Control
and Cybernetics
VOL. 15 (1986) No. 1

# Exact finite — dimensional predictors for certain diffusion with non-linear drift

by

TADEUSZ BANEK

Department of Technical Education
University M. Curie-Skłodowska
Narutowicza 12
20-004 Lublin, Poland

Let $w_t$ and $b_t$ be independent Wiener processes, and consider the task of estimating a diffusion $x_t$ solving the stochastic DE $dx_t=f(x_t)\,dt+dw_t$, $x_0=x$, at the moment $\tau$, $\tau>t$, on the basis of noisy observations $\{y_s,\ 0\leqslant s\leqslant t\}$ defined by $dy_t=x_t\,dt+db_t$.

THEOREM. *If* $f'+f^2=az^2+bz+c$, $a\geqslant-1$ *then the unnormalized conditional density*

$$q^{\tau}_{t}(z, x)\propto P\{x_\tau\in dz\mid y_s,\ 0\leqslant s\leqslant t,\ \tau>t\}$$

*can be written explicitly in terms of a finite number of sufficient statistics.*

## 1. Introduction

We shall be concerned with the estimation of a "system process" $x_t$, $0\leqslant t\leqslant T$ which we assume to be defined as a stochastic diffusion process on the known probability space $(\Omega, F, P)$ solving the stochastic DE

$$dx_t=f(x_t)\,dt+dw_t, \qquad x_0=x. \tag{1.1}$$

It is further assumed that the system process cannot be observed directly. Instead we have available an "observation process" $y_t$ which is given by

$$dy_t=x_t\,dt+db_t, \qquad y_0=0. \tag{1.2}$$

where $(w_t, b_t, F_t)$ is a 2-dimensional Wiener process with respect to $F_t$, $F_t\subset F$ $0\leqslant t\leqslant T$, system of non-decreasing sub-$\sigma$-algebras of $F$.

Our available data is $\{y_s,\ 0\leqslant s\leqslant t\}$ and $x$, and using this data we wish to estimate $x_\tau$, at the moment $\tau$, $0\leqslant\tau\leqslant T$.

This problem will be called as follows

— the prediction problem if $\tau>t$

— the filtering problem if $\tau=t$

— the smoothing problem if $\tau<t$.

Virtually the only solutions of the prediction problem have been obtained to date for the linear dynamics case $f(z)=\alpha z+\beta$, but after Beneš paper [1] it became quite obvious, that even for some nonlinear $f$ the problem is finite-dimensional.

In his excellent paper [1], Beneš has solved the filtering problem, under condition

$$(*) \qquad f'+f^2=az^2+bz+c, \quad a\geqslant -1$$

writing explicity $q_t^t(z, x)$. In this paper we will study the prediction problem and under "Beneš condition" $(*)$ we will exhibit its exact solution. It turns out that $q_t^\tau(z, x)$ depends on $a$ drift $f$ in exactly the same way as $q_t^t(z, x)$ does. Dependence on sufficient statistics is obviously more complicated but the number of statistics is finite (the finite-dimensional computable solutions of the prediction problem are called finite-diemensional predictors). Applications of the solutions in real situations is considered and connection with smoothing is pointed out.

## 2. Solution of the prediction problem

Our main result is:

THEOREM. *Let $f'+f^2=az^2+bz+c$, $a\geqslant-1$.* Then

$$q_t^\tau(z, x)=\varphi(t)\left(\frac{R}{r}\right)^{1/2} \exp\left\{A_1+2\kappa^2 R-\eta+\int_x^z f(u)\,du\right\}, \tag{2.1}$$

where

$$r=r(\tau)\equiv a^{-1/2}\, th\, a^{1/2}(\tau-t) \tag{2.2}$$

$$\sigma=\sigma(t)\equiv (a-1)^{-1/2}\, th\,(a-1)^{1/2}\, t \tag{2.3}$$

$$A=A(\tau, t)\equiv\exp\left\{-a\int_t^\tau r(s)\,ds\right\} \tag{2.4}$$

$$B=B(\tau, t)\equiv\frac{b}{2}\int_t^\tau A(\tau, s)\,ds \tag{2.5}$$

$$A_1=A_1(\tau, t)\equiv\frac{1}{2}\int_t^\tau [c+ar(s)-bB(\tau, s)+aB^2(\tau, s)]\,ds \tag{2.6}$$

$$A_2=A_2(\tau, t)\equiv -B+a\int_t^\tau A(\tau, s)\,B(\tau, s)\,ds \tag{2.7}$$

$$A_3=A_3(\tau, t)\equiv\frac{1}{2}a\int_t^\tau A^2(\tau, s)\,ds \tag{2.8}$$

$$R=R(\tau, t)\equiv r\sigma\,[r+\sigma A^2+2r\sigma A_3]^{-1} \tag{2.9}$$

$$C=C(t, s)\equiv \exp\left\{-(a-1)\int_s^t \sigma(u)\,du\right\} \tag{2.10}$$

$$\mu=\mu_t\equiv xC(t,0)-\frac{1}{2}b\int_0^t \sigma(s)\,C(t,s)\,ds+\int_0^t \sigma(s)\,C(t,s)\,dy_s \tag{2.11}$$

$$\kappa=\kappa(\tau, z, t, x)\equiv\frac{1}{2}\,[r\sigma A_2+\mu r+(z+B)\,\sigma A]\,(r\sigma)^{-1} \tag{2.12}$$

$$\eta=\eta(\tau, z, t, x)\equiv\frac{1}{2}\,[\mu^2\,\sigma^{-1}+(z+B)^2\,r^{-1}] \tag{2.13}$$

For definition of $\varphi(t)$ see proof.

P r o o f. Under condition (*) Beneš [1, Theorem 8] showed that

$$q_t^t(\xi, x)=\varphi(t)\exp\left\{-\frac{(\xi-\mu)^2}{2\sigma}+\int_x^\xi f(u)\,du\right\} \tag{2.14}$$

where $\varphi(t)$ is time — function which does not depend on $x$ and/or $\xi$, $\sigma$ and $\mu$ are statistics satisfying (2.3) (2.11). We shall show now that

$$q_t^\tau(z, x)=\int_{-\infty}^{+\infty}\Gamma(\tau, z, t, \xi)\,q_t^t(\xi, x)\,d\xi \tag{2.15}$$

where $\Gamma(\tau, z, t, \xi)$ is the fundamental solution of the PDE

$$\frac{\partial u}{\partial t}=\frac{1}{2}\,\frac{\partial^2 u}{\partial z^2}-\frac{\partial}{\partial z}(fu)\,. \tag{2.16}$$

To do this, let us denote

$$A_t\equiv\exp\left\{\int_0^t x_s\,dy_s-\frac{1}{2}\int_0^t x_s^2\,ds\right\}.$$

From Kallianpur—Striebel formula ([4], Corollary, pp. 800) we have

$$q_t^\tau(z, x)\,dz=E1_{\{z\leqslant x_\tau<z+dz\}}\,A_t \tag{2.15'}$$

where $1_A$ is an indicator of the set $A$.

Using well known properties of Markov processes we can express the right hand side of (2.15′) as follows

$$\int_{\xi=-\infty}^{\xi=+\infty} E1_{\left\{\substack{\xi\leqslant x_t<\xi+d\xi\\ z\leqslant(x_\tau-x_t)+x_t<z+dz}\right\}}A_t=$$

$$=\int_{\xi=-\infty}^{\xi=+\infty} E1_{\{\xi\leqslant x_t<\xi+d\xi\}}\,A_t\cdot E1_{\left\{\substack{x_t=\xi\\ z-\xi\leqslant x_\tau-x_t<z-\xi+d(z-\xi)}\right\}}.$$

Now we recognize from $K$–$S$ formula that

$$E1_{\{\xi\leqslant x_t<\xi+d\xi\}}\,\Lambda_t=q_t^t(\xi,x)\,d\xi$$

and from the Markov diffusion theory that

$$E1_{\left\{\substack{x_t=\xi\\ z-\xi\leqslant x_\tau-x_t<z-\xi+d(z-\xi)}\right\}}=\Gamma(\tau,z,t,\xi)\,dz\,. \qquad \blacksquare$$

This completes the proof of (2.15). Note that the equality (2.15) has an obvious interpretation from parabolic PDE theory point of view namely, it is a solution of the Cauchy problem for (2.16) with the initial condition $u(t,\xi)=q_t^t(\xi,x)$, see [3]. Equation (2.16) is not directly integrable, but let us define $v(\tau,z)\equiv u(\tau,z)\exp\{+-\int_\xi^z f(u)\,du\}$ and calculate both sides of (2.16)

$$\frac{\partial u}{\partial\tau}=\frac{\partial v}{\partial\tau}\exp\left\{\int_\xi^z f(u)\,du\right\}$$

$$\frac{\partial^2 u}{\partial z^2}=\left[\frac{\partial^2 v}{\partial z^2}+2f\frac{\partial v}{\partial z}+(f'+f^2)v\right]\exp\left\{\int_\xi^z f(u)\,du\right\} \tag{2.16'}$$

$$\frac{\partial(fu)}{\partial z}=\left[(f'+f^2)\,v+f\frac{\partial v}{\partial z}\right]\exp\left\{\int_\xi^z f(u)\,du\right\}.$$

Substituting this formulas into (2.16) we get the equation for $v(\tau,z)$:

$$\frac{\partial v}{\partial\tau}=\frac{1}{2}\frac{\partial^2 v}{\partial z^2}-\frac{1}{2}(f'+f^2)\,v=\frac{1}{2}\frac{\partial^2 v}{\partial z^2}-\frac{1}{2}(az^2+bz+c)\,v \tag{2.17}$$

which is now directly integrable in view of the quadratic potential term. Integrating (2.17) we get the formulas:

$$\Gamma(\tau,z,t,\xi)=(2\pi r)^{-1/2}\exp\left\{-\frac{(z-m)^2}{2r}+\lambda+\int_\xi^z f(u)\,du\right\} \tag{2.18}$$

where $r$ is given by (2.2),

$$\frac{dm}{d\tau}=-arm-\frac{1}{2}br\,,\quad m(t)=\xi \tag{2.19}$$

$$\frac{d\lambda}{d\tau}=-\frac{1}{2}(c+ar+bm+am^2),\quad \lambda(t)=0\,. \tag{2.20}$$

Using formulas (2.6), (2.7), (2.8), (2.19) we can express the solution of (2.20) in the form

$$\lambda\equiv\lambda(\tau,t,\xi)=\Lambda_1+\Lambda_2\,\xi+\Lambda_3\,\xi^2 \tag{2.21}$$

which together with (2.14), (2.18), (2.19) makes it possible to use (2.15). Simple manipulations show that the right hand side of (2.15)

$$(2\pi r)^{-1/2}\,\varphi(t)\exp\left\{A_1+\int_x^z f(u)\,du\right\}\times$$
$$\times\int_{-\infty}^{+\infty}\exp\left\{A_2\,\xi+A_3\,\xi^2-\frac{(\xi-\mu)^2}{2\sigma}-\frac{(z+B-A\xi)^2}{2r}\right\}d\xi\,. \quad (2.22)$$

This last integration leads directly to (2.1).

### 3. Reduced density

Examination of the formula (2.1) for the unnormalized conditional density shows that normalization will reduce dependence on all the sufficient statistics except:

$$M\equiv M(\tau,t)=\left[RA\left(A_2+\frac{\mu}{\sigma}+\frac{AB}{r}\right)-B\right](1-RA)^{-1} \quad (3.1)$$

depending on the observation process $y_t$ (through $\mu$) and

$$\sigma,\,r,\,A,\,B,\,C,\,R,\,N\equiv(1-RA^2)\,r^{-1} \quad (3.2)$$

which are deterministic time functions.

After a bit of simple calculation one can express the reduced density as

$$\exp\left\{-\frac{(z-M)^2}{2N^2}+\int_x^z f(u)\,du\right\}. \quad (3.3)$$

Except for the part involving $\int_x^z f(u)\,du$, the reduced conditional density is determined like the Gaussian, by the quadratic function in its exponent.

### 4. Applications

It is almost trivial to say that the linear stochastic diffusion process (called also the Ornstein–Uhlenbeck process) is very popular in many areas of investigations. The linear models used in control, filtering and prediction theories allow to formulate and solve relevant problems rigorously. What is even more important for applications, many problems not of this (linear) form are approximately modelled as linear and the solutions are used as approximations of the solutions for the original situations. However, assumption of linearity is a serious restriction imposed on the model of a systems. It is very important to extend the scope of the relevant mathematical methods and explicity obtained results. We begin this section by presenting examples of functions $f$ satisfying the Beneš condition $(*)$ (other examples and extensions to the multidimensional case can be found in [1])

(!) $$f(z)=\theta\,(Ae^{\theta z}-Be^{-\theta z})/(Ae^{\theta z}+Be^{-\theta z});$$

for which $a=b=0$, $c=\theta^2$. For $A=B$, $f(z)=\theta$ th $\theta_z$

(!!) the linear case $f(z)=\alpha z+\beta$ meets the hypothesis $(*)$, with $a=\alpha^2$, $b=2\alpha\beta$, $c=\alpha+\beta^2$.

As we see, the condition $(*)$ is satisfied for large classes of interesting nonlinear drifts $f$ including the linear case (!!) and this fact implies usefulness of the diffusion (1.1), $(*)$. Statement of the prediction problem and its solution presented in section 2 and 3 allows us to refer to some practical situations. Two of them are shown below:

A. Delayed observations

Let us assume that the signal process $x_t$ is — as previously — given by (1.1), $(*)$, but the observation process

$$\tilde{y}_t=\int_0^t x_{s-h}\,ds+b_t \tag{4.1}$$

and initial condition $x=x_s$ for $-h\leqslant s\leqslant 0$, where $h>0$ is the time delay appearing in a information channel and/or a computational machinery calculating the conditional density $q_t^t(z, x)$. Now let us define the "new" observation process

$$\bar{y}_{t-h}=\int_0^{t-h} x_s\,ds+b_{t-h} \tag{4.2}$$

of the form (1.2), where $\bar{b}_{t-h}\equiv b_t$.

Introducing the new variable $u=s+h$ we have from (4.2)

$$\bar{y}_{t-h}=\int_h^t x_{u-h}\,du+b_{t-h}.$$

Defining $\tilde{\bar{y}}_t\equiv\bar{y}_{t-h}-xh$, we see that $\tilde{\bar{y}}_t\equiv\tilde{y}_t$, hence for the problem of estimating $x_t$ on the basis of observations (4.1) our formulas (2.1), (3.3) are applicable with the substitutions $\tau\equiv t+h$ and $y_t\equiv\bar{y}_t=xh+\tilde{y}_t$.

B. Double observations

For the same signal process $x_t$ as previously, we have available data which consist of two components, the first

$$\{y_s,\ 0\leqslant s\leqslant t\} \quad \text{where } y_s=\int_0^s x_u\,du+b_s$$

and the second

$$\{x_\sigma, \quad 0 \leqslant \sigma \leqslant t-h, \quad h>0\}.$$

This model of observation is useful, for instance, in certain economical situations which occur oftentimes in practice Following [2], [5], [6] we only mention demand processes which are perfectly observable up to (for example) yesterday, are noisy today and have to be predicted for the next few days Obviously our result applies with the substitutions "$h$" instead of "$t$" and "$x_{t-h}$" instead of "$x$" in the formulas of sections 2 and 3

## 5. Smoothing

Let $\bar{\Gamma}(\tau, z, t, \xi)$, $\tau<t$, be the fundamental solution of the backward PDE:

$$\frac{\partial u}{\partial \tau}+\frac{1}{2}\frac{\partial^2 u}{\partial z^2}+f\frac{\partial u}{\partial z}=0$$

From the theory of PDE ([3], theorem 15 pp. 28) we know that $\bar{\Gamma}(\tau, z, t, \xi)= =\Gamma(t, \xi, \tau, z)$ where $\Gamma(\tau, z, t, \xi)$ is given by (2.18), (2.2), (2.19), (2.20). Again using formula (2.15) with $\bar{\Gamma}(\tau, z, t, \xi)$ instead of $\Gamma(\tau, z, t, \xi)$ we can calculate $q_t^\tau(z, x)$ for $\tau<t$.

**Acknowledgements**

Author would like to thank the anonymous referee for giving many valuable comments.

## References

[1] Beneš V. E. Exact finite — dimensional filters for certain diffusion with nonlinear drift. *Stochastic*, **5** (1981).

[2] Gedymin O. Private communications. 1985.

[3] Friedman A. Partial differential equations of parabolic type. Prentice-Hall, 1964.

[4] Kallianpur G., Striebel C. Estimation of stochastic systems: arbitrary system process with additive white noise observation errors. *Ann. Math. Stat.* **39** (1968).

[5] Aoki M. Optimal control and system theory in dynamic economic analysis. North-Holland 1976.

[6] Albouy M. La régulation économique dans l'entreprise. Dunod, Paris 1972.

*Received, February 1985.*

## Dokładne prodyktory skończenie wymiarowe dla pewnych procesów dyfuzji z nieliniowym dryftem

W pracy podano rozwiązanie problemu prognozy (exstrapolacji) dla pewnej klasy nieliniowych procesów dyfuzji z liniową obserwacją. Omówiono także zastosowanie otrzymanego rozwiązania oraz wskazano że nieznaczna modyfikacja dowodu zasadniczego twierdzenia prowadzi do rozwiązania problemu wygładzania (interpolacji).

## Точные конечномерные предикторы для некоторых диффузий с нелинейным дрифтом

Рассмотрено задачу оценивания решения $x_t$ стохастического уравнения

$$dx_t = f(x_t)\,dt + dw_t, \quad x_0 = x \text{ в момент } \tau, \ \tau > t.$$

Оценивание ведётся на основе возмущенного процесса наблюдения $y_t$. Дано достаточное условие того, чтобы функция условной плотности вероятности $x_t$ могла быть выражена при помощи конечного числа достаточных статистик.