n(y_{n-1}, w_n)||^2/y_{n-1}\}\} \le \mathbf{E}\{C_n^\phi(w_n)(1 + \mathbf{E}\{||y_{n-1}||^2\})\} \le \infty$. we have $y_0 = \eta$ and $\mathbf{E}\{||\eta||^2\} < \infty$ for all $n \ge 0$.
Let $\hat{y}_{n-1}$ be the CMNF estimate given Z^{n-1}. Suppose that $\mathbf{E}\{||\hat{y}_{n-1}||^2\} < \infty$. Let us show that $\{||\hat{y}_n||^2\} < \infty$. Consider the basic predictor $\bar{y}_n = \xi_n(\hat{y}_{n-1})$. $\mathbf{E}\{||\bar{y}_n||^2\} = \mathbf{E}\{||\xi_n(\hat{y}_{n-1})||^2\} \le C_n^\xi(1 + \mathbf{E}\{||\hat{y}_{n-1}||^2\}) < \infty$.
From the obtained results it follows that $x_1 = col(y_n, \bar{y}_n)$ is a random vector with a finite second order moment. Applying Lemma to $z = x_1$ with $x = y_n$, $y = \bar{y}_n$ we obtain the result $\tilde{y}_n = F_n\bar{y}_n + f_n$, where $F_n = cov(y_n, \bar{y}_n)cov^+(\bar{y}_n, \bar{y}_n)$, $f_n = \mathbf{E}\{y_n\} - F_n\mathbf{E}\{\bar{y}_n\}$. Then
$\mathbf{E}\{||\tilde{y}_n||^2\} = tr\{W(F_n\mathbf{E}\{\bar{y}_n\bar{y}_n^T\}F_n^T + 2f_n\mathbf{E}\{\bar{y}_nF_n^T)\} + ||f_n||^2 < \infty$.
Now let us show the existence of H_n and h_n. Let $\Delta\tilde{y}_n = y_n - \tilde{y}_n$. From Lemma it follows $\mathbf{E}\{\Delta\tilde{y}_n\} = 0$ and $\tilde{K}_n = cov(\Delta\tilde{y}_n, \Delta\tilde{y}_n) = K_n - cov(y_n, \bar{y}_n)cov^+(\bar{y}_n, \bar{y}_n)cov(\bar{y}_n, y_n) = K_n - F_ncov(\bar{y}_n, y_n)$ where $K_n = cov(y_n, y_n)$.
Let condition 3.a) be fulfiled, then
$\mathbf{E}\{||\nu_n||^2\} = \mathbf{E}\{||\pi_n(\tilde{y}_n, z_n)||^2\} \le C_n^\pi(1 + \mathbf{E}\{||\{y_{n-1}||^2\}) < \infty$.
Let now 3.b) be fulfiled, then taking into account the independence of $\{v_n\}$ and $\{y_n\}$ we have $\mathbf{E}\{||z_n||^2\} = \mathbf{E}\{||\psi_n(y_n, v_n)||^2\} \le \mathbf{E}\{C_n^\psi(v_n)\}(1 + \mathbf{E}\{||y_n||^2\}) < \infty$. Then we obtain $\mathbf{E}\{||\pi_n(\tilde{y}_n, z_n)||^2\} \le C_n^\pi(1 + \mathbf{E}\{||\tilde{y}_n||^2 + ||z_n||^2\}) < \infty$. So if 3.a) or 3.b) take place then $\mathbf{E}\{||\nu_n||^2\} < \infty$.
Applying Lemma to $x_2 = col(\Delta\tilde{y}_n, \nu_n)$ we obtain $\hat{y}_n = \tilde{y}_n + H_n\nu_n + h_n$, where $H_n = cov(\Delta\tilde{y}_n, \nu_n)cov^+(\nu_n, \nu_n)$ and $h_n = \mathbf{E}\{\Delta\tilde{y}_n - H_n\nu_n\} = -H_n\mathbf{E}\{\nu_n\} since \tilde{y}_n$ is unbiased.
$\mathbf{E}\{||y_n||^2\} \le 3(\mathbf{E}\{||\tilde{y}_n||^2\} + ||H_n||^2\mathbf{E}\{||\nu_n||^2\} + ||h_n||^2) < \infty$. Besides this $\mathbf{E}\{\Delta\hat{y}_n\} = \mathbf{E}\{y_n - \hat{y}_n\} = 0$ and $\hat{K}_n = cov(\Delta\hat{y}_n, \Delta\hat{y}_n) = \tilde{K}_n - cov(\Delta\hat{y}_n, \nu_n)cov^+(\nu_n, \nu_n)cov(\nu_n, \Delta\hat{y}_n) = \tilde{K}_n - H_ncov(\nu_n, \Delta\hat{y}_n)$.
Now let $n = 1$, then $y_{n-1} = \eta$ and $\hat{y}_0 = \mathbf{E}\{\eta\}$. $\mathbf{E}\{||\hat{y}_0||^2\} = ||\mathbf{E}\{\eta\}||^2 \le \mathbf{E}\{||\eta|||^2 < \infty$, hence $\hat{y}_1$ exists and $\mathbf{E}\{||\hat{y}_1||^2\} < \infty$. Now the proof follows from the mathematical induction principle.

6 Appendix B.

Proof of Proposition 2. Suppose that $\mathbf{E}\{||y_{n-1}||^k\} < \infty$ and $\mathbf{E}\{||\hat{y}_{n-1}||^k\} < \infty$ for all $k \ge 1$ and some $n \ge 1$. Let us show that y_n and $\hat{y}_n$ have the moments of arbitrary order. We shall use the following inequality: $(||x|| + ||y||)^r \le C_r(||x||^r + ||y||^r)$, $r \ge 1$, $C_r = 2^{r-1}$. It follows from the C_r-inequality (Loeve,1960): $(|x| + |y|)^r \le C_r(|x|^r + |y|^r)$ and the fact that $||x|| = (x^TWx)^{1/2} = |\tilde{x}|$, where $\tilde{x} = W^{1/2}x$.
$||y_n||^k = ||a_n(y_{n-1}) + b_n(y_{n-1})w_n||^k \le C_k(||a_n(y_{n-1})||^k + ||b_n(y_{n-1})w_n||^k) \le$
$C_k(C_N^1)^k[(1 + ||y_{n-1}||^{\alpha_n})^k + (1 + ||y_{n-1}||^{\alpha_n})^k||w_n||^k] \le$
$D_n^1[1 + ||y_{n-1}||^{\beta_n} + (1 + ||y_{n-1}||^{\beta_n})||w_n||^k]$, where $\beta = \alpha_nk$. Hence $\mathbf{E}\{||\hat{y}_n||^k\} \le D_n^1(1 + \mathbf{E}\{||y_{n-1}||^{\beta_n}\})(1 + \mathbf{E}\{||w_n||^k\}) < \infty$, since $\mathbf{E}\{||w_n||^k\} < \infty$ and y_{n-1} and w_n are independent.

$\bar{y}_n = \xi_n(\hat{y}_{n-1})$, then $\mathbf{E}\{||\bar{y}_n||^k\} = \mathbf{E}\{\xi_n(\hat{y}_{n-1})||^k\} \leq$
$(C_n^2)^k \mathbf{E}\{(1 + ||\hat{y}_{n-1}||^{\alpha_n})^k\} \leq D_n^2(1 + \mathbf{E}\{||\hat{y}_{n-1}||^{\beta_n}\}) < \infty$.
From the obtained inequalities it follows that $\mathbf{E}\{||x_1||^k\} < \infty$ for $k \geq 1$, where $x_1 = col(y_n, \bar{y}_n)$. Then by applying Lemma we obtain $\tilde{y}_n = F_n\bar{y}_n + f_n$, $||F_n|| + ||f_n|| < \infty$ and $\mathbf{E}\{||\tilde{y}_n||^k\} \leq C_k(||F_n||^k \mathbf{E}\{||\bar{y}_n||^k\} + ||f_n||^k) < \infty$.
$||\pi_n(y,z)||^k \leq (C_n^2)^k(1 + ||y||^{\alpha_n} + ||z||^{\alpha_n})^k \leq D_n^3(1 + ||y||^{\beta_n} + ||z||^{\beta_n})$. Then
$\mathbf{E}\{||\pi_n(\tilde{y}_n, z_n)||^k\} \leq D_n^3(1 + \mathbf{E}\{||\tilde{y}_n||^{\beta_n}\} + \mathbf{E}\{||z_n||^{\beta_n}\}) < \infty$ since
$\mathbf{E}\{||z_n||^{\beta_n}\} = \mathbf{E}\{||\psi_n(y_n) + v_n||^{\beta_n}\} \leq$
$C_{\beta_n}[\mathbf{E}\{||\psi_n(y_n)||^{\beta_n}\} + \mathbf{E}\{||v_n||^{\beta_n}\}] \leq$
$C_{\beta_n}(C_n^2)^{\beta_n}\mathbf{E}\{1 + ||y_n||^{\alpha_n})^{\beta_n}\} + C_{\beta_n}\mathbf{E}\{||v_n||^{\beta_n}\} \leq$
$D_n^4(1 + \mathbf{E}\{||y_n||^{\gamma_n}\}) + C_{\beta_n}\mathbf{E}\{||v_n||^{\beta_n}\} < \infty$
where $\gamma_n = \alpha_n\beta_n$; $\mathbf{E}\{||v_n||^{\beta_n}\} < \infty$ since v_n is a gaussian vector. Hence the vector $x_2 = col(\Delta\tilde{x}_n, \nu_n)$ has finite second-order moments and consequently H_n, h_n exist. From this, the desired result follows for $n > 1$. Now, if $n = 1$, then $y_{n-1} = y_0 = \eta$ and $\hat{y}_0 = \mathbf{E}\{\eta\}$. If η is a gaussian vector then $\mathbf{E}\{||y_0||^k\} < \infty$ and $\mathbf{E}\{$ —— $\hat{y}_0||^k\} = ||\mathbf{E}\{\eta\}||^k < \infty$. Now the desired result follows from the mathematical induction principle.

References

1. **Ahlberg,J.H., Nilson,E.N., and Walsh,J.L.,** 1967 The theory of splines and their applications. N.-Y., Academic Press.

2. **Andrade Netto,M.L., Gimeno,L., and Mendes,M.J.,** 1978 A new spline algorithms for nonlinear filtering of discrete-time systems. *7th IFAC World Congress.* Preprint 3, London, Pergamon Press, 2123-2130.

3. **Benes,V.E.,** 1981 Exact finite-dimensional nonlinear filters for certain diffusions with nonlinear drift. *Stochastics*, 5, 65-92.

3. **Cox,M.G.,** 1978 The numerical evaluation of a spline from its B-spline representation, *J. Inst. Maths. Applics.*, 21,135-143.

4. **Davis,M.H.A., and Marcus,S.I.,** 1981 An introduction to nonlinear filtering. Stochastic systems: the mathematics of filtering and identification and applications. Dotrecht, the Netherlands, 565-572.

5. **Daum,F.E.,** 1986 Exact finite-dimensional nonlinear filters. *IEEE Trans. on Automat. Control*, AC-31, 616-622.

6. **Grebennikov,A.I.,** 1983 Spline method and the solution of ill-conditioned problems of the approximation theory. Moscow, Moscow State University Press (in Russian).

7. **Hazewinkel,M., Marcus,S.I., and Sussan,H.J.,** 1983 Nonexistence of finite-dimensional filters for conditional statistics of the cubic sensor problem. *Syst. Contr. Lett.*, 3, 331-340.

8. **Liptser,R.S., and Shiryayev,A.N.,** 1977 Statistics of random processes. N.-Y.,Springer Verlag.

9. **Loeve,M.,** 1960 Probability theory. N.J., Princeton,D. Van Nostrand Company Inc.

10. **Pankov,A.R.,** 1990 Synthesis of conditionally-optimal filters by computer simulation. *Analysis and synthesis of dynamic systems under uncertainty*, Moscow, MAI Press, 69-75 (in Russian).

11. **Pankov,A.R.,** 1990 Recursive dynamic system trajectory estimation by means of regression nonlinear filter. *Statistical methods in flight control*, Moscow, MAI Press, 117-118 (in Russian).

12. **Pankov,A.R.,** 1990 Synthesis of conditionally-minimax filters for nonlinear stochastic systems by method of modeling. *The All-Union Conference on Dynamic, Control and Operation Research*, Moscow, MAI Press, 117-118 (in Russian).

13. **Pugachev,V.S.,** 1979 Estimation of variables and parameters in discrete-time nonlinear systems. *Automat. and Remote Control*, 40,512-521.

14. **Pugachev,V.S., and Sinitsyn,I.N.,** 1990 Stochastic differential systems, London, J.Wiley & Sons.

15. **Raol,S.R., and Sinha,N.K.,** 1987 Conditionally optimal state estimation for systems, governed by difference equations. *Canadian Electr. Engin. J.*, 12, 71-77.

16. **Sage,A.P., and Melsa,J.L.,** 1972 Estimation theory with applications to communications and control. N.-Y., McGrow-Hill.

17. **Sawitzki,G.,** 1981 Finite-dimensional filter systems in discrete time. *Stochastics*, 5, 107-114.

18. **Takeuchi,Y., and Akhashi,H.,** 1981 Nonlinear filtering formulas for discrete-

time observations. *SIAM J. Contr. Optimiz.*, 19, 244-261.

19. **Tam,L.-M., Wong,W.S., and Yau,S.S.-T.,** 1990 On necessary and sufficient conditions for finite dimensionality of estimation algebras. *SIAM J. Contr. Optimiz.*, 28, 173-185.

20. **Villalobos,M., and Wahba,G.,** 1987 Inequality-constrained multivariate smoothing splines with applications to the estimation of posterior probabilities. *J. of Amer. Statist. Association*, 82, 239-248.

21. **Wegman,E., and Wright,I.,** 1983 Splines in statistics. *J. of Amer. Statist. Association*, 78, 351-366.

22. **Wong,W.S.,** 1983 New classes of finite-dimensional nonlinear filters. *Syst. Contr. Lett.*, 3, 155-164.