Acta Sci. Math., 48 (1985), 13—23

Parameter estimation and Kalman filtering in noisy background

MÁTYÁS ARATÓ

Dedicated to Professor Károly Tandori on his 60th birthday

The parameter estimation (identification) problem will be discussed in the presence of additive coloured noise for multidimensional stochastic processes with continuous time. We shall investigate the case when the signal and noise both are Ornstein—Uhlenbeck processes. To obtain the main tool in such investigations, the Radon—Nikodym derivative, we use the method of Kalman filtering and the important remark that the Riccati equation can be solved explicitely. Some asymptotic results for the white noise case will be discussed (the reader may compare with BALAKRISHNAN [6]—[8]).

1. Introduction. Let us have an observed process $\xi(t)$ in the form

$$\xi(t) = \boldsymbol{\theta}(t, \alpha) + \boldsymbol{\varepsilon}(t, \boldsymbol{\beta}), \ 0 \leqq t \leqq T, \tag{1}$$

where $\boldsymbol{\alpha}, \boldsymbol{\beta}$ denote the unknown parameters, generally vector valued, which one wishes to estimate. The $\boldsymbol{\theta}(t, \boldsymbol{\alpha})$ process means the signal, while $\boldsymbol{\varepsilon}(t, \boldsymbol{\beta})$ models the error and both they are completely specified once $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are given. When the time, t, is continuous the basic tool in statistical theory is the Radon—Nikodym derivative of the probability measures induced by the processes $\xi(t)$ and $\boldsymbol{\theta}(t)$, with respect to a standard measure, e.g. the Wiener measure in the Gaussian case. It turns out that the derivatives even in the most simple cases have complicated form. For the stationary Gaussian process case the reader can find different methods in HAJEK [10], IBRAGIMOV and ROZANOV [13], PISARENKO [19]—[21], PISARENKO and ROZANOV [22], while for diffusional type Gaussian processes we may mention ARATÓ [2]—[4], BALAKRISHNAN [6]—[8], KUTOJANC [15], LIPSTSER and SHIRYAEV [16]. In these works mostly the scalar case was studied.

In this paper we assume that $\boldsymbol{\theta}(t, \boldsymbol{\alpha})$ and $\boldsymbol{\varepsilon}(t, \boldsymbol{\beta})$ both are first order multidimensional, autoregressive processes, the so-called Ornstein—Uhlenbeck processes, i.e.,

Received July 3, 1984.

they satisfy the stochastic differential equations

$$d\boldsymbol{\theta}(t) = -\boldsymbol{\alpha}\boldsymbol{\theta}(t)\,dt + \mathbf{c}_1^{1/2}\,d\mathbf{w}_1(t), \quad t \geqq 0, \tag{2}$$

$$d\boldsymbol{\varepsilon}(t) = -\boldsymbol{\beta}\boldsymbol{\varepsilon}(t)\,dt + \mathbf{c}_2^{1/2}\,d\mathbf{w}_2(t), \quad t \geqq 0,$$

where $\mathbf{w}_1(t)$, $\mathbf{w}_2(t)$ are standard, independent Wiener processes and $-\boldsymbol{\alpha}$, $-\boldsymbol{\beta}$ have eigenvalues with negative real parts, $\mathbf{c}_1 = \mathbf{c}_1^{1/2}(\mathbf{c}_1^{1/2})^*$ and $\mathbf{c}_2 = \mathbf{c}_2^{1/2}(\mathbf{c}_2^{1/2})^*$ are positive semidefinite symmetrical matrices. The $\boldsymbol{\varepsilon}(t)$ process is called *coloured noise*.

In the mathematical literature it was customary to take $\varepsilon(t)$ as a Wiener process. In the earlier engineering literature the white noise process was introduced for $\varepsilon(t)$ in a formal way as a stationary stochastic process with constant spectral density. For more rigorous treatment of the white noise case Balakrishnan introduced the Radon—Nikodym derivative of weak distributions, see [6]—[7].

In the present paper we are motivated by statistical considerations more than control theoretical, to develope some practically useful and computationally efficient closed-form expressions for estimators of the drift parameter α in (2). The expressions are developed for some limiting values in terms of the spectral characteristics of the processes for the purpose to obtain approximations in the white noise case too. First of all we remark that the Riccati equations of Kalman filtering can be solved explicitely and so one can get the Radon—Nikodym derivatives. With limiting it is possible to obtain those formulae derived earlier in the one dimensional case by Balakrishnan. The result that the filtering equations can be solved explicitely in the case of stochastic equations with constant coefficients may have other practical and theoretical consequences.

2. Explicite Kalman filtering. We assume that in (2) the processes $\boldsymbol{\theta}(t, \boldsymbol{\alpha})$ and $\boldsymbol{\varepsilon}(t, \boldsymbol{\beta})$ are p-dimensional and the Wiener processes $\mathbf{w}_1(t)$, $\mathbf{w}_2(t)$ are independent, p-dimensional and standard, i.e.,

$$\mathbf{E}\mathbf{w}_i(t) = 0, \quad \mathbf{E}\big(\mathbf{w}_i(t)\mathbf{w}_i(t)^*\big) = \mathbf{I}_p \cdot t, \quad \mathbf{w}_i(0) = 0, \quad i = 1, 2,$$

where asterisk means the transposed and $\mathbf{I}_p$ is the p-dimensional unit matrix. $\mathbf{w}_1(t)$ and $\mathbf{w}_2(t)$ are independent of $\boldsymbol{\theta}(0)$, $\boldsymbol{\varepsilon}(0)$ and further $\boldsymbol{\theta}(t)$ and $\boldsymbol{\varepsilon}(t)$ are stationary, which means that for the covariance functions $\big($assuming $\mathbf{E}\boldsymbol{\theta}(t) = \mathbf{E}\boldsymbol{\varepsilon}(t) = 0\big)$

$$B_\theta(t) = \mathbf{E}\boldsymbol{\theta}(t+s)\boldsymbol{\theta}^*(s), \quad B_\varepsilon(t) = \mathbf{E}\boldsymbol{\varepsilon}(t+s)\boldsymbol{\varepsilon}^*(s),$$

we have

$$\boldsymbol{\alpha}B_\theta(0) + B_\theta(0)\boldsymbol{\alpha}^* = \mathbf{c}_1, \quad B_\theta(t) = e^{-\boldsymbol{\alpha}|t|}B_\theta(0), \tag{3}$$

$$\boldsymbol{\beta}B_\varepsilon(0) + B_\varepsilon(0)\boldsymbol{\beta}^* = \mathbf{c}_2, \quad B_\varepsilon(t) = e^{-\boldsymbol{\beta}|t|}B_\varepsilon(0)$$

From (1) and (2) we get

$$d\boldsymbol{\theta}(t) = -\boldsymbol{\alpha}\boldsymbol{\theta}(t)\,dt + \mathbf{c}_1^{1/2}\,d\mathbf{w}_1(t), \tag{4}$$

$$\begin{aligned} d\boldsymbol{\xi}(t) &= d\boldsymbol{\theta}(t) + d\boldsymbol{\varepsilon}(t) = \\ &= -(\boldsymbol{\alpha}-\boldsymbol{\beta})\boldsymbol{\theta}(t)\,dt - \boldsymbol{\beta}\boldsymbol{\xi}(t)\,dt + \mathbf{c}_1^{1/2}\,d\mathbf{w}_1(t) + \mathbf{c}_2^{1/2}\,d\mathbf{w}_2(t), \end{aligned} \tag{5}$$

where $\boldsymbol{\xi}(t)$ is the observable, while $\boldsymbol{\theta}(t)$ is the unobservable component of the vector process $(\boldsymbol{\xi}^*(t), \boldsymbol{\theta}^*(t))$. The Kalman filtering equations for $\mathbf{m}(t)=E(\boldsymbol{\theta}(t)|\mathscr{F}_t^{\xi})$ and $\gamma(t)=\mathbf{E}(\boldsymbol{\theta}(t)-\mathbf{m}(t))(\boldsymbol{\theta}(t)-\mathbf{m}(t))^*$ are given by (see LIPTSER and SHIRYAEV [16] Th. 12.7)

$$\begin{aligned} d\mathbf{m}(t) &= -\boldsymbol{\alpha}\mathbf{m}(t)\,dt + [\mathbf{c}_1+\gamma(t)(\boldsymbol{\beta}-\boldsymbol{\alpha})^*][\mathbf{c}_1+\mathbf{c}_2]^{-1}\big[d\boldsymbol{\xi}(t) - \\ &\quad -((\boldsymbol{\beta}-\boldsymbol{\alpha})\mathbf{m}(t)-\boldsymbol{\beta}\boldsymbol{\xi}(t))\,dt\big], \end{aligned} \tag{6}$$

$$\begin{aligned} \dot\gamma(t) &= \frac{d\gamma(t)}{dt} = -\boldsymbol{\alpha}\gamma(t) - \gamma(t)\boldsymbol{\alpha}^* - [\mathbf{c}_1+\gamma(t)(\boldsymbol{\beta}-\boldsymbol{\alpha})^*][\mathbf{c}_1+\mathbf{c}_2]^{-1}\times \\ &\quad \times[\mathbf{c}_1+\gamma(t)(\boldsymbol{\beta}-\boldsymbol{\alpha})^*]^* + \mathbf{c}_1 = -[\boldsymbol{\alpha}+\mathbf{c}_1(\mathbf{c}_1+\mathbf{c}_2)^{-1}(\boldsymbol{\beta}-\boldsymbol{\alpha})]\gamma(t) - \\ &\quad -\gamma(t)[\boldsymbol{\alpha}+\mathbf{c}_1(\mathbf{c}_1+\mathbf{c}_2)^{-1}(\boldsymbol{\beta}-\boldsymbol{\alpha})]^* - \gamma(t)(\boldsymbol{\beta}-\boldsymbol{\alpha})^*(\mathbf{c}_1+\mathbf{c}_2)^{-1}(\boldsymbol{\beta}-\boldsymbol{\alpha})\gamma(t) - \\ &\quad -\mathbf{c}_1(\mathbf{c}_1+\mathbf{c}_2)^{-1}\mathbf{c}_1 + \mathbf{c}_1 = \\ &= \mathbf{a}\gamma(t) + \gamma(t)\mathbf{a}^* - \gamma(t)\mathbf{A}^*(\mathbf{c}_1+\mathbf{c}_2)^{-1}\mathbf{A}\gamma(t) + \mathbf{b}\mathbf{b}^*. \end{aligned} \tag{7}$$

The solution of the Riccati equation (7) may be given in the form (see ARATÓ [4], Lemma 2 in section 1.8)

$$\gamma(t) = e^{\tilde{\mathbf{a}}t}\Big[c_0 + \int_0^t e^{\tilde{\mathbf{a}}^*u}\mathbf{A}^*(\mathbf{c}_1+\mathbf{c}_2)^{-1}\mathbf{A}e^{\tilde{\mathbf{a}}u}\,du\Big]^{-1}e^{\tilde{\mathbf{a}}^*t} + \mathbf{c} \tag{8}$$

where $\mathbf{c}$ is the positive semidefinite solution of the "algebraic Riccati equation":

$$\mathbf{ac}+\mathbf{ca}^* - \mathbf{cA}^*(\mathbf{c}_1+\mathbf{c}_2)^{-1}\mathbf{Ac} + \mathbf{bb}^* = 0, \tag{9}$$

and the constants $\mathbf{c}_0, \tilde{\mathbf{a}}, \tilde{\mathbf{a}}^*$ are given by

$$\begin{gathered} \mathbf{c}_0^{-1} = \gamma(0) - \mathbf{c}, \\ \tilde{\mathbf{a}} = \mathbf{a} - \mathbf{cA}^*(\mathbf{c}_1+\mathbf{c}_2)^{-1}\mathbf{A}, \quad \tilde{\mathbf{a}}^* = \mathbf{a}^* - \mathbf{A}^*(\mathbf{c}_1+\mathbf{c}_2)^{-1}\mathbf{Ac}. \end{gathered} \tag{10}$$

In the scalar case we have the special form

$$\gamma(t) = e^{-\tilde{A}t}[c_0 + B(1-e^{-\tilde{A}t})/\tilde{A}]^{-1} + c, \tag{8'}$$

where from the stationarity

$$\begin{gathered} A = 2\frac{\alpha c_2+\beta c_1}{c_1+c_2}, \quad B = \frac{(\beta-\alpha)^2}{c_1+c_2}, \quad b = \frac{c_1c_2}{c_1+c_2}, \\ c = \frac{-A+\sqrt{A^2+4Bb}}{2B} \end{gathered} \tag{9'}$$

and

$$c_0^{-1} = \gamma(0) - c, \quad \gamma(0) = \frac{c_1 c_2}{2(\alpha c_2 + \beta c_1)}, \quad \tilde{A} = A + 2c\beta. \tag{10'}$$

From (6)—(10) by integration we can prove the following statement.

Theorem 1. *The conditional expectation of* $\boldsymbol{\theta}(t)$ *under condition* $\xi(s)$, $0 \leqq s \leqq t$, *given in* (4)—(5), *has the form*

$$\begin{aligned} \mathbf{E}(\boldsymbol{\theta}(t)|\mathscr{F}_t^\xi) = \mathbf{m}(t) &= \exp\Big\{-\int_0^t \mathbf{g}(u)\,du\Big\}\Big\{\mathbf{m}(0) + \int_0^t \exp\Big(\int_0^s \mathbf{g}(u)\,du\Big)[\mathbf{h}(s)\boldsymbol{\beta}\xi(s)\,ds + \\ &+ \mathbf{h}(s)\,d\xi(s)]\Big\} = \exp\Big\{-\int_0^t \mathbf{g}(u)\,du\Big\}\Big\{\mathbf{m}(0) + \exp\Big(\int_0^t \mathbf{g}(u)\,du\Big)\mathbf{h}(t)\xi(t) + \\ &+ \mathbf{h}(0)\xi(0) + \int_0^t \exp\Big(\int_0^s \mathbf{g}(u)\,du\Big)[\mathbf{h}(s)\boldsymbol{\beta} - \mathbf{h}'(s)]\xi(s)\,ds - \\ &- \int_0^t \mathbf{g}(s)\exp\Big(\int_0^s \mathbf{g}(u)\,du\Big)\mathbf{h}(s)\xi(s)\,ds\Big\}, \end{aligned} \tag{11}$$

where

$$\mathbf{g}(u) = \boldsymbol{\alpha} + \mathbf{h}(u)(\boldsymbol{\beta} - \boldsymbol{\alpha}),$$
$$\mathbf{h}(u) = [\mathbf{c}_1 + \gamma(u)(\boldsymbol{\beta} - \boldsymbol{\alpha})^*](\mathbf{c}_1 + \mathbf{c}_2)^{-1},$$

and $\gamma(t)$ *is given in* (8)—(10).

Specially in the scalar case we have

$$\begin{aligned} m(t) &= \exp(-h_1(t))\Big\{m(0) + (\beta-\alpha)^{-1}\exp(h_1(t))\Big[\frac{A}{2} + B\gamma(t) - \alpha\Big]\xi(t) + \\ &-(\beta-\alpha)^{-1}[A/2 + B\gamma(0) - \alpha]\xi(0) + (\beta-\alpha)^{-1}\int_0^t \exp(h_1(s))\Big[A(\beta-\alpha)/2 + \\ &-\beta\alpha - \frac{A^2}{4} - bB + B\gamma(s)(\beta - 2A + \alpha) - B^2\gamma^2(s)\Big]\xi(s)\,ds\Big\}, \end{aligned} \tag{11'}$$

where

$$m(0) = \frac{c_1\beta}{c_1\beta + c_2\alpha}\xi(0)$$

$$h_1(t) = \int_0^t [A/2 + B\gamma(s)]\,ds = (A/2 + Bc)t + \log\big(c_0 + B(1 - e^{-\tilde{A}t})/\tilde{A}\big)/c_0.$$

To calculate $\mathbf{m}(0) = \mathbf{E}(\boldsymbol{\theta}(0)|\xi(0))$ and $\gamma(0) = \operatorname{cov}(\boldsymbol{\theta}(0), \boldsymbol{\theta}(0)|\xi(0))$ one can use the Gaussianness and stationarity of the process $(\boldsymbol{\theta}(t), \xi(t))$. Let $\boldsymbol{\theta} = \boldsymbol{\theta}(0)$, $\xi = \xi(0)$

be a Gaussian random vector, then

$$m(0) = D_{\theta\xi} D_{\xi\xi}^{-1} \xi, \tag{12}$$
$$\gamma(0) = D_{\theta\theta} - D_{\theta\xi} D_{\xi\xi}^{-1} D_{\xi\theta},$$

where

$$\tilde{B}(0) = \begin{pmatrix} \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\theta}) & \operatorname{cov}(\boldsymbol{\theta}, \xi) \\ \operatorname{cov}(\xi, \boldsymbol{\theta}) & \operatorname{cov}(\xi, \xi) \end{pmatrix} = \begin{pmatrix} D_{\theta\theta} & D_{\theta\xi} \\ D_{\xi\theta} & D_{\xi\xi} \end{pmatrix}$$

is the joint covariance matrix. From stationarity we obtain

$$\tilde{\mathbf{A}} \tilde{B}(0) + \tilde{B}(0) \mathbf{A}^* = -\tilde{\mathbf{B}}_w, \tag{13}$$

where

$$\tilde{\mathbf{A}} = \begin{pmatrix} -\alpha & 0 \\ \beta - \alpha & -\beta \end{pmatrix}, \quad \tilde{\mathbf{B}}_w = \begin{pmatrix} \mathbf{c}_1 & \mathbf{c}_1^{1/2}(\mathbf{c}_1^{1/2} + \mathbf{c}_2^{1/2})^* \\ (\mathbf{c}_1^{1/2} + \mathbf{c}_2^{1/2})(\mathbf{c}_1^{1/2})^* & \mathbf{c}_1 + \mathbf{c}_2 \end{pmatrix},$$

and the unknown $\tilde{B}(0)$ may be gotten from the linear relations

$$\boldsymbol{\alpha} D_{\theta\theta} - D_{\theta\theta} \boldsymbol{\alpha}^* = \mathbf{c}_1,$$
$$\boldsymbol{\alpha} D_{\theta\xi} - D_{\theta\theta}(\boldsymbol{\beta} - \boldsymbol{\alpha})^* + D_{\xi\theta} \boldsymbol{\beta}^* = \mathbf{c}_1^{1/2}(\mathbf{c}_1^{1/2} + \mathbf{c}_1^{1/2})^*,$$
$$(\boldsymbol{\alpha} - \boldsymbol{\beta}) D_{\theta\xi} + \boldsymbol{\beta} D_{\theta\xi} + D_{\theta\xi}(\boldsymbol{\beta} - \boldsymbol{\alpha})^* + D_{\xi\xi} \boldsymbol{\beta}^* = (\mathbf{c}_1 + \mathbf{c}_2).$$

Explicite solution of (13) is not known in the general case.

To find different approximations for the maximum likelihood method one has to dicuss special cases and find simple forms for $\mathbf{m}(t)$ and its integral. Below we shall investigate two problems: a) the realization of $\xi(s)$ is given in $-\infty < s \leqq t$, b) the white noise approximation, when

$$\boldsymbol{\beta} \mathbf{c}_2^{-1} \boldsymbol{\beta}^* \to \boldsymbol{\beta}_0, \quad \text{if} \quad \boldsymbol{\beta} \to \infty, \tag{14}$$

where the convergence of matrix $\mathbf{A}$ is understood in the norm $\|\mathbf{A}\|^2 = \mathrm{Sp}\, \mathbf{A}\mathbf{A}^*$, and $\boldsymbol{\beta}$ has full rank, p.

a) Taking the observation interval $-T \leqq s \leqq t$, $(T > 0)$, we obtain from (12) that

$$\mathbf{m}_T(-T) = \mathbf{a}_1 \xi(-T), \quad \gamma_T(-T) = \mathbf{b}_1,$$

with constant matrices $\mathbf{a}_1$ and $\mathbf{b}_1$. Assuming that the matrix

$$\mathbf{g} = \left[\boldsymbol{\alpha} - (\mathbf{c}_1 + \mathbf{c}(\boldsymbol{\beta} - \boldsymbol{\alpha})^* (\mathbf{c}_1 + \mathbf{c}_2)^{-1} (\boldsymbol{\beta} - \boldsymbol{\alpha}))\right]$$

has eigenvalues with positive real parts from (8) and (11) one can obtain

$$\gamma_T(t) \to \mathbf{c}, \quad \text{if} \quad T \to \infty, \tag{15}$$
$$e^{-\int_{-T}^{t} g(u)\, du} = \exp\{-g \cdot (t + T)\} \to 0, \quad \text{if} \quad T \to \infty.$$

So we proved.

Theorem 2. *The limit function of $m_T(t)$, when $T\to\infty$, under the condition (15) has the from*

$$\tilde{\mathbf{m}}(t) = \mathbf{h}\cdot\xi(t) + \int_{-\infty}^{t} e^{\mathbf{g}(s-t)}[\mathbf{h}\cdot\boldsymbol{\beta}]\xi(s)\,ds - \int_{-\infty}^{t} \mathbf{g}e^{\mathbf{g}(s-t)}\mathbf{h}\xi(s)\,ds, \tag{16}$$

where

$$\mathbf{h} = [\mathbf{c}_1+\mathbf{c}(\boldsymbol{\beta}-\boldsymbol{\alpha})^*][\mathbf{c}_1+\mathbf{c}_2]^{-1}.$$

Remark. In the special case when $\alpha=\beta$ we have $g=\alpha$, $h=c_1[c_1+c_2]^{-1}$ and

$$\tilde{\mathbf{m}}(t) = c_1(c_1+c_2)^{-1}\xi(t),$$

as it could be expected.

b) If $\boldsymbol{\beta}\mathbf{c}_2^{-1}\boldsymbol{\beta}^*\to\boldsymbol{\beta}_0$, when $\boldsymbol{\beta}\to\infty$, and $\boldsymbol{\beta}_0$ is positive definite we obtain from (7)—(10)

$$\mathbf{b}\mathbf{b}^*\to\mathbf{c}_1,\quad \mathbf{a}\to-\boldsymbol{\alpha}, \tag{17}$$

$$\mathbf{A}^*(\mathbf{c}_1+\mathbf{c}_2)^{-1}\mathbf{A}\to\boldsymbol{\beta}_0,$$

further, $\tilde{\mathbf{c}}$ is the solution of the equation

$$-\boldsymbol{\alpha}\tilde{\mathbf{c}}-\tilde{\mathbf{c}}\boldsymbol{\alpha}^*-\tilde{\mathbf{c}}\boldsymbol{\beta}_0\mathbf{c}+\mathbf{c}_1 = 0 \tag{18}$$

and

$$\tilde{\mathbf{a}}\to-\boldsymbol{\alpha}-\tilde{\mathbf{c}}\boldsymbol{\beta}_0, \tag{19}$$

$$\mathbf{c}_0^{-1}\to\tilde{\mathbf{c}}_0^{-1} = \gamma(0)-\tilde{\mathbf{c}}.$$

In this case we have

Theorem 3. *In the "white noise" case approximation*

$$\gamma_\beta(t)\sim\tilde{\gamma}(t) = \exp[(-\boldsymbol{\alpha}-\tilde{\mathbf{c}}\boldsymbol{\beta}_0)t]\Big[\tilde{\mathbf{c}}_0+\int_0^t \exp[-(\boldsymbol{\alpha}+\tilde{\mathbf{c}}\boldsymbol{\beta}_0)^*u]\boldsymbol{\beta}_0\exp[-(\boldsymbol{\alpha}+ \tag{20}$$

$$+\tilde{\mathbf{c}}\boldsymbol{\beta}_0)u]\,du\Big]^{-1}\exp[(-\boldsymbol{\alpha}-\tilde{\mathbf{c}}\boldsymbol{\beta}_0)^*t]+\tilde{\mathbf{c}},$$

and

$$\mathbf{m}_\beta(t)\sim\tilde{\mathbf{m}}(t) = \exp\Big(-\boldsymbol{\alpha}t+\int_0^t\tilde{\gamma}(u)\boldsymbol{\beta}_0\,du\Big)\Big\{\tilde{\mathbf{m}}(0)+\int_0^t\exp\Big(-\boldsymbol{\alpha}s+ \tag{21}$$

$$+\int_0^s\tilde{\gamma}(u)\boldsymbol{\beta}_0\,du\Big)\tilde{\gamma}(s)\boldsymbol{\beta}_0\xi(s)\,ds\Big\},$$

when $\boldsymbol{\beta}\to\infty$, *and* $\boldsymbol{\beta}\mathbf{c}_2^{-1}\boldsymbol{\beta}^*\to\boldsymbol{\beta}_0$ *is positive definite.*

Remark. In the same way can be treated the case when the noise power/signal power, i.e., $B_\varepsilon(0)\cdot B_\theta^{-1}(0)$ tends to 0.

3. Parameter estimation. The parameters $\mathbf{c}_1$, $\mathbf{c}_2$ can be determined in two steps, if we want to handle them separately. Using the fact that $\xi(t)$ is a diffusional type process we obtain (with probability 1)

(22) $$\Sigma(\xi(t_i)-\xi(t_{i-1}))(\xi(t_i)-\xi(t_{i-1}))^* \to (\mathbf{c}_1+\mathbf{c}_2)T,$$
$$0 = t_0 \leqq t_1 \leqq \ldots \leqq t_n = T,$$

when $\max(t_i-t_{i-1}) \to 0$. On the other side in representation (6) the process

(23) $$d\tilde{\mathbf{w}}(t) = d\xi(t)-[(\beta-\alpha)\mathbf{m}(t)-\beta\xi(t)]\,dt$$

is a Wiener process with parameters

$$E\tilde{\mathbf{w}}(t) = 0, \quad E\tilde{\mathbf{w}}(t)\tilde{\mathbf{w}}^*(t) = (\mathbf{c}_1+\mathbf{c}_2)t.$$

This gives (see e.g. Th. 7.17 in [16]) that (with probability 1)

(24) $$\Sigma(\mathbf{m}(t_i)-\mathbf{m}(t_{i-1}))(\mathbf{m}(t_i)-\mathbf{m}(t_{i-1}))^* \to$$
$$\to \int_0^T [\mathbf{c}_1+((\beta-\alpha)\gamma(s))][\mathbf{c}_1+\mathbf{c}_2]^{-1}[c_1+\gamma(s)(\beta-\alpha)^*]\,ds,$$

if $\max(t_i-t_{i-1}) \to 0$. The last relation could be used if α and β were known.

To get separately $\mathbf{c}_1$ and $\mathbf{c}_2$ one can use (22), (24) and some preliminary estimation of α and β. For this purpose let us use that the covariance matrix function and the spectral density function have the form

(25) $$B_\xi(t) = \mathbf{E}\xi(t+s)\xi^*(s) = e^{-\beta|t|}B_\varepsilon(0)+e^{-\alpha|t|}B_\theta(0),$$
$$f_\xi(\lambda) = \frac{1}{2\pi}(i\lambda\mathbf{I}_p+\alpha)^{-1}\mathbf{c}_1[(-i\lambda\mathbf{I}_p+\alpha)^*]^{-1}+$$
$$+\frac{1}{2\pi}(i\lambda\mathbf{I}_p+\beta)^{-1}\mathbf{c}_2[(-i\lambda\mathbf{I}_p+\beta)^*]^{-1},$$

with condition (3).

Let $\hat{B}(t)$ denote the empirical covariance function

$$\hat{B}(t) = \frac{1}{T-t}\int_0^{T-t} \xi(s)\xi^*(s+t)\,ds,$$

then equating at the points t_1, t_2 to the theoretical values we have

(26) $$\hat{B}(t_1) = e^{\hat{\alpha}|t_1|}\cdot B_\theta(0)+e^{\hat{\beta}|t_1|}\cdot B_\varepsilon(0),$$
$$\hat{B}(t_2) = e^{\hat{\alpha}|t_2|}\cdot B_\theta(0)+e^{\hat{\beta}|t_2|}\cdot B_\varepsilon(0)$$

which give, together with (3) a system for the estimators $\hat{\alpha}$ and $\hat{\beta}$. We note here that these estimators are not efficient.

To improve estimators $\hat{\alpha}$ and $\hat{\beta}$, the solutions of system (26), one can use the sequential estimation, proposed by LIPTSER and SHIRYAEV [16] (chapter 12) in a modified form. E.g. let us assume that $\boldsymbol{\alpha}$ is a random vector variable with

$$\mathbf{E}\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}} \quad \text{and} \quad \gamma_\alpha(0) = \operatorname{cov}(\boldsymbol{\alpha}, \boldsymbol{\alpha}).$$

Further, from (6), with the approximation only in the diffusion coefficient $\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}$ and $\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}$ we get

$$d\mathbf{m}(t) = -\boldsymbol{\alpha}\mathbf{m}(t)\,dt+[\mathbf{c}_1+\gamma(t)(\hat{\boldsymbol{\beta}}-\hat{\boldsymbol{\alpha}})][\mathbf{c}_1+\mathbf{c}_2]^{-1}\,[d\boldsymbol{\xi}(t)- -((\hat{\boldsymbol{\beta}}-\hat{\boldsymbol{\alpha}})\mathbf{m}(t)-\hat{\boldsymbol{\beta}}\boldsymbol{\xi}(t))\,dt] \tag{27}$$

where $\gamma(t)$ is given by (8)—(9) with parameters $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\beta}}$. Assuming that $P(\boldsymbol{\alpha}<u|\mathbf{m}(0))$ is Gaussian and the fourth moment of $\boldsymbol{\alpha}$ exists from theorem 12.8 in [16] we obtain for

$$\tilde{\boldsymbol{\alpha}}(t) = \mathbf{E}(\boldsymbol{\alpha}|\mathscr{F}_t^m), \quad \tilde{\gamma}_\alpha(t) = \mathbf{E}((\boldsymbol{\alpha}(t)-\boldsymbol{\alpha})(\boldsymbol{\alpha}(t)-\boldsymbol{\alpha})^*|\mathscr{F}_t^m), \tag{28}$$

$$\tilde{\boldsymbol{\alpha}}(t) = \Big\{\mathbf{I}+\gamma_\alpha(0)\int_0^t \mathbf{m}^*(s)[\mathbf{c}_1+\gamma(s)(\hat{\boldsymbol{\beta}}-\hat{\boldsymbol{\alpha}})](\mathbf{c}_1+\mathbf{c}_2)^{-1}[\mathbf{c}_1+ +\gamma(s)(\hat{\boldsymbol{\beta}}-\hat{\boldsymbol{\alpha}})]^*\mathbf{m}(s)\,ds\Big\}^{-1}\Big\{\hat{\boldsymbol{\alpha}}+\gamma_\alpha(0)\int_0^t \mathbf{m}^*(s)[\mathbf{c}_1+\gamma(s)(\hat{\boldsymbol{\beta}}- -\hat{\boldsymbol{\alpha}})](\mathbf{c}_1+\mathbf{c}_2)^{-1}[\mathbf{c}_1+\gamma(s)(\hat{\boldsymbol{\beta}}-\hat{\boldsymbol{\alpha}})]^*\,d\mathbf{m}(s)\Big\},$$

(29)

$$\tilde{\gamma}_\alpha(t) = \Big\{\mathbf{I}+\gamma_\alpha(0)\int_0^t \mathbf{m}(s)[\mathbf{c}_1+\gamma(s)(\hat{\boldsymbol{\beta}}-\hat{\boldsymbol{\alpha}})](\mathbf{c}_1+\mathbf{c}_2)^{-1}[\mathbf{c}_1+\gamma(s)(\hat{\boldsymbol{\beta}}-\hat{\boldsymbol{\alpha}})]^*\mathbf{m}(s)ds\Big\}^{-1}\gamma_\alpha(0).$$

To improve $\hat{\boldsymbol{\beta}}$ one can use (5) with $\hat{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\alpha}}(t)$ and the same sequential procedure. This approximation may be compared with those which were proposed by LJUNG [17], see also HANNAN [11, 12], in the discrete time case. The derivation here seems more simple, but no optimality is proved.

Now let us return to the maximum likelihood estimators. It is known that generally no finite system of sufficient statistics exists if $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are unknown.

From a Girsanov's type theorem (see [16] Th. 7.19) and (6) we obtain for the Radon—Nikodym derivative (assuming $\boldsymbol{\xi}(0)=0$)

$$\frac{d\mathbf{P}_\xi}{d\mathbf{P}_{\tilde{w}}}(\boldsymbol{\xi}(t)) = \exp\Big\{-\int_0^t [(\boldsymbol{\alpha}-\boldsymbol{\beta})\mathbf{m}(s)+\boldsymbol{\beta}\boldsymbol{\xi}(s)]^*(\mathbf{c}_1+\mathbf{c}_2)^{-1}d\boldsymbol{\xi}(s)- -\frac{1}{2}\int_0^t [(\boldsymbol{\alpha}-\boldsymbol{\beta})\mathbf{m}(s)+\boldsymbol{\beta}\boldsymbol{\xi}(s)]^*(\mathbf{c}_1+\mathbf{c}_2)^{-1}[(\boldsymbol{\alpha}-\boldsymbol{\beta})\mathbf{m}(s)+\boldsymbol{\beta}\boldsymbol{\xi}(s)]\,ds\Big\}. \tag{30}$$

To find the distribution of the exponent in (30) in special cases has a very short history see ARATÓ and BENCZÚR [5], NOVIKOV [18], KONCZ [14].

Example. In the approximate "white noise" case (see case b) in §2.) we obtain for the Radon—Nikodym derivative (30) that it equals to

$$\exp\left\{-\frac{1}{2}\int_0^t [-\boldsymbol{\beta}_0\tilde{\mathbf{m}}(s)+\boldsymbol{\beta}_0\xi(s)]^*[-\boldsymbol{\beta}_0\tilde{\mathbf{m}}(s)]+\boldsymbol{\beta}_0\xi(s)]\,ds\right\}, \tag{31}$$

where $\tilde{\mathbf{m}}(s)$ and $\tilde{\gamma}(s)$ are given by (20) and (21), respectively. Neglecting in $\tilde{\mathbf{m}}(t)$ with the term $\tilde{\mathbf{m}}(0)$, i.e., assuming that

$$\tilde{\mathbf{m}}(t) = \int_0^t \exp\left\{-\boldsymbol{\alpha}(t-s)-\int_0^t \tilde{\gamma}(u)\boldsymbol{\beta}_0\,du\right\}\tilde{\gamma}(s)\boldsymbol{\beta}_0\xi(s)\,ds = \int_0^t f(t,s)\xi(s)\,ds, \tag{32}$$

and taking $\tilde{\gamma}(t)=\tilde{\mathbf{c}}$ we obtain for the log likelihood function

$$L(\boldsymbol{\alpha},\boldsymbol{\beta}_0) = -\frac{1}{2}\int_0^t \left[-\int_0^s \bar{e}^{(\boldsymbol{\alpha}+\tilde{\mathbf{c}}\boldsymbol{\beta}_0)(s-u)}\tilde{\mathbf{c}}\boldsymbol{\beta}_0\xi(u)\,du+\xi(s)\right]^*\boldsymbol{\beta}_0^*\boldsymbol{\beta}_0\times \tag{33}$$

$$\times\left[\xi(s)-\int_0^s \bar{e}^{(\boldsymbol{\alpha}+\tilde{\mathbf{c}}\boldsymbol{\beta}_0)(s-u)}\tilde{\mathbf{c}}\boldsymbol{\beta}_0\xi(u)\,du\right]ds.$$

From (33) the system of transcendental equations for the maximum likelihood estimators of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ is given in the following way

$$\frac{\partial L(\boldsymbol{\alpha},\boldsymbol{\beta}_0)}{\partial\boldsymbol{\alpha}} = 0,\quad \frac{\partial L(\boldsymbol{\alpha},\boldsymbol{\beta}_0)}{\partial\boldsymbol{\beta}} = 0.$$

Specially in the scalar case one gets

$$\tilde{\mathbf{c}} = \frac{-\alpha+\sqrt{\alpha^2+c_1\beta_0}}{\beta_0},\quad \tilde{\gamma}(0) = \frac{c_1}{2\alpha},$$

$$\tilde{c}_0^{-1} = \frac{c_1}{2\alpha}-\tilde{c},\quad \tilde{A}_0 = \alpha+\beta_0\tilde{c},$$

$$\tilde{\gamma}(t) = \bar{e}^{2\tilde{A}_0t}[\tilde{c}_0+\beta_0(1-\bar{e}^{2\tilde{A}_0t})/2\tilde{A}_0]^{-1}+\tilde{c},$$

and this gives with $\tilde{\gamma}(t)\approx\tilde{c}$

$$L(\alpha,\beta_0) = -\frac{1}{2}\beta_0^2\int_0^t \left[\xi(s)-\int_0^s \bar{e}^{A_0(s-u)}\xi(u)\tilde{c}\,du\right]^2 ds.$$

The maximum likelihood equations are the following

$$\beta_0 \int_0^t \Big[\xi(s) - \int_0^s \bar{e}^{\tilde{A}_0(s-u)} \tilde{c}\xi(u)\,du\Big]^2 ds + \beta_0^2 \int_0^t \Big[\xi(s) - \int_0^s \bar{e}^{\tilde{A}_0(s-u)} \tilde{c}\xi(u)\,du\Big] \Big\{ \int_0^s \exp(-\tilde{A}_0 s +$$

$$+A_0 u) c \frac{1}{2} c_1 (\alpha^2 + c_1\beta_0)^{-1/2} \xi(u)\,du - \beta_0^{-2} \int_0^s \bar{e}^{A_0(s-u)} \times \tag{34}$$

$$\times \Big[\frac{1}{2} c_1 (\alpha^2 + c_1\beta_0)^{-1/2} \beta_0 - (-\alpha + (\alpha^2 + c_1\beta_0)^{1/2})\Big] \xi(u)\,du \Big\} ds = 0$$

$$\int_0^t \Big[\xi(s) - \int_0^s \bar{e}^{\tilde{A}_0(s-u)} \tilde{c}\xi(u)\,du\Big] \Big[\int_0^s \bar{e}^{\tilde{A}_0(s-u)} \alpha (\alpha^2 + c_1\beta_0)^{-1/2} (-\alpha + \tag{35}$$

$$+(\alpha^2 + c_1\beta_0)^{1/2}) \xi(u)\,du\Big] ds = 0.$$

If β_0 is apriori given (known) and only α must be estimated from (35) we obtain the equation

$$\left(-\alpha + \sqrt{\alpha^2 + c_1\beta_0}\right) \int_0^t \left(\int_0^s \exp\left\{ \frac{\sqrt{\alpha^2 + c_1\beta_0}}{\beta_0} (u-s) \right\} \xi(u)\,du \right)^2 ds =$$

$$= \int_0^t \left(\int_0^s \exp\left\{ \frac{\sqrt{\alpha^2 + c_1\beta_0}}{\beta_0} (u-s) \right\} \xi(u)\,du \right) \xi(s)\,ds.$$

A similar equation, derived in another way, was given by PISARENKO [19], [21].

References

[1] M. ARATÓ, On the sufficient statistics of stationary Gaussian random processes, *Theory Probab. Appl.*, **6** (1961), 199—201. (Russian)

[2] M. ARATÓ, Exact formulas for density measure of elementary Gaussian processes, *Studia Sci. Math. Hungar.*, **5** (1970), 17—27. (Russian)

[3] M. ARATÓ, On parameter estimation in the presence of noise, *Theory Probab. Appl.*, **29** (1984), 599—604.

[4] M. ARATÓ, *Linear stochastic systems with constant coefficients*, Lecture Notes in Control and Information, **45**, Springer-Verlag (Berlin, 1982).

[5] M. ARATÓ and A. BENCZÚR, Distribution function of the damping parameter of stationary Gaussian processes, *Studia Sci. Math. Hungar.*, **5** (1970), 443—456. (Russian)

[6] A. V. BALAKRISHNAN, *Applied Functional Analysis*, Springer-Verlag (Berlin, 1976).

[7] A. V. BALAKRISHNAN, Parameter estimation in stochastic differential systems, *Developments in Statistics*, **1**, 1—32, (Academic Press, 1978).

[8] A. V. BALAKRISHNAN, *Stochastic filtering and Control*, Optimization Software (Los Angeles, 1981).

[9] K. O. DZHAPARIDZE, On the estimation of the spectral parameters of a Gaussian stationary process with rational spectral density, *Theory Probab. Appl.*, **15** (1970), 531—538.

[10] J. HAJEK, On linear statistical problems in stochastic processes, Czechoslovak Math. J., **12** (1962), 404—444.

[11] E. J. HANNAN, The convergence of some recursions, *Annals of Statistics*, **4** (1976), 1258—1270.

[12] E. J. HANNAN, Recursive estimation based on ARMA models, *Ann. Statist.*, **8** (1980), 762—777.

[13] I. A. IBRAHIMOV and J. A. ROZANOV, *Гауссовские случайные процессы*, Наука (Москва, 1970).

[14] K. KONCZ, Lineáris együtthatójú diffúziós folyamatok paramétereinek becsléséről, *Alkalmaz. Mat. Lapok*, to appear. (Hungarian)

[15] J. A. KUTOJANC, *Оценивание параметров случайных процессов*, Издательство АН (Ереван, 1980).

[16] R. LIPTSER and A. SHIRYAEV, *Статистика случайных процессов*, Наука (Москва, 1974).

[17] L. LJUNG, Strong convergence of a stochastic approximation algorithm, *Ann. Statist.*, **6** (1978), 680—696.

[18] A. A. NOVIKOV, On the estimation of parameters of diffusion processes, *Studia Sci. Math. Hungar.*, **6** (1971), 201—209. (Russian)

[19] V. F. PISARENKO, On the problem of discovering a random signal in noisy background, *Radiotechn. i Èlektron.*, **6** (1961), 515—528. (Russian)

[20] V. F. PISARENKO, On the estimation of parameters of stationary Gaussian process with a rational spectral density $|P(i\lambda)|^{-2}$, *Litovsk. Mat. Sb.*, **2** (1963), 159—167. (Russian)

[21] V. F. PISARENKO, On the computation of the likelihood ratio for Gaussian processes with a rational spectrum, *Theory Probab. Appl.*, **10** (1965), 299—303. (Russian)

[22] V. PISARENKO and J. ROZANOV, On certain problems for stationary processes leading to integral equations related to the Wiener—Hopf equation, *Problemy Peredači Informacii*, **14** (1963), 113—135. (Russian)

DEPARTMENT OF PROBABILITY AND STATISTICS
LORÁND EÖTVÖS UNIVERSITY
MÚZEUM KRT. 6—8.
1088 BUDAPEST, HUNGARY