Listy Biometryczne – Biometrical Letters
Vol. 30 (1993), No. 2, 65-71

A discrete state approximation in the nonlinear filtering problem

V. Lobatch

Department of Mathematical Modelling and Data Analysis
Belorussian State University

Summary

A computable approximation to the nonlinear filtering problem, where the system and data models are given by $dx = \alpha(t)\,x\,dt + \sigma(t)$, $dz = h(t,x)dt + r(t)\nu(t)$, is treated. The approximation (with approximation parameter h) gives estimates which converge to the optimal filter.

1. Introduction

In the ECG signal analysis with the aim of the enhancement of its potentials the modern statistical methods of the signal analysis are used (Visser, Molenaar, 1988; Madhavan, 1989; Kadhim Kadhim et al., 1988). The filtering theory is a base of these methods. A linear model has been used as a suitable first approximate model in many practical biometrical problems. But in some problems such as the description of the phenomenon of biological growth, it is necessary to use only nonlinear models for the identical description of these phenomena (Visser, Malenaar, 1988). Therefore the optimal or suboptimal nonlinear filtering algorithms can be used in these problems.

In the present investigation, we propose a method of the solution of nonlinear filtering problem which can be applied to obtain the ECG signal from the noisy observations.

Let $w(t)$, $v(t)$ denote two independent Wiener processes and define the processes $x(t)$, $z(t)$ by the Ito equations

$$dx(t) = a(t)\,x(t)\,dt + \sigma(t)\,dw(t)\ , \tag{1}$$

Key words: nonlinear filtering problem, ECG signal analysis, least squares estimation

$$dz(t) = h(t, x)\, dt + r(t)\, \nu(t) \ , \tag{2}$$

where $x(0) = x_o$ – a Gaussian random variable with parameters $\mu_o, \gamma_o, y(0) = 0, t < T$, with T denoting an arbitrary, but fixed positive number. It is assumed that $f(t)$, $\sigma(t)$, $r(t)$, $h(t, x)$ are continuous functions and function $h(t, x)$ is of the form

$$h(t, x) = x\, g(t, x) \ , \tag{3}$$

where $g(t, x)$ satisfies the Lipshitz condition with constant L.

The observed process (2) is given by a nonlinear equation. The nonlinearity of this equation is noninversible. This situation is typical in the theory of extremal systems. For extremal systems, the filtering problem, that is estimation of the state variable $x(t)$ based on the observed process $z(t)$, $0 < t < T$, is denoted by Wonham (1970) as an open (unsolved) problem.

If $\mathcal{F}_t$ is the σ-algebra generated by $z(\tau)$, $\tau \le t$, then the least squares estimate of $x(t)$ based upon the observations $z(\tau)$, $\tau \le t$, is given by the conditional expectation $E(x(t)|\mathcal{F}_t)$. Although stochastic differential equations for $E(x(t)|\mathcal{F}_t)$ are known, the problem how to effectively compute $E(x(t)|\mathcal{F}_t)$ is still in bad shape. It is known (Lipster and Shiryayev, 1977, 1978) that causal least squares state estimation for system (1) and (2) requires, in general, real-time computation of the solution of an infinite set of coupled stochastic differential equations in order to generate the estimate $\hat{x}(t) = E(x(t)|\mathcal{F}_t)$ of the state variable $x(t)$. As a result nonlinear filters (which compute $\hat{x}(t)$, $t \ge 0$) have not become practical, yet.

2. A useful approximating process

In this section we describe a particular approximating process which is to be used in Section 3. Let D_h denote the finite difference grid on $(0,t)$, $0 < t < T$, with difference parameter h (Lobatch, 1989):

$$t_k = kh, \quad h = 2^{-n}t, \quad k = 0,1,\ldots,2^n = N \ . \tag{4}$$

Let us define a process $\{\xi_N^{(i)}(\tau)\}$ as follows

$$\xi_N^{(i)}(\tau) = x(\tau)\, g_i(\tau,\theta)\, d\tau + r(\tau)\, d\nu(\tau), \; t_i < \tau < t_{i+1}, \; i = \overline{0, N-1} \ , \tag{5}$$

$$\theta_i = x(t_i), \; i = \overline{0, N-1} \ ,$$

$$g_i(\tau,\theta) = g(\tau,\theta_i), \; i = \overline{0, N-1}, \; \xi_N(\tau) = \{\xi_N^{(i)}(\tau), \; i = \overline{0, N-1}\} \ . \tag{6}$$

Theorem 1. The following limit holds

$$\lim \xi_N^{(i)}(\tau) \overset{\text{a.m.}}{=} z(\tau) . \qquad (7)$$

Proof. It is known that the Wiener process has the following property

$$|w(\tau) - w(s)| < \beta \cdot |\tau - s|^{\alpha}, \qquad (8)$$

where $0 < \alpha < \frac{1}{2}$, β – random variable, $P\{\beta < \infty\} = 1$. Thus, with probability 1,

$$|z(t) - \xi_N(t)| = |\sum_{i=0}^{N-1} \int_{\tau_i}^{\tau_i+1} x(s) [g(x(s)) - g(x(\tau_i))] ds| < \sum_{i=0}^{N-1} \int_{\tau_i}^{\tau_i+1} |x(s)| \cdot L \cdot |x(s) - x(\tau_i)| ds$$

$$= \sum_{i=0}^{N-1} |x(t_i)| \cdot L \cdot |x(t_i) - x(\tau_i)| \cdot |\tau_{i+1} - \tau_i| ,$$

where $t_i = \tau_i + l(\tau_{i+1} - \tau_i),\ 0 < l < 1$.

Then

$$|x(t_i) - x(\tau_i)| \le A_1 \int_{\tau_i}^{t_i} |x(s) - x(\tau_i)| d\tau + A_1 |t_i - \tau_i| \cdot |x(\tau_i)| + A_2 \cdot |w(t_i) - w(\tau_i)| \le$$

$$A_1 \int_{\tau_i}^{t_i} |x(s) - x(\tau_i)| ds + A_1 |t_i - s_i| \cdot |x(t_i)| + A_2 \cdot \beta \cdot |t_i - \tau_i|^{\alpha} =$$

$$A_1 \int_{\tau_i}^{t_i} |x(s) - x(\tau_i)| ds + \gamma \cdot |t_i - \tau_i|^{\alpha}$$

with probability 1, where

$$A_1 = \max_{0 < t < N} |a(t)|, \quad A_2 = \max_{0 < t < T} |\sigma(t)| ,$$

$$\gamma = \max_{0 < i < N} \{ A_1 |t_{i+1} - t_i|^{1-\alpha} \cdot x(t_i) \} + A_2 \beta.$$

Using the Bellman lemma (see Hartman, 1964) we have

$$|x(\tau) - x(t_i)| \overset{\text{a.m.}}{\le} \gamma \cdot |t_{i+1} - t_i|^{\alpha} + A_1 \int_{t_i}^{t_i+1} \exp \{A_1(t_{i+1} - \tau)\} \gamma \cdot |\tau - t_i|^{\alpha} d\tau \le \gamma_i \cdot |t_{i+1} - t_i|$$

with probability 1, where

$$\gamma_i = \gamma + A_1 \int_{t_i}^{t_i+1} \exp\{A_1(t_{i+1} - \tau)\}\, d\tau .$$

Then

$$|z(t) - \zeta_N(t)| \overset{\text{a.m.}}{\leq} \sum_{i=0}^{N-1} |x(t_{i+1})| \cdot A_1 \cdot \gamma_1 |t_{i+1} - t_i|^{\alpha} \cdot |t_{i+1} - t_i| \overset{\text{a.m.}}{\leq}$$

$$\max_{0 \leq i \leq N-1} (A_1 \cdot |x(t_{i+1})| \cdot \gamma_1 |t_{i+1} - t_i|^{\alpha}) \cdot \sum_{i=0}^{N-1} |t_{i+1} - t_i| \overset{\text{a.m.}}{\leq}$$

$$\max_{0 \leq i \leq N-1} (A_1 \cdot |x(t_{i+1})| \cdot \gamma_1 \cdot |t_{i+1} - t_i|^{\alpha}) \cdot t \xrightarrow[N \to \infty]{\text{a.m.}} 0 .$$

Theorem 1 is proved.

Theorem 2. It holds w.p. 1

$$p\{y | \mathcal{F}_{t,N}^{\xi}\} \to p\{y | \mathcal{F}_t^z\} , \tag{9}$$

where $y \in R^1$, $\mathcal{F}_{t,N}^{\xi}$, $\mathcal{F}_t^z$ denote σ-algebras generated by $\xi_{\tau,N}$, $\tau \leq t$, and z_τ, $\tau \leq t$, respectively.

Proof. Let D_n denote the division of [0,t] defined by (4). Then we have

$$\{\xi(0), \xi(t_1),\dots,\xi(t_N)\} \in \{\xi(0), \xi(t_1),\dots,\xi(t_{N_1})\} ,$$

$N_1 = 2^{n+1} = 2N$. Then

$$\mathcal{F}_{t,N}^{\xi} \in \mathcal{F}_{t,N_1}^{\xi} \quad ,$$

where $\mathcal{F}_{t,N}^{\xi} = \sigma\{\xi(0), \xi(t_1),\dots,\xi(t_n)\}$.

To show (9) we only need to prove that

$$\mathcal{F}_{t,N}^{\xi} \to \mathcal{F}_t^z \ . \tag{10}$$

Using $\xi_{t,N} \to z(t)$ we have (10). Next, it is necessary to show that

$$S_N = p\{y | \mathcal{F}_{t,N}^{\xi}\}$$

is a martingale. Using

$$\mathcal{F}_{t,2^{N-1}} \in \mathcal{F}_{t,2^{N-1}}, N = 2^n ,$$

we obtain

$$E\{S_{2^n}|\mathcal{F}^{\xi}_{t,2^{n-1}}, \mathcal{F}^{\xi}_{t,2^{n-2}},\dots,\mathcal{F}^{\xi}_{t,2^0}\} = E\{S_{2^n}\,|\mathcal{F}^{\xi}_{t,2^{n-1}}\} = p\{y|\mathcal{F}^{\xi}_{t,2^{n-1}}\} = S_{2^n} .$$

Using the convergence theorem of martingales we have

$$p\{y|\mathcal{F}^{\xi}_{t,N}\} \to p\{y|\mathcal{F}^{z}_{t}\} .$$

The theorem 2 is proved.

Remark. It is well known (Busy, 1965) that w.p. 1

$$p\{y|\mathcal{F}^z_t\} = p_t\{y\} = \frac{E(\exp\{\int_0^t h(\tau, x(\tau))\frac{1}{R}\,dz(\tau) - \frac{1}{2R}\int_0^t h^2(\tau, x(\tau))\,d\tau\}|x(t) = y, \mathcal{F}^z_t)}{E(\exp\{\int_0^t h(\tau, x(\tau))\frac{1}{R}\,dz(\tau) - \frac{1}{2R}\int_0^t h(\tau, x(\tau))\,d\tau\}|\mathcal{F}^z_t)}, \tag{11}$$

where $E\{\cdot, \mathcal{F}^z_t\}$ denotes the expectation over $(\cdot)$ given $z(\tau)$, $\tau \le t$, $p_t(y)$ is a density of $x(t)$, t is fixed, $R = r^2$.

3. The filtration theorem

Theorem 3. The least squares estimate of $x(t)$ based upon the observation $\{\xi_{\tau,N}\} = \{\xi^{(i)}_{\tau,N},\ t_i \le \tau \le t_{i+1},\ i = \overline{0, N-1}\}$ is defined by

$$\hat{x}_N(t) = \int \hat{x}_N(\theta)\, p\{\theta|\mathcal{F}^{\xi}_{t,N}\}\, d\theta , \tag{12}$$

where $\hat{x}_{t,N}(\theta)$ is a set of Kalman – Busy filters and the conditional density $p\{\theta|\mathcal{F}^{\xi}_{t,N}\}$ is defined in (11).

Proof. It is well known that the conditional expectation $E\{x(t)|\mathcal{F}^{\xi}_{t,N}\}$ is the least squares estimate of $x(t)$ based upon the observation $\{y(\tau),\ 0 \le \tau \le t\}$. Using the properties of the conditional expectation we have

$$\{x(t)|\mathcal{F}^{\xi}_{t,N}\} + \int E\{x(t)|\mathcal{F}^{\xi}_{t,N}, \theta\}\, p\{\theta|\mathcal{F}^{\xi}_{t,N}\}\, d\theta ,$$

where $E\{x(t)|\mathcal{F}^{\xi}_{t,N}, \theta\}$ is the conditional expectation of $x(t)$ based on $\{\xi_{\tau,N}, \tau \le t\}$ and θ.

The optimal Kalman – Busy filter estimator and error covariance update equations are given by Lipster and Shiryayev (1977, 1978):

$$d\hat{x}_N(t,\theta) = a(t)\,\hat{x}_N(t,\theta)\,dt + g_N(t,\theta)\,\gamma_t(\theta)\frac{1}{R}\,(d\xi_{t,N} - \hat{x}_N(t,\theta)\,g_N(t,\theta)\,\gamma_t(\theta) =$$

$$2a(t)\,\gamma_t(\theta) + \sigma^2(t) - [\,g_N(t,\theta)\,\gamma_t(\theta)]^2\,\frac{1}{R}\,, \tag{13}$$

$$t_k \le t \le t_{k+1},\quad k = 0,1,2,\ldots,N-1,\quad x_N(0,0) = \mu_0,\quad \gamma_0(\theta)\ .$$

Lainiotis (1971) showed that

$$p\,\{\,\theta\,|\,\mathcal{F}^{\xi}_{t,N}\} = \frac{\exp\{-\frac{1}{2R}\int_0^t(\hat{x}_N(\tau,\theta)\,g_N(\tau,\theta))^2\,d\tau + \int_0^t\hat{x}_N(\tau,\theta)\,d\,\xi_N(\tau)\,\}\,p(\theta\,|\,t_N)}{\int\exp\{-\frac{1}{2R}\int_0^t(\hat{x}_N(\tau,\theta)\,g_N(\tau,\theta))^2\,d\tau + \int_0^t\hat{x}_N(\tau,\theta)\,d\,\xi_N(\tau)\,\}\,p(\theta\,|\,t_N)\,d\theta} \tag{14}$$

where

$$p(\theta\,|\,t_N) = \frac{dP\,\{\,x(t_N) < \theta\,\}}{d\theta}$$

and $\hat{x}_N(\tau,\theta)$ is defined by equations (12) and (13). Theorem 3 is proved.

Theorem 4. It holds w.p. 1

$$\lim \hat{x}_N(t) = \hat{x}(t)\ ,$$

where $x(t) = E\,\{\,x(t)\,|\,\mathcal{F}^{z}_{t}\,\}$.

Proof. The conditional expectations $\hat{x}_N(t,\theta)$ are defined by (12), (13). Using (14) we have

$$\int\hat{x}_N(t,\theta)\cdot p\,\{\,\theta\,|\,\mathcal{F}^{\xi}_{t,N}\}\,d\theta < C\,,\quad C = \text{const.}\ ,$$

w.p. 1. Using the Lebesque theorem on majorante covergence we obtain

$$\int\hat{x}_N(t,\theta)\,p\,\{\,\theta\,|\,\mathcal{F}^{\xi}_{t,N}\}\,d\theta = \int E\,\{x(t)\,|\,\mathcal{F}^{\xi}_{t,N},\,\theta\,\}\,p\,\{\,\theta\,|\,\mathcal{F}^{\xi}_{t,N}\}\,d\theta \xrightarrow{\text{w.p.1}}$$

$$\int E\,\{x(t)\,|\,\mathcal{F}^{\xi}_{t},\,\theta\,\}\,p\,\{\,\theta\,|\,\mathcal{F}^{\xi}_{t}\}\,d\theta = \hat{x}(t)\,.$$

We have shown that

$$\hat{x}_N(t) \xrightarrow{\text{w.p.1}} \hat{x}(t)\ .$$

The theorem 4 is proved.

Remark on computation. Let us define a process denoted by $\xi_{t,N}$ which converges to $x(t)$ for $N \to \infty$. The method used in this paper can be used with the process $x(t)$ replacing approximation $\xi_{t,N}$ in the formula (12).

REFERENCES

Busy R.S.(1965). Nonlinear filtering theory. *IEEE Trans. Autom. Control* **10**, 198-201.

Hartman Ph. (1964). *Ordinary Differential Equations.* Wiley, New York.

Kadhim Kadhim N., Taha Saleem M.R., Mahmoud Walled A. (1988). A new method for filtering and segmentation of the ECG signal. *Proc. 10th Annu. Int. Conf. IEEE Eng. Med. and Biol. Soc.,* New Orleans, 154-155.

Lainiotis D.G. (1971). Optimal nonlinear estimation. *Int. J. Control* **14**, 1137-1148.

Lipster R.S., Shiryayev A.N. (1977). *Statistics of Random Processes, I: General Theory.* Springer – Verlag, New York.

Lipster R.S., Shiryayev A.N. (1978). *Statistics of Random Processes, II: Applications.* Springer – Verlag, New York.

Lobatch V.I. (1989). Optimization of the filtering process in the extremal systems. In: *Proceedings of the International Conference on Stochastic Methods in Experimental Sciences COSMEX'89.* Technical Universtity of Wrocław, 55-56.

Madhavan G. (1989). Adaptive filtering of brain evoked potentials. *22nd Asilomar Conf. Signals, Syst. and Comput., Conf. Rec.* Vol. 1-2, San Jose, Calif., 924-928.

Visser H., Molenaar J. (1988). Kalman filter analysis in dendroclimatology. *Biometrics* **44**, 929-940.

Wonham W.M. (1970). Random differential equations in control theory. In: A.T. Bharucha – Reid, Ed., *Probabilistic Methods in Applied Mathematics*, Vol. 2, Academic Press, New York, 131-212.

Received 28 May 1992; revised 20 December 1992