STATISTICS OF CONDITIONALLY GAUSSIAN RANDOM SEQUENCES

R. SH. LIPTSER and A. N. SHIRYAYEV
INSTITUTE OF PROBLEMS CONTROL, STEKLOV INSTITUTE, MOSCOW

1. Statement of the problem

We are given, on a probability space $(\Omega, \mathscr{F}, P)$, a random sequence (θ_t, ξ_t),

$$\theta_t = \big(\theta_1(t), \cdots, \theta_k(t)\big), \ \xi_t = \big(\xi_1(t), \cdots, \xi_\ell(t)\big), \tag{1.1}$$

for $t = 0, 1, \cdots$, defined by the system of recursive equations

$$\begin{aligned} \theta_{t+1} &= a_0(t, \omega) + a_1(t, \omega)\theta_t + b(t, \omega)\Delta_1(t + 1), \\ \xi_{t+1} &= A_0(t, \omega) + A_1(t, \omega)\theta_t + B(t, \omega)\Delta_2(t + 1), \end{aligned} \tag{1.2}$$

where $\Delta_1(t)$ and $\Delta_2(t)$ are Gaussian and, in general, mutually dependent random vectors; while the vectors $a_i(t, \omega)$ and $A_i(t, \omega)$ and the matrices $b(t, \omega)$ and $B(t, \omega)$ are, for each t, $\mathscr{F}_t^\xi = \sigma\{\omega : \xi_0, \cdots, \xi_t\}$ measurable. The system (1.2) is to be solved for the initial conditions (θ_0, ξ_0), which are assumed to be independent of the processes $\Delta_1(t)$ and $\Delta_2(t)$, $t = 0, 1, \cdots$.

In the sequel, θ_t will be treated as a vector with unobservable components, and ξ_t as a vector with observable components. The statistical problems we wish to consider involve the construction of optimal (in the mean square sense) estimates of the unobservable process θ_t in terms of observations on the process ξ_t.

One can distinguish the following three basic problems of estimation, which will be called the problems of filtering, interpolation, and extrapolation.

Filtering. By filtering is understood the problem of estimating the unobservable vector θ_t by means of observations on the values of $\xi^t = (\xi_0, \xi_1, \cdots, \xi_t)$. We put $\Pi_\alpha(t) = P\{\theta_t \leqq \alpha \,|\, \mathscr{F}_t^\xi\}$ (where for vectors $x = (x_1, \cdots, x_k)$, $y = (y_1, \cdots, y_k)$ the inequality $x \leqq y$ is taken to mean that $x_i \leqq y_i$ for all $i = 1, \cdots, k$), let $m(t) = \mathbf{M}(\theta_t \,|\, \mathscr{F}_t^\xi)$, and

$$\gamma(t) = \operatorname{Cov}(\theta_t \,|\, \mathscr{F}_t^\xi) = \mathbf{M}\{(\theta_t - m(t))(\theta_t - m(t))^* \,|\, \mathscr{F}_t^\xi\}. \tag{1.3}$$

It is well known that $m(t) = \mathbf{M}(\theta_t \,|\, \mathscr{F}_t^\xi)$ is the optimal (in the mean square sense) estimate of θ_t by means of $\xi^t = (\xi_0, \xi_1, \cdots, \xi_t)$, and that

$$\operatorname{tr} \mathbf{M}\gamma(t) = \sum_{i=1}^{k} \mathbf{M}[\theta_i(t) - m_i(t)]^2, \tag{1.4}$$

where $m(t) = \big(m_1(t), \cdots, m_k(t)\big)$ is the error corresponding to this estimate.

In the general case the problem of finding the form of the distribution $\Pi_\alpha(t)$ and its parameters $m(t)$ and $\gamma(t)$ is a very difficult one. However, if one assumes in addition that the conditional distribution $\Pi_\alpha(0) = P\{\theta_0 \leqq \alpha \mid \xi_0\}$ is normal, $\mathcal{N}(m, \gamma)$, then it turns out that the *a posteriori* distribution $\Pi_\alpha(t) = P\{\theta_t \leqq \alpha \mid \mathcal{F}_t^\xi\}$ is also normal, for each $t = 1, 2, \cdots$.

(It is assumed that conditions are imposed upon the coefficients of the system (1.2) which guarantee the existence of second moments for θ_t and ξ_t, $t = 1, 2, \cdots$ (see Section 2).)

Moreover, under this assumption the distribution $P\{\theta_0 \leqq \alpha_0, \theta_1 \leqq \alpha_1, \cdots, \theta_t \leqq \alpha_t \mid \mathcal{F}_t^\xi\}$ is also normal, $t = 1, 2, \cdots$, which is the justification for calling the process (θ_t, ξ_t), $t = 0, 1, \cdots$, conditionally Gaussian.

For conditionally Gaussian processes the solution of the filtering problem reduces to finding the parameters $m(t)$ and $\gamma(t)$ of the normal distribution $P\{\theta_t \leqq \alpha \mid \mathcal{F}_t^\xi\}$. In Theorem 3.2 recursion relations are found for these parameters which generalize the familiar equations of Kalman and Bucy, [6], [7] obtained in a similar but substantially simpler framework (see equations (3.31) and (3.32)).

In recent years a considerable number of papers have been devoted to deriving recursion relations for $m(t)$ and $\gamma(t)$ under various assumptions on the coefficients of the system (1.2), particularly in connection with problems of optimal control. The appropriate bibliography can be found, for example, in Nahi's book [11] and in proceedings of conferences on automatic control [12]. A systematic investigation of the problem of estimating the unobservable components θ_t of a Markov process (θ_t, ξ_t) by means of certain values of the observable process ξ_t, $t = 0, 1, \cdots$, was carried out in [2], [3], many results of which will be used in this paper.

Interpolation. By interpolation is understood the construction of the best (in the mean square sense) estimates of the vector θ_t by means of observations of $\xi^\tau = (\xi_0, \cdots, \xi_\tau)$, where $\tau \geqq t$. We put $\Pi_\alpha(t, \tau) = P\{\theta_t \leqq \alpha \mid \mathcal{F}_\tau^\xi\}$,

$$\begin{aligned} m(t, \tau) &= \mathbf{M}(\theta_t \mid \mathcal{F}_\tau^\xi), \\ \gamma(t, \tau) &= \mathrm{Cov}\,(\theta_t \mid \mathcal{F}_\tau^\xi). \end{aligned} \tag{1.5}$$

In Sections 4 and 5, once again under the assumption that the conditional distribution $P(\theta_0 \leqq \alpha_0 \mid \xi_0)$ is normal, we will derive for the system (1.2) direct (in terms of τ for fixed $t \leqq \tau$) and inverse (in terms of t for fixed $\tau \geqq t$) interpolation equations for the parameters $m(t, \tau)$ and $\gamma(t, \tau)$ of the conditionally Gaussian distribution $\Pi_\alpha(t, \tau)$.

Extrapolation. By extrapolation is understood the problem of estimating the vectors θ_τ, ξ_τ by means of observations of $\xi^t = (\xi_0, \cdots, \xi_t)$, where $\tau > t$. As in the case of interpolation, one can derive equations in τ (direct equations) as well as equations in t (inverse equations) for the optimal extrapolative estimates

$$n_1(\tau, t) = \mathbf{M}(\theta_\tau \mid \mathcal{F}_t^\xi), \qquad n_2(\tau, t) = \mathbf{M}(\xi_\tau \mid \mathcal{F}_t^\xi). \tag{1.6}$$

In Section 7, we consider cases in which it is possible to obtain recursion relations for the estimates $n_1(\tau, t)$ and $n_2(\tau, t)$. Examples of the use of these relations are given in Section 8.

The contents of this paper, section by section, are as follows. Section 2 contains a theorem on normal correlation which we need. In Section 3, we state assumptions concerning the coefficients of system (1.2) and the initial conditions (θ_0, ξ_0). The normality of the conditional distribution $\Pi_\alpha(t)$ is proved for all $t = 1, 2, \cdots$, recursion relations are derived for $m(t)$ and $\gamma(t)$, and their properties are studied. Section 4 is devoted to the application of the recursion relations to the derivation of estimates for the unknown parameters. In Sections 5 and 6, respectively, we derive direct and inverse equations for interpolation. An application of them is given in Section 7. The derivation of the equations for extrapolation is carried out in Section 8.

In this paper only the case of discrete time is considered. For the case of continuous time, the analog of system (1.2) is Itô's system of equations

$$\begin{aligned} d\theta_t &= [a_0(t, \omega) + a_1(t, \omega)\theta_t]\, dt + b(t, \omega)\, dW_1(t), \\ d\xi_t &= [A_0(t, \omega) + A_1(t, \omega)\theta_t]\, dt + B(t, \omega)\, dW_2(t), \end{aligned} \tag{1.7}$$

where $W_1(t) = (W_{1,1}(t), \cdots, W_{1,k}(t))$ and $W_2(t) = (W_{2,1}(t), \cdots, W_{2,\ell}(t))$ are, in general, mutually dependent Wiener processes. Many results for this case can be obtained from the results discussed in this paper by a formal passage to the limit. Without dwelling on the details involved in the consideration of continuous time, we shall merely note here the papers [4], [5], [8], [9], [10], and [13], which are devoted to the study of the problems of filtering, interpolation, and extrapolation for systems of the type (1.7).

2. A theorem on normal correlation

THEOREM 2.1. *Let θ, ξ, where $\theta = (\theta_1, \cdots, \theta_k)$, $\xi = (\xi_1, \cdots, \xi_\ell)$, be jointly Gaussian vectors with means $\mathbf{M}\theta$, $\mathbf{M}\xi$, and covariances $d_{1,1} = \operatorname{Cov}(\theta, \theta)$, $d_{1,2} = \operatorname{Cov}(\theta, \xi)$, $d_{2,2} = \operatorname{Cov}(\xi, \xi)$.*

Then the vector of conditional expectations $\mathbf{M}(\theta \mid \xi)$ and the conditional covariance matrix $\operatorname{Cov}(\theta \mid \xi)$ are given by

$$\mathbf{M}(\theta \mid \xi) = \mathbf{M}\theta + d_{1,2} d_{2,2}^{+}(\xi - \mathbf{M}\xi), \tag{2.1}$$

$$\operatorname{Cov}(\theta \mid \xi) = d_{1,1} - d_{1,2} d_{2,2}^{+} d_{1,2}^{*}, \tag{2.2}$$

where $d_{2,2}^{+}$ is the pseudo-inverse of the matrix $d_{2,2}$.

PROOF. We turn first of all to the definition and properties of the pseudo-inverse matrix $d_{2,2}^{+}$.

If A is a symmetric nonnegative definite matrix, then, as is known (Gantmacher, [1], p. 34) there exists a (possibly nonunique) matrix T such that $T^*T = A$. The matrix

$$A^{+} = T^*(TT^*)^{-2}T \tag{2.3}$$

is called the pseudo-inverse of A. If $A = 0$, then by definition we put $A^+ = 0$.

It is not hard to verify that the pseudo-inverse matrix, as defined here, has the following properties:

$$\begin{gathered} AA^+A = A, \qquad A^+AA^+ = A^+, \\ (A^+)^* = (A^*)^+, \qquad A^+A(A^+)^* = (A^+)^*, \qquad (A^+)^+ = A, \end{gathered} \tag{2.4}$$

and there exists a matrix W such that $A^+ = A^*WA^*$.

It follows from the first and last properties that the pseudo-inverse matrix $A^+ = T^*(TT^*)^{-2}T$ is defined uniquely, independent of the method of representing A in the form $A = T^*T$.

If the symmetric matrix A is positive definite and hence nonsingular, the pseudo-inverse matrix A^+ coincides with the inverse A^{-1}, that is, $A^+ = A^{-1}$.

Thus, the pseudo-inverse $d_{2,2}^+$ can be defined as follows:

$$d_{2,2}^+ = \begin{cases} d_{2,2}^{-1}, & \text{if } d_{2,2} \text{ is nonsingular,} \\ T^*(TT^*)^{-2}T, & \text{if } d_{2,2} \text{ is singular,} \end{cases} \tag{2.5}$$

where $T^*T = d_{2,2}$.

We now proceed directly to the proof of the theorem. Putting $\eta = (\theta - \mathbf{M}\theta) - C(\xi - \mathbf{M}\xi)$, we shall show that the matrix $C = C(k \times \ell)$ can be so chosen that $\mathbf{M}\eta(\xi - \mathbf{M}\xi)^* = 0$. Then by virtue of the lack of correlation between, and hence the independence of the jointly Gaussian vectors η and $\xi - \mathbf{M}\xi$,

$$0 = \mathbf{M}\eta = \mathbf{M}(\eta \mid \xi) = [\mathbf{M}(\theta \mid \xi) - \mathbf{M}\theta] - C[\xi - \mathbf{M}\xi], \tag{2.6}$$

and therefore

$$\mathbf{M}(\theta \mid \xi) = \mathbf{M}\theta + C(\xi - \mathbf{M}\xi). \tag{2.7}$$

If the matrix $d_{2,2}$ is nonsingular, then to satisfy the equality $\mathbf{M}\eta(\xi - \mathbf{M}\xi)^* = 0$ it suffices to put $C = d_{1,2}d_{2,2}^{-1}$.

Suppose now that $d_{2,2}$ is singular and $d_{2,2} = T^*T$, where $T = T(r \times \ell)$, and r is the rank of the matrix $d_{2,2}$, $r \geqq 1$. (Note that in contrast with the matrix $d_{2,2} = T^*T$ which is an $\ell \times \ell$ matrix, TT^* is a nonsingular $r \times r$ matrix.) Obviously, there exists a Gaussian vector $\zeta = (\zeta_1, \cdots, \zeta_r)$ with independent components satisfying $\mathbf{M}\zeta_i = 0$, $\mathbf{M}\zeta_i^2 = 1$, such that the components of $\xi = (\xi_1, \cdots, \xi_\ell)$ are linear combinations of the ζ_i. Since $d_{2,2} = T^*T$, it follows that $\xi = \mathbf{M}\xi + T^*\zeta$.

Let $\tilde{d}_{1,2} = \mathbf{M}[(\theta - \mathbf{M}\theta)\zeta^*]$. Then $d_{1,2} = \tilde{d}_{1,2}T$; and putting $C = d_{1,2}d_{2,2}^+$, we find that by virtue of (2.5)

$$\begin{aligned} C\, d_{2,2} &= d_{1,2}T^*(TT^*)^{-2}T\, d_{2,2} \\ &= \tilde{d}_{1,2}TT^*(TT^*)^{-2}T\, d_{2,2} \\ &= \tilde{d}_{1,2}(TT^*)(TT^*)^{-2}(TT^*)T = \tilde{d}_{1,2}T = d_{1,2}. \end{aligned} \tag{2.8}$$

This completes the proof of formula (2.1).

To prove the representation (2.2), we note that since $\eta = (\theta - \mathbf{M}\theta) - d_{1,2}d_{2,2}^{+}(\xi - \mathbf{M}\xi)$, we have

$$
\begin{aligned}
\text{(2.9)}\quad \operatorname{Cov}(\theta\,|\,\xi) &= \mathbf{M}\{[\theta - \mathbf{M}(\theta\,|\,\xi)][\theta - \mathbf{M}(\theta\,|\,\xi)]^{*}\,|\,\xi\} \\
&= \mathbf{M}(\eta\eta^{*}\,|\,\xi) = \mathbf{M}(\eta\eta^{*}) \\
&= \mathbf{M}\{([\theta - \mathbf{M}\theta] - d_{1,2}d_{2,2}^{+}[\xi - \mathbf{M}\xi]) \\
&\qquad ([\theta - \mathbf{M}\theta] - d_{1,2}d_{2,2}^{+}[\xi - \mathbf{M}\xi])^{*}\} \\
&= d_{1,1} - d_{1,2}d_{2,2}^{+}d_{1,2}^{*} - d_{1,2}d_{2,2}^{+}d_{1,2}^{*} \\
&\qquad + d_{1,2}[d_{2,2}^{+}d_{2,2}d_{2,2}^{+}]\,d_{1,2}^{*}.
\end{aligned}
$$

But by the second property in (2.4) $d_{2,2}^{+}d_{2,2}d_{2,2}^{+} = d_{2,2}^{+}$; therefore, $\operatorname{Cov}(\theta\,|\,\xi) = d_{1,1} - d_{1,2}d_{2,2}^{+}d_{1,2}^{*}$.

COROLLARY 2.1. *It follows from* (2.1) *that the regression function of* $\theta = (\theta_1, \cdots, \theta_k)$ *on* $\xi = (\xi_1, \cdots, \xi_\ell)$, *that is,* $\mathbf{M}(\theta\,|\,\xi_1, \cdots, \xi_\ell)$, *is a linear function of* $\xi_1, \cdots, \xi_\ell$, *and* $\operatorname{Cov}(\theta\,|\,\xi)$ *does not depend upon the value of* ξ.

COROLLARY 2.2. *If* $k = \ell = 1$ *and* $\mathbf{D}\xi > 0$, *then*

$$\text{(2.10)}\qquad \mathbf{M}(\theta\,|\,\xi) = \mathbf{M}\theta + \frac{\operatorname{Cov}(\theta, \xi)}{\mathbf{D}\xi}(\xi - \mathbf{M}\xi),$$

$$\text{(2.11)}\qquad \mathbf{D}(\theta\,|\,\xi) = \mathbf{D}\theta - \frac{\operatorname{Cov}^2(\theta, \xi)}{\mathbf{D}\xi}.$$

Putting $\sigma_\theta = +\sqrt{\mathbf{D}\theta}$, $\sigma_\xi = +\sqrt{\mathbf{D}\xi}$ and introducing the correlation coefficient

$$\text{(2.12)}\qquad \rho = \frac{\operatorname{Cov}(\theta, \xi)}{\sigma_\theta\sigma_\xi},$$

formulas (2.10) and (2.11) can also be expressed in the following form, which is frequently used in statistics:

$$\text{(2.13)}\qquad \mathbf{M}(\theta\,|\,\xi) = \mathbf{M}\theta + \rho\frac{\sigma_\theta}{\sigma_\xi}(\xi - \mathbf{M}\xi),$$

$$\text{(2.14)}\qquad \mathbf{D}(\theta\,|\,\xi) = \sigma_\theta^2(1 - \rho^2).$$

COROLLARY 2.3. *If* $\theta = b_1\varepsilon_1 + b_2\varepsilon_2$, *and* $\xi = B_1\varepsilon_1 + B_2\varepsilon_2$, *where* $\varepsilon_1, \varepsilon_2$ *are independent Gaussian random variables, with* $\mathbf{M}\varepsilon_i = 0$, $\mathbf{D}\varepsilon_i = 1$, $i = 1, 2$ *and* $B_1^2 + B_2^2 > 0$, *then*

$$
\text{(2.15)}\qquad
\begin{aligned}
\mathbf{M}(\theta\,|\,\xi) &= \frac{b_1B_1 + b_2B_2}{B_1^2 + B_2^2}\,\xi, \\
\mathbf{D}(\theta\,|\,\xi) &= \frac{(B_1b_2 - b_1B_2)}{B_1^2 + B_2^2}.
\end{aligned}
$$

REMARK 2.1. Let (θ, ξ), $\theta = (\theta_1, \cdots, \theta_k)$, $\xi = (\xi_1, \cdots, \xi_\ell)$ be a random vector and let $\mathscr{G}$ be some σ-subalgebra of $\mathscr{F}$, $\mathscr{G} \subseteq \mathscr{F}$. Let us assume that the conditional (with respect to $\mathscr{G}$) distribution of the vector (θ, ξ) is Gaussian with means $\mathbf{M}(\theta\,|\,\mathscr{G})$, $\mathbf{M}(\xi\,|\,\mathscr{G})$ and covariances $d_{1,1} = \operatorname{Cov}(\theta, \theta\,|\,\mathscr{G})$, $d_{1,2} = \operatorname{Cov}$

$(\theta, \xi \mid \mathscr{G})$, and $d_{2,2} = \operatorname{Cov}(\xi, \xi \mid \mathscr{G})$. Then the vector of conditional expectations $\mathbf{M}(\theta \mid \xi, \mathscr{G})$ and the conditional covariance matrix $\operatorname{Cov}(\theta \mid \xi, \mathscr{G})$ are given by

$$\begin{aligned} \mathbf{M}(\theta \mid \xi, \mathscr{G}) &= \mathbf{M}(\theta \mid \mathscr{G}) + d_{1,2} d_{2,2}^{+}[\xi - \mathbf{M}(\xi \mid \mathscr{G})], \\ \operatorname{Cov}(\theta \mid \xi, \mathscr{G}) &= d_{1,1} - d_{1,2} d_{2,2}^{+} d_{1,2}^{*}. \end{aligned} \tag{2.16}$$

We shall use this result, whose proof is like that of (2.1) and (2.2), repeatedly in the sequel.

3. Recursion relations for filtering in the conditional Gaussian case

Instead of the system (1.2), which describes the evolution of the process (θ_t, ξ_t), $t = 0, 1, \cdots$, we shall consider the equivalent system

$$\begin{aligned} \theta_{t+1} &= a_0(t, \omega) + a_1(t, \omega)\theta_t + b_1(t, \omega)\varepsilon_1(t+1) + b_2(t, \omega)\varepsilon_2(t+1), \\ \xi_{t+1} &= A_0(t, \omega) + A_1(t, \omega)\theta_t + B_1(t, \omega)\varepsilon_1(t+1) + B_2(t, \omega)\varepsilon_2(t+1), \end{aligned} \tag{3.1}$$

where

$$\varepsilon_1(t) = (\varepsilon_{1,1}(t), \cdots, \varepsilon_{1,k}(t)), \qquad \varepsilon_2(t) = (\varepsilon_{2,1}(t), \cdots, \varepsilon_{2\ell}(t)) \tag{3.2}$$

are independent Gaussian random variables with $\mathbf{M}\varepsilon_{i,j}(t) = 0$ and

$$\mathbf{M}\varepsilon_{i_1,j_1}(s)\varepsilon_{i_2,j_2}(t) = \delta(i_1, i_2)\delta(j_1, j_2)\delta(t, s), \tag{3.3}$$

where $\delta(x, y)$ is the Kronecker symbol:

$$\delta(x, y) = \begin{cases} 1, & x = y, \\ 0, & x \neq y. \end{cases} \tag{3.4}$$

The vector functions

$$\begin{aligned} a_0(t, \omega) &= (a_{0,1}(t, \omega), \cdots, a_{0,k}(t, \omega)), \\ A_0(t, \omega) &= (A_{0,1}(t, \omega), \cdots, A_{0,\ell}(t, \omega)), \end{aligned} \tag{3.5}$$

and the matrices

$$\begin{aligned} &b_1(t, \omega) = \|b_{i,j}^{(1)}(t, \omega)\|, \quad b_2(t, \omega) = \|b_{i,j}^{(2)}(t, \omega)\|, \quad B_1(t, \omega) = \|B_{i,j}^{(1)}(t, \omega)\|, \\ &B_2(t, \omega) = \|B_{i,j}^{(2)}(t, \omega)\|, \quad a_1(t, \omega) = \|a_{i,j}^{(1)}(t, \omega)\|, \quad A_1(t, \omega) = \|A_{i,j}^{(1)}(t, \omega)\|, \end{aligned} \tag{3.6}$$

in (3.1) have orders, respectively, $k \times k$, $k \times \ell$, $\ell \times k$, $\ell \times \ell$, $k \times k$ and $\ell \times k$.

The functions $a_{0,i}(t, \omega)$, $A_{0,j}(t, \omega)$, $a_{i,j}^{(1)}(t, \omega)$, $A_{i,j}^{(1)}(t, \omega)$, $b_{i,j}^{(1)}(t, \omega)$, $b_{i,j}^{(2)}(t, \omega)$, $B_{i,j}^{(1)}(t, \omega)$ and $B_{i,j}^{(2)}(t, \omega)$ are assumed, for each value of t, to be $\mathscr{F}_t^\xi$ measurable, where $\mathscr{F}_t^\xi = \sigma\{\omega: \xi_0, \cdots, \xi_t\}$ is the σ-algebra generated by the random variables $\xi_0, \cdots, \xi_t$. For notational simplicity, we will omit the symbol ω, and sometimes the symbol t, when we write down these functions.

The system (3.1) is to be solved under the initial conditions (θ_0, ξ_0), where the random vector (θ_0, ξ_0) is assumed to be independent of the processes $\varepsilon_1(t)$, $\varepsilon_2(t)$, $t = 0, 1, \cdots$.

Throughout this paper, as concerns the coefficients in the system (3.1) and the distributions of the initial values (θ_0, ξ_0), we assume the following:

(a) if $g(t, \omega)$ is any one of the functions $a_{0,i}$, $A_{0,j}$, $b_{i,j}^{(1)}$, $b_{i,j}^{(2)}$, $B_{i,j}^{(1)}$, $B_{i,j}^{(2)}$, then $\mathbf{M}|g(t, \omega)|^2 < \infty$;

(b) with probability 1, $|a_{i,j}^{(1)}(t, \omega)| \leqq C$, $|A_{i,j}^{(1)}(t, \omega)| \leqq C$;

(c) $\mathbf{M}(\|\theta_0\|^2 + \|\xi_0\|^2) < \infty$, where $\|x\|^2 = \Sigma_i x_i^2$;

(d) the conditional distribution $\Pi_\alpha(0) = \mathbf{P}(\theta_0 \leqq \alpha \mid \xi_0)$ is, with probability 1, normal with parameters m and γ (which depend upon ξ_0).

It follows from assumptions (a) through (c) that $\mathbf{M}(\|\theta_t\|^2 + \|\xi_t\|^2) < \infty$ for every $t < \infty$.

THEOREM 3.1. *Suppose that the assumptions* (a) *through* (d) *are fulfilled. Then for every $t > 0$ the conditional distribution $\Pi_\alpha(t)$ is* (*with probability* 1) *normal.*

PROOF. We will carry out the proof by induction. Let us assume that the distribution $\Pi_\alpha(t)$ is normal, $\mathcal{N}(m(t), \gamma(t))$. We put

$$\Pi_{\beta x, \alpha}(\tau, t) = P\{\theta_\tau \leqq \beta, \xi_\tau \leqq x \mid \mathcal{F}_t^\xi, \theta_t = \alpha\}. \tag{3.7}$$

By virtue of (3.1), the distribution $\Pi_{\beta x, \alpha}(t+1, t)$ is Gaussian, with the vector of mathematical expectations

$$\mathcal{A}_0 + \mathcal{A}_1\alpha = \begin{pmatrix} a_0 + a_1\alpha \\ A_0 + A_1\alpha \end{pmatrix}, \tag{3.8}$$

and mathematical covariance

$$\mathcal{B} = \begin{pmatrix} b \circ b & b \circ B \\ (b \circ B)^* & B \circ B \end{pmatrix}, \tag{3.9}$$

where

$$b \circ b = b_1 b_1^* + b_2 b_2^*, \qquad b \circ B = b_1 B_1^* + b_2 B_2^*, \qquad B \circ B = B_1 B_1^* + B_2 B_2^*. \tag{3.10}$$

Let $v_t = (\theta_t, \xi_t)$ and $z = (z_1, \cdots, z_{k+\ell})$. Then the conditional characteristic function of the vector v_{t+1} is given by

$$\mathbf{M}[\exp\{i\langle z, v_{t+1}\rangle\} \mid \mathcal{F}_t^\xi, \theta_t] = \exp\{i\langle z, \mathcal{A}_0 + \mathcal{A}_1\theta_t\rangle - \tfrac{1}{2}\langle \mathcal{B}z, z\rangle\}; \tag{3.11}$$

therefore

$$\mathbf{M}[\exp\{i\langle z, v_{t+1}\rangle\} \mid \mathcal{F}_t^\xi] = \exp\{i\langle z, \mathcal{A}_0\rangle - \tfrac{1}{2}\langle \mathcal{B}z, z\rangle\}\mathbf{M}[\exp\{i\langle z, \mathcal{A}_1\theta_t\rangle\} \mid \mathcal{F}_t^\xi]. \tag{3.12}$$

By the induction hypothesis $\Pi_\alpha(t) \sim \mathcal{N}(m(t), \gamma(t))$. Therefore,

$$\mathbf{M}[\exp\{i\langle z, \mathcal{A}_1\theta_t\rangle\} \mid \mathcal{F}_t^\xi] = \exp\{i\langle z, \mathcal{A}_1 m(t)\rangle - \tfrac{1}{2}\langle \mathcal{A}_1\gamma(t)\mathcal{A}_1^* z, z\rangle\} \tag{3.13}$$

$$\mathbf{M}[\exp\{i\langle z, v_{t+1}\rangle\} \mid \mathcal{F}_t^\xi] = \exp[i\langle z, \mathcal{A}_0 + \mathcal{A}_1 m(t)\rangle] \cdot \exp[-\tfrac{1}{2}\langle \mathcal{B}z, z\rangle - \tfrac{1}{2}\langle \mathcal{A}_1\gamma(t)\mathcal{A}_1^* z, z\rangle].$$

Thus, each of the conditional distributions

$$(3.14)\qquad P\{\theta_{t+1} \leqq \beta, \xi_{t+1} \leqq x \mid \mathscr{F}_t^\xi\}, \quad P\{\theta_{t+1} \leqq \beta \mid \mathscr{F}_t^\xi\}, \quad P\{\xi_{t+1} \leqq x \mid \mathscr{F}_t^\xi\}$$

is Gaussian. Further let

$$(3.15)\qquad \eta = [\theta_{t+1} - M(\theta_{t+1} \mid \mathscr{F}_t^\xi)] - C[\xi_{t+1} - M(\xi_{t+1} \mid \mathscr{F}_t^\xi)].$$

We may find an $\mathscr{F}_t^\xi$-measurable matrix C (see the demonstration of Theorem 1) such that

$$(3.16)\qquad M\{\eta[\xi_{t+1} - M(\xi_{t+1} \mid \mathscr{F}_t^\xi)]^* \mid \mathscr{F}_t^\xi\} = 0 \quad \text{(a.s.)}$$

From this we obtain that the vectors η and ξ_{t+1} are conditionally independent (under the conditions $\mathscr{F}_t^\xi$) and

$$\begin{aligned}(3.17)\qquad & M[\exp i\langle z, \theta_{t+1}\rangle \mid \mathscr{F}_t^\xi, \xi_{t+1}] \\ &= M[\exp (i\langle z, M(\theta_{t+1} \mid \mathscr{F}_t^\xi) \\ &\quad + C(\xi_{t+1} - M(\xi_{t+1} \mid \mathscr{F}_t^\xi)\rangle + i\langle z, \eta\rangle) \mid \mathscr{F}_t^\xi, \xi_{t+1}] \\ &= \exp [i\langle z, M(\theta_{t+1} \mid \mathscr{F}_t^\xi) \\ &\quad + C(\xi_{t+1} - M(\xi_{t+1} \mid \mathscr{F}_t^\xi))\rangle] M[\exp i\langle z, \eta\rangle \mid \mathscr{F}_t^\xi, \xi_{t+1}] \\ &= \exp [i\langle z, M(\theta_{t+1} \mid \mathscr{F}_t^\xi) \\ &\quad + C(\xi_{t+1} - M(\xi_{t+1} \mid \mathscr{F}_t^\xi))\rangle] M[\exp i\langle z, \eta\rangle \mid \mathscr{F}_t^\xi] \quad \text{(a.s.)}\end{aligned}$$

But $P(\eta \leqq y \mid \mathscr{F}_t^\xi)$ is Gaussian and therefore the distribution $P(\theta_{t+1} \leqq y \mid \mathscr{F}_t^\xi)$ is Gaussian also.

The parameters of this distribution can be found from the recursive equations (3.18) and (3.19) which follow.

THEOREM 3.2. *Suppose that assumptions* (a) *through* (d) *are fulfilled. Then the parameters $m(t)$ and $\gamma(t)$ of the normal distribution $\Pi_\alpha(t)$ can be determined from the recursive equations*

$$\begin{aligned}(3.18)\qquad m(t+1) = {} & [a_0(t) + a_1(t)m(t)] + [b \circ B(t) + a_1(t)\gamma(t)A_1^*(t)] \\ & \cdot [B \circ B(t) + A_1(t)\gamma(t)A_1^*(t)]^+ [\xi_{t+1} - A_0(t) - A_1(t)m(t)],\end{aligned}$$

$$\begin{aligned}(3.19)\qquad \gamma(t+1) = {} & [a_1(t)\gamma(t)a_1^*(t) + b \circ b(t)] \\ & - [b \circ B(t) + a_1(t)\gamma(t)A_1^*(t)][B \circ B(t) + A_1(t)\gamma(t)A_1^*(t)]^+ \\ & \cdot [b \circ B(t) + a_1(t)\gamma(t)A_1^*(t)]^*\end{aligned}$$

with the initial conditions $m(0) = m$, $\gamma(0) = \gamma$.

PROOF. Let us first determine the parameters of the distribution

$$(3.20)\qquad \Pi_{\beta x}(t+1, t) = P\{\theta_{t+1} \leqq \beta, \xi_{t+1} \leqq x \mid \mathscr{F}_t^\xi\}.$$

Since

$$(3.21)\qquad \Pi_{\beta x}(t+1, t) = \mathbf{M}(\Pi_{\beta x, \theta_t}(t+1, t) \mid \mathscr{F}_t^\xi),$$

by virtue of (3.8) we have

$$\mathbf{M}(\theta_{t+1}\,|\,\mathscr{F}_t^\xi) = a_0(t) + a_1(t)m(t),$$
$$\mathbf{M}(\xi_{t+1}\,|\,\mathscr{F}_t^\xi) = A_0(t) + A_1(t)m(t). \tag{3.22}$$

To find the covariance matrix, we use the fact that according to (3.1) and (3.8),

$$\theta_{t+1} - \mathbf{M}(\theta_{t+1}\,|\,\mathscr{F}_t^\xi) = a_1(t)[\theta_t - m(t)] + b_1(t)\varepsilon_1(t+1) + b_2(t)\varepsilon_2(t+1),$$
$$\xi_{t+1} - \mathbf{M}(\xi_{t+1}\,|\,\mathscr{F}_t^\xi) = A_1(t)[\theta_t - m(t)] + B_1(t)\varepsilon_1(t+1) + B_2(t)\varepsilon_2(t+1). \tag{3.23}$$

Then

$$d_{1,1} = \operatorname{Cov}(\theta_{t+1}\,|\,\mathscr{F}_t^\xi) = a_1(t)\gamma(t)a_1^*(t) + b\circ b(t),$$
$$d_{1,2} = \operatorname{Cov}(\theta_{t+1}, \xi_{t+1}\,|\,\mathscr{F}_t^\xi) = a_1(t)\gamma(t)A_1^*(t) + b\circ B(t), \tag{3.24}$$
$$d_{2,2} = \operatorname{Cov}(\xi_{t+1}\,|\,\mathscr{F}_t^\xi) = A_1(t)\gamma(t)A_1^*(t) + B\circ B(t).$$

Since under conditioning by $\mathscr{F}_t^\xi$ the distribution of the vector $(\theta_{t+1}, \xi_{t+1})$ is normal, we have by virtue of the remark following Theorem 2.1,

$$\mathbf{M}(\theta_{t+1}\,|\,\mathscr{F}_t^\xi, \xi_{t+1}) = \mathbf{M}(\theta_{t+1}\,|\,\mathscr{F}_t^\xi) + d_{1,2}d_{2,2}^+[\xi_{t+1} - \mathbf{M}(\xi_{t+1}\,|\,\mathscr{F}_t^\xi)] \tag{3.25}$$

and

$$\operatorname{Cov}(\theta_{t+1}\,|\,\mathscr{F}_t^\xi, \xi_{t+1}) = d_{1,1} - d_{1,2}d_{2,2}^+d_{1,2}^*. \tag{3.26}$$

Inserting the expressions for $\mathbf{M}(\theta_{t+1}\,|\,\mathscr{F}_t^\xi)$, $\mathbf{M}(\xi_{t+1}\,|\,\mathscr{F}_t^\xi)$, and $d_{1,1}$, $d_{1,2}$, $d_{2,2}$ in the right sides of these formulae, we obtain (3.18) and (3.19).

COROLLARY 3.1. Let $a_0(t, \omega) = a_0(t) + a_2(t)\xi_t$, $A_0(t, \omega) = A_0(t) + A_2(t)\xi_t$, $a_1(t, \omega) = a_1(t)$, $A_1(t, \omega) = A_1(t)$, $b_i(t, \omega) = b_i(t)$, $B_i(t, \omega) = B_i(t)$, $i = 1, 2$, where the functions $a_j(t)$, $A_j(t)$, $b_i(t)$, $B_i(t)$, for $j = 0, 1, 2$, and $i = 1, 2$, are functions of t only (not depending upon ω). Then if the vector (θ_0, ξ_0) is Gaussian, the process (θ_t, ξ_t), $t = 0, 1, \cdots$, will also be Gaussian. In this case, the covariance $\gamma(t)$ does not depend upon "chance", and thus $\operatorname{tr}\gamma(t)$ yields the mean square error in filtering the vector θ_t by means of observations on $(\xi_0, \cdots, \xi_t)$.

COROLLARY 3.2. Suppose that the sequence (θ_t, ξ_t), where $\theta_t = (\theta_1(t), \cdots, \theta_k(t))$, $(\xi_1(t), \cdots, \xi_\ell(t))$, $t = 0, 1, \cdots$, satisfies the system of equations

$$\theta_{t+1} = a_0(t, \omega) + a_1(t, \omega)\theta_t + b_1(t, \omega)\varepsilon_1(t+1) + b_2(t, \omega)\varepsilon_2(t+1),$$
$$\xi_t = \tilde{A}_0(t-1, \omega) + \tilde{A}_1(t-1, \omega)\theta_t + \tilde{B}_1(t-1, \omega)\varepsilon_1(t) + \tilde{B}_2(t-1, \omega)\varepsilon_2(t), \qquad t \geqq 1, \tag{3.27}$$

where $a_0(t, \omega)$ and $\tilde{A}_0(t-1, \omega)$ as well as the coefficients of θ_t, $\varepsilon_1(t)$, $\varepsilon_2(t)$, $\varepsilon_1(t+1)$, and $\varepsilon_2(t+1)$ are the same as in (3.1), and $\mathbf{P}\{\theta_1 \leqq \alpha\,|\,\mathscr{F}_1^\xi\} \sim \mathscr{N}(m_1, \gamma_1)$.

Although this system of equations for (θ_t, ξ_t), $t = 0, 1, \cdots$, does not formally fit into the scheme (3.1), nevertheless in seeking equations for $m(t) =$

$\mathbf{M}(\theta_t|\mathscr{F}_t^\xi)$ and $\gamma(t) = \operatorname{Cov}(\theta_t|\mathscr{F}_t^\xi)$, one can make use of results obtained for the system (3.1).

In fact, we find from (3.27) that

$$\begin{aligned}\xi_{t+1} = \tilde{A}_0(t,\omega) + \tilde{A}_1(t,\omega)[&a_0(t,\omega) + a_1(t,\omega)\theta_t \\ &+ b_1(t,\omega)\varepsilon_1(t+1) + b_2(t,\omega)\varepsilon_2(t+1)] \\ &+ \tilde{B}_1(t,\omega)\varepsilon_1(t+1) + \tilde{B}_2(t,\omega)\varepsilon_2(t+1).\end{aligned} \tag{3.28}$$

Putting

$$\begin{aligned}A_0(t,\omega) &= \tilde{A}_0(t,\omega) + \tilde{A}_1(t,\omega)a_0(t,\omega), \\ A_1(t,\omega) &= \tilde{A}_1(t,\omega)a_1(t,\omega), \\ B_1(t,\omega) &= \tilde{A}_1(t,\omega)b_1(t,\omega) + \tilde{B}_1(t,\omega), \\ B_2(t,\omega) &= \tilde{A}_1(t,\omega)b_2(t,\omega) + \tilde{B}_2(t,\omega),\end{aligned} \tag{3.29}$$

we find that (θ_t, ξ_t) satisfies the system (3.1), and consequently $m(t)$ and $\gamma(t)$ satisfy the equations (3.18), (3.19).

COROLLARY 3.3. (Kalman-Bucy filter). Suppose that the Gaussian sequence (θ_t, ξ_t), $\theta_t = (\theta_1(t), \cdots, \theta_k(t))$, $\xi_t = (\xi_1(t), \cdots, \xi_\ell(t))$, for $t = 0, 1, \cdots$, satisfies the system of equations

$$\begin{aligned}\theta_{t+1} &= a_0(t) + a_1(t)\theta_t + b_1(t)\varepsilon_1(t+1) + b_2(t)\varepsilon_2(t+1), \\ \xi_t &= A_0(t) + A_1(t)\theta_t + B_1(t)\varepsilon_1(t) + B_2(t)\varepsilon_2(t),\end{aligned} \tag{3.30}$$

where the coefficients are vector functions and matrix functions depending only on t. Then equations (3.31) and (3.32) below determine the evolution of the optimal estimate $m(t)$ and the conditional covariance $\gamma(t)$.

By Corollary 3.3, $m(t)$ and $\gamma(t)$ satisfy the system of equations

$$\begin{aligned}m(t+1) = [a_0(t) + a_1(t)m(t)]\mathrm{P}_\gamma(t)Q_\gamma^+(t)[&\xi_{t+1} - A_0(t+1) \\ &- A_1(t+1)a_0(t) - A_1(t+1)a_1(t)m(t)],\end{aligned} \tag{3.31}$$

$$\gamma(t+1) = [a_1(t)\gamma(t)a_1^*(t) + b\circ b(t)] - P_\gamma(t)Q_\gamma^+(t)P_\gamma^*(t), \tag{3.32}$$

where

$$\begin{aligned}P_\gamma(t) = b_1(t)[A_1(t+1)b_1(t) &+ B_1(t+1)]^* \\ &+ b_2(t)[A_1(t+1)b_2(t) + B_2(t+1)],\end{aligned} \tag{3.33}$$

$$\begin{aligned}Q_\gamma(t) = [A_1(t+1)b_1(t) + B_1(t+1)]&[A_1(t+1)b_1(t) + B_1(t+1)]^* \\ + [A_1(t+1)b_2(t) + B_2(t+1)]&[A_1(t+1)b_2(t) \\ + B_2(t+1)]^* &+ A_1(t+1)a_1(t)\gamma(t)a_1^*(t)A_1^*(t+1).\end{aligned} \tag{3.34}$$

By means of the theorem on normal correlation, we obtain the following expressions for the initial conditions $m(0) = \mathbf{M}(\theta_0|\xi_0)$ and $\gamma(0) = \operatorname{Cov}(\theta_0|\xi_0)$:

$$\begin{aligned}m(0) = \mathbf{M}\theta_0 + \operatorname{Cov}\theta_0\cdot A_1^*(0)[A_1(0)\operatorname{Cov}\theta_0 A_1^*(0) &+ B\circ B(0)]^+ \\ &\cdot[\xi_0 - A_0(0) - A_1(0)\mathbf{M}\theta_0],\end{aligned} \tag{3.35}$$

$$\gamma(0) = \operatorname{Cov}\theta_0 - \operatorname{Cov}\theta_0 \cdot A_1^*(0)[A_1(0)\operatorname{Cov}\theta_0 A_1^*(0) + B \circ B(0)]^+ A_1(0)\operatorname{Cov}\theta_0.$$

Equations (3.31) and (3.32) are called the Kalman-Bucy equations.

We shall now point out a number of properties of the processes $m(t)$ and $\gamma(t)$ which are, respectively, the optimal estimate in the mean square sense, and the conditional covariance.

PROPERTY 3.1. *For every t, the variables $m(t)$ and $\theta_t - m(t)$ are uncorrelated, that is,*

$$\mathbf{M}\{m^*(t)[\theta_t - m(t)]\} = \mathbf{M}\{[\theta_t - m(t)]^* m(t)\} = 0, \tag{3.36}$$

and so

$$\mathbf{M}\theta_t^*\theta_t = \mathbf{M}m_t^* m_t + \mathbf{M}\{(\theta_t - m(t))^*(\theta_t - m(t))\}. \tag{3.37}$$

PROPERTY 3.2. *The conditional covariance $\gamma(t)$ does not depend explicitly upon the values of the coefficients $a_0(t, \omega)$ and $A_0(t, \omega)$.*

PROPERTY 3.3. *Suppose that with the exception of $a_0(t, \omega)$ and $A_0(t, \omega)$, all the coefficients in the system (3.1) as well as $\gamma(0)$ do not depend upon ω. Then the conditional covariance $\gamma(t)$ is a function of t only (not depending upon ω) and*

$$\gamma(t) = \mathbf{M}\{(\theta_t - m(t))(\theta_t - m(t))^*\}. \tag{3.38}$$

The distribution of the filtering error $\Delta_t = \theta_t - m(t)$ is normal, $\mathcal{N}(0, \gamma(t))$.

PROPERTY 3.4. *For every $t = 0, 1, 2, \cdots$, $m(t) = \mathbf{M}(\theta_t | \mathcal{F}_t^\xi)$ is the optimal, in the mean square sense, unbiased $(\mathbf{M}m(t) = \mathbf{M}\theta_t)$ estimate, that is,*

$$\operatorname{tr}\mathbf{M}\{(\theta_t - m(t))(\theta_t - m(t))^*\} = \inf_{\hat\theta_t \in \hat\Delta_t} \mathbf{M}\{(\theta_t - \hat\theta_t)(\theta_t - \hat\theta_t)^*\}, \tag{3.39}$$

where $\hat\Delta_t = \{\hat\theta_t : \mathbf{M}\hat\theta_t = \mathbf{M}\theta_t, \mathbf{M}\|\hat\theta_t\|^2 < \infty, \hat\theta_t$ is $\mathcal{F}_t^\xi$ measurable$\}$.

PROPERTY 3.5. *Let $G(t) = \|G_{i,j}(t)\|$ be a symmetric nonnegative definite matrix, depending upon t, and of order $k \times k$. Put*

$$g_t(x_1, \cdots, x_k) = \sum_{i,j} G_{i,j}(t) x_i x_j. \tag{3.40}$$

Then the estimate $m(t)$ is also optimal in the class $\hat\Delta_t$ in the sense that

$$\mathbf{M}g_t(\theta_t - m(t)) = \inf_{\hat\theta_t \in \hat\Delta_t} \mathbf{M}g_t(\theta_t - \hat\theta_t). \tag{3.41}$$

In fact, for all $x = (x_1, \cdots, x_k)$ and $a = (a_1, \cdots, a_k)$

$$\begin{aligned} g_t(x) &= g_t(a) + \sum_{i=1}^k (x_i - a_i) \left.\frac{\partial g_t}{\partial x_i}\right|_{x_i = a_i} + \sum_{i,j=1}^k (x_i - a_i)(x_j - a_j) G_{i,j}(t) \\ &= g_t(a) + 2\sum_{i=1}^k (x_i - a_i)\left(\sum_{j=1}^k G_{j,i}(t) a_j\right) \\ &\qquad + \sum_{i,j=1}^k (x_i - a_i)(x_j - a_j) G_{i,j}(t). \end{aligned} \tag{3.42}$$

But the matrix $G(t)$ is nonnegative definite. Therefore,

$$g_t(x) \geqq g_t(a) + 2 \sum_{i=1}^{k} (x_i - a_i) \left(\sum_{j=1}^{k} G_{j,i}(t) a_j(t) \right). \tag{3.43}$$

We choose $x = \theta_t - \hat{\theta}_t$, $a = \theta_t - m(t)$. Then

$$\begin{aligned} g_t(\theta_t - \hat{\theta}_t) \geqq\ & g_t(\theta_t - m(t)) \\ & + 2 \sum_{i=1}^{k} (m_i(t) - \theta_i(t)) \cdot \sum_{j=1}^{k} G_{j,i}(t)(\theta_j(t) - m_j(t)) \end{aligned} \tag{3.44}$$

and $\mathbf{M} g_t(\theta_t - \hat{\theta}_t) \geqq \mathbf{M} g_t(\theta_t - m(t))$, since for all $i, j = 1, \cdots, k$,

$$\begin{aligned} & \mathbf{M}[m_i(t) - \hat{\theta}_i(t)][\theta_i(t) - m_i(t)] \\ & \quad = \mathbf{M}\{[m_i(t) - \hat{\theta}_i(t)] \mathbf{M}[(\theta_i(t) - m_i(t)) | \mathscr{F}_t^\xi]\} = 0. \end{aligned} \tag{3.45}$$

PROPERTY 3.6. *Let us assume that $\gamma(0)$ and all the coefficients in (3.1), with the exception of $a_0(t, \omega)$ and $A_0(t, \omega)$, do not depend upon ω. In this case, the preceeding property can be substantially generalized, as follows.*

Let $g_t(x)$, $x \in R^k$, be a continuous nonnegative function which is convex from below, that is, for all $x \in R^k$, $a \in R^k$

$$g_t(x) \geqq g_t(a) + \sum_{i=1}^{k} (x_i - a_i) \lambda_i^{(t)}(a_1, \cdots, a_k). \tag{3.46}$$

We put

$$\begin{aligned} \tilde{\Delta}_t = \{\tilde{\theta}_t \colon \mathbf{M}\tilde{\theta}_t = \mathbf{M}\theta_t, \mathbf{M} g_t(\theta_t - \tilde{\theta}_t) < \infty, \mathbf{M}\lambda_i^{(t)}(\theta_t - \tilde{\theta}_t) < \infty, \\ i = 1, \cdots, k\}, \end{aligned} \tag{3.47}$$

a class of $\mathscr{F}_t^\xi$ measurable estimates $\tilde{\theta}_t$ of the vector θ_t. Then if $m(t) \in \tilde{\Delta}_t$, the estimate $m(t)$ is optimal in the class $\tilde{\Delta}_t$, that is

$$\mathbf{M} g_t(\theta_t - m(t)) = \inf_{\tilde{\theta}_t \in \tilde{\Delta}_t} \mathbf{M} g_t(\theta_t - \tilde{\theta}_t). \tag{3.48}$$

For the proof, we put $x = \theta_t - \tilde{\theta}_t$, $a = \theta_t - m(t)$ in (3.46). Then

$$g_t(\theta_t - \tilde{\theta}_t) \geqq g_t(\theta_t - m(t)) + \sum_{i=1}^{k} [m_i(t) - \tilde{\theta}_i(t)] \lambda_i^{(t)}(\theta_t - m(t)). \tag{3.49}$$

We shall show that under the stated hypotheses we have, for all i,

$$\mathbf{M}[m_i(t) - \theta_i(t)] \lambda_i^{(t)}(\theta_t - m(t)) = 0. \tag{3.50}$$

According to Theorem 3.1, the conditional distribution $P\{\theta_t - m(t) \leqq y | \mathscr{F}_t^\xi\}$ is normal, $\mathscr{N}(0, \gamma(t))$. But if $\gamma(0)$ and all the coefficients in the system (3.1), with the exception of $a_0(t, \omega)$ and $A_0(t, \omega)$, do not depend upon ω, then $\gamma(t)$ depends only upon t. Hence,

$$P\{\theta_t - m(t) \leqq y | \mathscr{F}_t^\xi\} = P\{\theta_t - m(t) \leqq y\}. \tag{3.51}$$

It follows that

$$\begin{aligned}(3.52)\qquad & \mathbf{M}[m_i(t) - \tilde{\theta}_i(t)][\lambda_i^{(t)}(\theta_t - m(t))] \\ &= \mathbf{M}[m_i(t) - \tilde{\theta}_i(t)]\mathbf{M}[\lambda_i^{(t)}(\theta_t - m(t))] = 0,\end{aligned}$$

since $\mathbf{M}m_i(t) = \mathbf{M}\tilde{\theta}_i(t) = \mathbf{M}\theta_i(t)$.

REMARK 3.1. Let us put

$$\begin{aligned}\Pi_{\beta,\alpha}(\tau, t) &= P\{\theta_\tau \leqq \beta \mid \mathscr{F}_t^\xi, \theta_t = \alpha\}, \\ m_\alpha(\tau, t) &= \mathbf{M}(\theta_\tau \mid \theta_t = \alpha, \mathscr{F}_\tau^\xi), \qquad (3.53) \\ \gamma_\alpha(\tau, t) &= \operatorname{Cov}(\theta_\tau \mid \theta_t = \alpha, \mathscr{F}_\tau^\xi),\end{aligned}$$

where $\tau \geqq t$. Then under the assumptions (a) through (d), the conditional distribution $\Pi_{\beta,\alpha}(\tau, t)$ is also normal, $\mathscr{N}(m_\alpha(\tau, t), \gamma_\alpha(\tau, t))$, and its parameters satisfy the system of equations relative to τ (compare with (3.18), (3.19)):

$$\begin{aligned}(3.54)\qquad m_\alpha(\tau + 1, t) &= [a_0(\tau) + a_1(\tau)m_\alpha(\tau, t)] + [b\circ B(\tau) + a_1(\tau)\gamma_\alpha(\tau, t)A_1^*(\tau)] \\ &\quad \cdot [B\circ B(\tau) + A_1(\tau)\gamma_\alpha(\tau, t)A_1^*(\tau)]^+[\xi_{\tau+1} - A_0(\tau) - A_1(\tau)m_\alpha(\tau, t)],\end{aligned}$$

$$\begin{aligned}(3.55)\qquad \gamma_\alpha(\tau + 1, t) &= [a_1(\tau)\gamma_\alpha(\tau, t)a_1^*(\tau) + b\circ b(\tau)] \\ &\quad - [b\circ B(\tau) + a_1(\tau)\gamma_\alpha(\tau, t)A_1^*(\tau)] \\ &\quad \cdot [B\circ B(\tau) + A_1(\tau)\gamma_\alpha(\tau, t)A_1^*(\tau)] \\ &\quad \cdot [b\circ B(\tau) + a_1(\tau)\gamma_\alpha(\tau, t)A_1^*(\tau)]^*,\end{aligned}$$

with initial conditions $m_\alpha(t, t) = \alpha$, $\gamma_\alpha(t, t) = 0$.

From equation (3.55), solved for the initial condition $\gamma_\alpha(t, t) = 0$, it follows that $\gamma_\alpha(\tau, t)$, $\tau \geqq t$, does not depend upon α. We will therefore put $\tilde{\gamma}(\tau, t) \equiv \gamma_\alpha(\tau, t)$ for $\tau \geqq t$.

Although under assumptions (a) through (d) the process (θ_t, ξ_t), $t = 0, 1, \cdots$, is not Gaussian, nevertheless it will be conditionally Gaussian in the following sense.

THEOREM 3.3. *Suppose that assumptions* (a) *through* (b) *are fulfilled. Then the process* (θ_t, ξ_t), $t = 0, 1, \cdots$, *is conditionally Gaussian, that is, for every* t *the conditional distribution* $P\{\theta_0 \leqq \alpha_0, \theta_1 \leqq \alpha_1, \cdots, \theta_t \leqq \alpha_t \mid \mathscr{F}_t^\xi\}$ *is Gaussian with probability* 1.

We defer the proof of this theorem to Section 5.

In Theorem 3.4, established below, we give a special representation for the process ξ_t, $t = 0, 1, \cdots$, which will be used in the sequel.

THEOREM 3.4. *Suppose that assumptions* (a) *through* (d) *are fulfilled. Then for every* t, *there exists a Gaussian vector* $\bar{\varepsilon}(t) = (\bar{\varepsilon}_1(t), \cdots, \bar{\varepsilon}_\ell(t))$ *with* $\mathbf{M}\bar{\varepsilon}(t) = 0$, $\mathbf{M}\bar{\varepsilon}(t)\bar{\varepsilon}^*(s) = \delta(t, s)E(\ell \times \ell)$, *such that the following representation of the vector* ξ_{t+1} *is valid with probability* 1:

$$\begin{aligned}(3.56)\qquad \xi_{t+1} &= A_0(t, \omega) + A_1(t, \omega)m(t) \\ &\quad + [B\circ B(t, \omega) + A_1(t, \omega)\gamma(t)A_1^*(t, \omega)]^{1/2}\bar{\varepsilon}(t + 1).\end{aligned}$$

If, moreover, the matrix $B\circ B(t, \omega) + A_1(t, \omega)\gamma(t)A_1^*(t, \omega)$ *is nonsingular* (*with probability* 1), $t = 0, 1, \cdots$, *then*

$$(3.57)\qquad \mathscr{F}_t^\xi = \mathscr{F}_t^{(\bar{\varepsilon}, \xi_0)}.$$

PROOF. We first assume that for all $t = 0, 1, \cdots$ the matrix $B \circ B(t)$ is positive definite. Then since the matrix $A_1(t)\gamma(t)A_1^*(t)$ is at least nonnegative definite, the matrix $[B \circ B + A_1\gamma(t)A_1^*]^{1/2}$ is positive definite, and therefore the random variable

$$(3.58) \qquad \bar{\varepsilon}(t+1) = [B \circ B(t, \omega) + A_1(t, \omega)\gamma(t)A_1^*(t, \omega)]^{-1/2} \cdot [A_1(t, \omega)(\theta_t - m(t)) + B_1(t, \omega)\varepsilon_1(t+1) + B_2(t, \omega)\varepsilon_2(t+1)]$$

can be defined.

Under conditioning by $\mathscr{F}_t^\xi$ the distribution of the vector θ_t is by Theorem 3.1 Gaussian, and the random vectors $\varepsilon_1(t+1)$ and $\varepsilon_2(t+1)$ do not depend upon $\xi^t = (\xi_0, \cdots, \xi_t)$. Therefore, the conditional (conditioned by $\mathscr{F}_t^\xi$) distribution of the vector $\bar{\varepsilon}(t+1)$ is likewise Gaussian, and it is not hard to calculate that

$$(3.59) \qquad \begin{aligned} \mathbf{M}[\bar{\varepsilon}(t+1) \mid \mathscr{F}_t^\xi] &= 0, \\ \mathrm{Cov}\,[\bar{\varepsilon}(t+1) \mid \mathscr{F}_t^\xi] &= E(\ell \times \ell). \end{aligned}$$

Hence, it is evident that the parameters of the conditional distribution of the vector $\bar{\varepsilon}(t+1)$ do not depend upon the condition. Consequently, the (unconditional) distribution of the vector $\bar{\varepsilon}(t+1)$ is also Gaussian, with $\mathbf{M}\bar{\varepsilon}(t+1) = 0$, $\mathrm{Cov}\,\bar{\varepsilon}(t+1) = E(\ell \times \ell)$.

Similarly, using Theorem 3.3, one can show that for any t the joint distribution of the vectors $(\bar{\varepsilon}(1), \cdots, \bar{\varepsilon}(t))$ is Gaussian, with $\mathrm{Cov}\,(\bar{\varepsilon}(u), \bar{\varepsilon}(v)) = \delta(u - v)E$. From this follows the independence of all the coordinates of the vectors $(\bar{\varepsilon}(1), \cdots, \bar{\varepsilon}(t))$.

The representation (3.56) follows from (3.58) and (3.1) in an obvious way.

To prove the assertion (3.57), we note that by virtue of (3.56) $\mathscr{F}_t^\xi \subseteqq \mathscr{F}_t^{(\bar{\varepsilon}, \xi_0)}$. If the matrix $B \circ B + A_1\gamma(t)A_1^*$ is nonsingular, then by (3.56)

$$(3.60) \qquad \bar{\varepsilon}(t) = (B \circ B(t-1, \omega) + A_1(t-1, \omega)\gamma(t-1)A_1^*(t-1, \omega))^{-1/2} \cdot (\xi_t - A_0(t-1, \omega) - A_1(t-1, \omega)m(t-1)).$$

It follows that $\mathscr{F}_t^\xi \supseteqq \mathscr{F}_t^{(\bar{\varepsilon}, \xi_0)}$. Hence, $\mathscr{F}_t^\xi = \mathscr{F}_t^{(\bar{\varepsilon}, \xi_0)}$.

Let us now assume that for some t the matrix $B \circ B$ is singular. The assertion of the theorem obviously remains valid for this case, so long as the matrix $B \circ B + A_1\gamma(t)A_1^*$ is nonsingular.

Thus, suppose that this matrix is singular. We construct a sequence of independent Gaussian random vectors $z(t) = (z_1(t), \cdots z_\ell(t))$, $\mathbf{M}z(t) = 0$, $\mathbf{M}z(t)z^*(t) = E(\ell \times \ell)$, which are also independent of the processes $\varepsilon_1(t), \varepsilon_2(t)$, $t \geqq 0$, and of θ_0, ξ_0. (Such a construction is always possible, though admittedly it may require extension of the basic probability space.) Put

$$(3.61) \qquad \bar{\varepsilon}(t+1) = D^+[A_1(\theta_t - m(t)) + B_1\varepsilon_1(t+1) + B_2\varepsilon_2(t+1)] + (E - D^+D)z(t+1),$$

where $D(t) = [B \circ B(t) + A_1(t)\gamma(t)A_1^*(t)]^{1/2}$. It is not hard to convince oneself that this sequence of random vectors has the properties stated in the theorem.

To verify the representation (3.56) it is obviously sufficient to show that

$$D(t)\bar{\varepsilon}(t+1) = A_1(t)[\theta_t - m(t)] + B_1(t)\varepsilon_1(t+1) + B_2(t)\varepsilon_2(t+1). \tag{3.62}$$

Multiplying the left and right side of (3.61) by D, we obtain

$$\begin{aligned} D\bar{\varepsilon}(t+1) = {} & [A_1(\theta_t - m(t)) + B_1\varepsilon_1(t+1) + B_2\varepsilon_2(t+1)] \\ & + (E - DD^+)[A_1(\theta_t - m(t)) + B_1\varepsilon_1(t+1) + B_2\varepsilon_2(t+1)] \\ & + D[E - D^+D]z(t+1). \end{aligned} \tag{3.63}$$

But by the first property in (2.4) of the pseudo-inverse, $D[E - D^+D] = D - DD^+D = 0$; consequently, with probability 1

$$D[E - D^+D]z(t+1) = 0. \tag{3.64}$$

Let us put

$$\zeta(t+1) = (E - DD^+)[A_1(\theta_t - m(t)) + B_1\varepsilon_1(t+1) + B_2\varepsilon_2(t+1)]. \tag{3.65}$$

Then

$$\begin{aligned} \mathbf{M}\zeta(t+1)\zeta^*(t+1) &= \mathbf{M}\{\mathbf{M}\zeta(t+1)\zeta^*(t+1) \mid \mathscr{F}_t^\xi\} \\ &= \mathbf{M}[(E - DD^+)DD^*(E - DD^+)] \\ &= \mathbf{M}[(DD^* - DD^+DD^*)(E - DD^+)] \\ &= \mathbf{M}[(DD^* - DD^*)(E - DD^+)] = 0. \end{aligned} \tag{3.66}$$

Thus, $\zeta(t+1) = 0$ with probability 1, which together with (3.64) proves (3.62).

REMARK 3.2. If the matrix $B \circ B + A_1\gamma(t)A_1^*$ is singular, then $\mathscr{F}_t^{(\varepsilon, \xi_0)} \supsetneqq \mathscr{F}_t^\xi$.

4. Estimation of the parameters

In this section, we shall consider various examples of parameter estimation which illustrate Theorems 3.1 and 3.2.

Let $\theta = (\theta_1, \cdots, \theta_k)$ be a Gaussian random vector with $\mathbf{M}\theta = m$, $\operatorname{Cov}\theta = \gamma$. We assume that θ is unobservable, and that observations are made on a sequence of ℓ dimensional vectors ξ_t, $t \geqq 0$, such that

$$\xi_{t+1} = A_0(t, \omega) + A_1(t, \omega)\theta + B(t, \omega)\varepsilon(t+1), \qquad \xi_0 = 0. \tag{4.1}$$

(It is further assumed that $\mathbf{M}\operatorname{tr} A_1(t, \omega)A_1^*(t, \omega) < \infty$, $t = 0, 1, \cdots$.)

The sequence (θ, ξ_t) being considered is obviously a special case of (3.1), with $a_0 = 0$, $a_1 = E(k \times k)$, $b_1 = 0$, $b_2 = 0$, $B_1 = B$, $B_2 = 0$.

From (3.5) and (3.6), we find that

$$m(t+1) = m(t) + \gamma(t)A_1^*[B \circ B + A_1\gamma(t)A_1^*]^+[\xi_{t+1} - A_0 - A_1 m(t)], \tag{4.2}$$

$$\gamma(t+1) = \gamma(t) - \gamma(t)A_1^*[B \circ B + A_1\gamma(t)A_1^*]^+ A_1\gamma(t) \tag{4.3}$$

with $m(0) = m$, $\gamma(0) = \gamma$.

THEOREM 4.1. *Suppose that assumptions* (a) *through* (c) *are fulfilled and that for all t the matrix $BB^*(t)$ is nonsingular. Then*

$$\gamma(t) = \left[E + \gamma \sum_{s=0}^{t-1} A_1^*(s, \omega)(BB^*)^{-1}(s, \omega)A_1(s, \omega)\right]^{-1} \gamma, \tag{4.4}$$

$$m(t) = \left[E + \gamma \sum_{s=0}^{t-1} A_1^*(s, \omega)(BB^*)^{-1}(s, \omega)A_1(s, \omega)\right]^{-1} \cdot\left[m + \gamma \sum_{s=0}^{t-1} A_1^*(s, \omega)(BB^*)^{-1}(s, \omega)\big(\xi_{s+1} - A_0(s, \omega)\big)\right]. \tag{4.5}$$

PROOF. From the obvious identity

$$(BB^*)^{-1}BB^* = (BB^* + A_1\gamma A_1^*)^{-1}(BB^* + A_1\gamma A_1^*), \tag{4.6}$$

it follows that

$$(BB^*)^{-1} = (BB^* + A_1\gamma A_1^*)^{-1} + (BB^* + A_1\gamma A_1^*)^{-1}A_1\gamma A_1^*(BB^*)^{-1}. \tag{4.7}$$

Multiplying both sides of (4.7) from the left by γA_1^*, and from the right by A_1, we find that

$$\gamma A_1^*(BB^*)^{-1}A_1 = \gamma A_1^*(BB^* + A_1\gamma A_1^*)^{-1}A_1 + \gamma A_1^*(BB^* + A_1\gamma A_1^*)^{-1}A_1\gamma A_1^*(BB^*)^{-1}A_1. \tag{4.8}$$

Let us assume that the matrix $\gamma(t)$ is nonsingular. Then from (4.8) we obtain

$$\gamma\gamma^{-1} + \gamma A_1^*(BB^*)^{-1}A_1 - \gamma A_1^*(BB^* + A_1\gamma A_1^*)^{-1}\gamma\gamma^{-1} - \gamma A_1^*(BB^* + A_1\gamma A_1^*)^{-1}A_1\gamma A_1^*(BB^*)^{-1}A_1 = E. \tag{4.9}$$

Hence,

$$\gamma[\gamma^{-1} + A_1^*(BB^*)^{-1}A_1] - \gamma A_1^*[BB^* + A_1\gamma A_1^*]^{-1}A_1\gamma[\gamma^{-1} + A_1^*(BB^*)^{-1}A_1] = E \tag{4.10}$$

or

$$[\gamma - \gamma A_1^*(BB^* + A_1\gamma A_1^*)^{-1}A_1\gamma][\gamma^{-1} + A_1^*(BB^*)^{-1}A_1] = E. \tag{4.11}$$

The first factor in (4.11) equals $\gamma(t + 1)$. Therefore, the matrix $\gamma^{-1}(t + 1)$ is defined and

$$\gamma^{-1}(t + 1) = \gamma^{-1}(t) + A_1^*(BB^*)^{-1}A_1. \tag{4.12}$$

From the foregoing it follows that if the matrix $\gamma(0) = \gamma$ is nonsingular, then for every t the matrix $\gamma(t)$ is likewise nonsingular and

$$\gamma(t + 1) = \left[\gamma^{-1} + \sum_{s=0}^{t} A_1^*(s, \omega)(BB^*)^{-1}(s, \omega)A_1(s, \omega)\right]^{-1}. \tag{4.13}$$

We at once obtain formula (4.4).

If the matrix $\gamma(0) = \gamma$ is singular, then putting $\gamma^\varepsilon(0) = \gamma + \varepsilon E$, $\varepsilon > 0$, we find from the preceding that

$$(4.14) \qquad \gamma^\varepsilon(t+1) = \left[E + \gamma^\varepsilon(0) \sum_{s=0}^{t} A_1^*(BB^*)^{-1}A_1\right]^{-1} \gamma^\varepsilon(0).$$

From (4.3), it follows that $\gamma(t+1) = \lim_{\varepsilon \downarrow 0} \gamma^\varepsilon(t+1)$. Therefore, (4.4) can be obtained from (4.14) by passing to the limit $\varepsilon \downarrow 0$.

To prove the representation (4.5), we first assume that the matrix $\gamma(0) = \gamma$ is nonsingular. As was shown above, in this case the matrix $\gamma(t)$ is nonsingular for all $t \geqq 0$, and from (4.2) and (4.3) we find that

$$\begin{aligned}(4.15) \quad & m(t+1) \\ &= [\gamma(t)\gamma^{-1}(t) - \gamma(t)A_1^*(BB^* + A_1\gamma(t)A_1^*)^{-1}A_1\gamma(t)\gamma^{-1}(t)]m(t) \\ &\qquad + \gamma(t)A_1^*(BB^* + A_1\gamma(t)A_1^*)^{-1}[\xi_{t+1} - A_0] \\ &= \gamma(t+1)\gamma^{-1}(t)m(t) + \gamma(t)A_1^*[BB^* + A_1\gamma(t)A_1^*]^{-1}[\xi_{t+1} - A_0].\end{aligned}$$

We shall show that

$$(4.16) \qquad \gamma(t)A_1^*[BB^* + A_1\gamma(t)A_1^*]^{-1} = \gamma(t+1)A_1^*(t)(BB^*)^{-1}(t).$$

To this end, we multiply both sides of (4.7) from the left by $\gamma(t)A_1^*(t)$. Then

$$\begin{aligned}(4.17) \quad & \gamma(t)A_1^*[BB^* + A_1\gamma(t)A_1^*]^{-1} \\ &= \gamma(t)A_1^*(BB^*)^{-1} - \gamma(t)A_1^*[(BB^*)^{-1} + A_1\gamma(t)A_1^*]^{-1}A_1\gamma(t)A_1^*(BB^*)^{-1}.\end{aligned}$$

By virtue of (4.3), the right side of (4.17) can be transformed to the following form:

$$\begin{aligned}(4.18) \quad & \gamma(t)A_1^*(BB^*)^{-1} - \gamma(t)A_1^*[BB^* + A_1\gamma(t)A_1^*]^{-1}A_1\gamma(t)A_1^*(BB^*)^{-1} \\ &= [\gamma(t)\gamma^{-1}(t) - \gamma(t)A_1^*(BB^* + A_1\gamma(t)A_1^*)^{-1}A_1\gamma(t)\gamma^{-1}(t)]\gamma(t)A_1^*(BB^*)^{-1} \\ &= \gamma(t+1)\gamma^{-1}(t)\gamma(t)A_1^*(BB^*)^{-1} = \gamma(t+1)A_1^*(BB^*)^{-1}.\end{aligned}$$

This establishes (4.16).

It follows from (4.15) and (4.16) that

$$\begin{aligned}(4.19) \quad & m(t+1) \\ &= \gamma(t+1)\gamma^{-1}(t)m(t) + \gamma(t+1)A_1^*(t)(BB^*)^{-1}(t)[\xi_{t+1} - A_0(t)].\end{aligned}$$

Thus,

$$(4.20) \quad \gamma^{-1}(t+1)m(t+1) = \gamma^{-1}(t)m(t) + A_1^*(t)(BB^*)^{-1}(t)[\xi_{t+1} - A_0(t)],$$

and so

$$(4.21) \quad \gamma^{-1}(t+1)m(t+1) = \gamma^{-1}m + \sum_{s=0}^{t} A_1^*(s)(BB^*)^{-1}(s)[\xi_{s+1} - A_0(s)].$$

Consequently,

$$(4.22) \qquad m(t+1) = \gamma(t+1)\left[\gamma^{-1}m + \sum_{s=0}^{t} A_1^*(s)(BB^*)^{-1}(s)[\xi_{s+1} - A_0(s)]\right],$$

which together with (4.4) leads to the required representation (4.5), for the case in which the matrix $\gamma(0) = \gamma$ is nonsingular. If this matrix is singular, then the proof of the representation (4.5) can be carried out in the same way as the proof of (4.4).

REMARK 4.1. Let $m_n(t)$ and $\gamma_n(t)$ be the parameters of the *a posteriori* distributions $P\{\theta \leqq \alpha | \mathscr{F}_t^\xi\}$, corresponding to normal $\mathscr{N}(m_n, \gamma_n)$ *a priori* distributions, where $\lim_{n\to\infty} \gamma_n^{-1} = 0$, and the vector m has bounded coordinates. If the matrix $\Sigma_{s=0}^t A_1^*(s)(BB^*)^{-1}(s)A_1(s)$ is nonsingular, then it is not hard to prove that the limits $m(t) = \lim_{n\to\infty} m_n(t)$ and $\gamma(t) = \lim_{n\to\infty} \gamma_n(t)$ exist (with probability 1) and

$$(4.23)\qquad \begin{aligned} \gamma(t+1) &= \left[\sum_{s=0}^t A_1^*(s)(BB^*)^{-1}(s)A_1(s)\right]^{-1}, \\ m(t+1) &= \left[\sum_{s=0}^t A_1^*(s)(BB^*)^{-1}(s)A_1(s)\right]^{-1} \\ &\quad \cdot\left[\sum_{s=0}^t A_1^*(s)(BB^*)^{-1}(s)(\xi_{s+1} - A_0(s)\right]. \end{aligned}$$

We remark that the estimate of the vector θ given by (4.23) coincides with the estimate obtained by the maximum likelihood method.

5. Direct equations for interpolation

As was mentioned in Section 1, by interpolation is understood the problem of constructing the best estimates (in the mean square sense) of the vector θ_t in terms of observations on $\xi^\tau = (\xi_0, \cdots, \xi_\tau)$, where $\tau \geqq t$. We suppose that the process (θ_t, ξ_t) satisfies the system (3.1) and the assumptions (a) through (d).

We put $\Pi_\alpha(t, \tau) = P\{\theta_t \leqq \alpha | \mathscr{F}_\tau^\xi\}$ and

$$(5.1)\qquad m(t, \tau) = \mathbf{M}(\theta_t | \mathscr{F}_\tau^\xi), \qquad \gamma(t, \tau) = \mathrm{Cov}\,(\theta_t | \mathscr{F}_\tau^\xi).$$

In this section, we shall deduce direct equations (in terms of τ for fixed $t \leqq \tau$) for the optimal (in the mean square sense) estimate of $m(t, \tau)$ and for the conditional covariance $\gamma(t, \tau)$.

From these equations, in particular, it will be evident how the estimate of the vector θ_t improves as the "datum" grows, that is, as $\tau - t \to \infty$. The derivation of inverse equations (in terms of t for fixed τ) will be carried out in Section 6.

THEOREM 5.1. *Suppose that the distribution* $\Pi_\alpha(t) = P\{\theta_t \leqq \alpha | \mathscr{F}_t^\xi\}$ *is Gaussian with probability* 1. *Then for all* $\tau > t$ *the distribution* $\Pi_\alpha(t, \tau)$ *will be Gaussian with probability* 1.

To prove this theorem we need the following.

LEMMA 5.1. *Under the assumptions of Theorem* 5.1, *the conditional expectation* $m_\alpha(\tau, t) = \mathbf{M}(\theta_\tau | \mathscr{F}_\tau^\xi, \theta_t = \alpha)$ *admits the representation*

$$(5.2)\qquad m_\alpha(\tau, t) = \varphi_t^\tau \alpha + \psi_t^\tau,$$

for matrices φ_t^τ *and vectors* ψ_t^τ *which do not depend on* α.

The matrices φ_t^τ *are given by*

$$\begin{aligned}
&\varphi_t^t = E(k \times k) \\
(5.3)\qquad &\varphi_t^\tau = \prod_{s=t}^{\tau-1} \{a_1(s) - [b \circ B(s) + a_1(s)\tilde{\gamma}(s, t)A_1^*(s)] \\
&\qquad\qquad [B \circ B(s) + A_1(s)\tilde{\gamma}(s, t)A_1^*(s)]^+ A_1(s)\},
\end{aligned}$$

where $\Pi_{s=t}^{\tau-1} A_s$ *represents the matrix product* $A_{\tau-1} \cdots A_t$. *The vectors* ψ_t^τ *are given by*

$$\begin{aligned}
&\psi_t^t = 0 \\
(5.4)\qquad &\psi_t^\tau = \sum_{s=t}^{\tau-1} \varphi_s^{\tau-1}\{a_0(s) + [b \circ B(s) + a_1(s)\tilde{\gamma}(s, t)A_1^*(s)] \\
&\qquad \cdot [B \circ B(s) + A_1(s)\tilde{\gamma}(s, t)A_1^*(s)]^+ (\xi_{s+1} - A_0(s)\}, \qquad \tau > t.
\end{aligned}$$

The matrices $\tilde{\gamma}(s, t)$, $s \geqq t$, *are determined from the equations*

$$\begin{aligned}
(5.5)\qquad \tilde{\gamma}(s, t) = {} & [a_1(s-1)\tilde{\gamma}(s-1, t)a_1^*(s-1) + b \circ b(s-1)] \\
& - [b \circ B(s-1) + a_1(s-1)\tilde{\gamma}(s-1, t)A_1^*(s-1)] \\
& \cdot [B \circ B(s-1) + A_1(s-1)\tilde{\gamma}(s-1, t)A_1(s-1)]^+ \\
& \cdot [b \circ B(s-1) + a_1(s-1)\tilde{\gamma}(s-1, t)A^*(s-1)]^*
\end{aligned}$$

with the initial condition $\tilde{\gamma}(t, t) = 0$.

PROOF. By the remark to Theorem 3.2, $m_\alpha(\tau, t)$ and $\gamma_\alpha(\tau, t) = \operatorname{Cov}(\theta_\tau | \mathscr{F}_\tau^\xi, \theta_t = \alpha)$ satisfy equations (3.54) and (3.55) with the initial conditions $m_\alpha(t, t) = \alpha$, $\gamma_\alpha(t, t) = 0$. It follows from (3.55) and the condition $\gamma_\alpha(t, t) = 0$ that $\gamma_\alpha(\tau, t)$ does not depend upon α. Put $\tilde{\gamma}(\tau, t) = \gamma_\alpha(\tau, t)$. By induction we obtain from (3.54) the representation (5.2), where φ_t^τ and ψ_t^τ are defined by (5.3) and (5.4).

PROOF OF THEOREM 5.1. We first show that the conditional distribution $P\{\theta_t \leqq \alpha, \xi_\tau \leqq x | \mathscr{F}_{\tau-1}^\xi\}$, $\tau > t$, is Gaussian. To do this, we calculate the conditional characteristic function

$$\begin{aligned}
(5.6)\qquad & \mathbf{M}\{\exp i[\langle Z_1, \theta_t\rangle + \langle Z_2, \xi_\tau\rangle] | \mathscr{F}_{\tau-1}^\xi\} \\
& = \mathbf{M}\{\exp i[\langle Z_1, \theta_t\rangle]\mathbf{M}[\exp i\langle Z_2, \xi_\tau\rangle | \mathscr{F}_{\tau-1}^\xi, \theta_t] | \mathscr{F}_{\tau-1}^\xi\}.
\end{aligned}$$

Obviously,

$$\begin{aligned}
(5.7)\qquad & \mathbf{M}\{\exp i\langle Z_2, \xi_\tau\rangle | \mathscr{F}_{\tau-1}^\xi, \theta_{\tau-1}\} \\
& = \exp\{i\langle Z_2, A_0(\tau-1) + A_1(\tau-1)\theta_{\tau-1}\rangle - \tfrac{1}{2}\langle B \circ B(\tau-1)Z_2, Z_2\rangle\}.
\end{aligned}$$

By the remark to Theorem 3.2,

$$(5.8)\qquad P\{\theta_{\tau-1} \leqq \beta | \mathscr{F}_{\tau-1}^\xi, \theta_t\} \sim \mathscr{N}\{m_{\theta_t}(\tau-1, t), \tilde{\gamma}(\tau-1, t)\},$$

and hence by virtue of (5.7)

$$\begin{aligned}(5.9)\qquad & \mathbf{M}\{\mathbf{M}[\exp i\langle Z_2, \xi_\tau\rangle \mid \mathscr{F}^\xi_{\tau-1}, \theta_{\tau-1}] \mid \mathscr{F}^\xi_{\tau-1}, \theta_t\} \\ & = \exp[i\langle Z_2, A_0(\tau - 1)\rangle - \tfrac{1}{2}\langle B\circ B(\tau - 1)Z_2, Z_2\rangle] \\ & \quad \mathbf{M}\{\exp i\langle Z_2, A_1(\tau - 1)\theta_{\tau-1}\rangle \mid \mathscr{F}^\xi_{\tau-1}, \theta_t\} \\ & = \exp[i\langle Z_2, A_0(\tau - 1)\rangle - \tfrac{1}{2}\langle B\circ B(\tau - 1)Z_2, Z_2\rangle] \\ & \quad \cdot\exp[i\langle Z_2, A_1(\tau - 1)m_{\theta_t}(\tau - 1, t)\rangle \\ & \qquad - \tfrac{1}{2}\langle A_1(\tau - 1)\tilde{\gamma}(\tau - 1, t)A_1^*(\tau - 1)Z_2, Z_2\rangle].\end{aligned}$$

By Lemma 5.1, $m_{\theta_t}(\tau - 1, t) = \varphi_t^{\tau-1}\theta_t + \psi_t^{\tau-1}$. Therefore,

$$\begin{aligned}(5.10)\qquad & \mathbf{M}[\exp i\langle Z_2, \xi_\tau\rangle \mid \mathscr{F}^\xi_{\tau-1}, \theta_t] \\ & = \exp[i\langle Z_2, A_0(\tau - 1) + A_1(\tau - 1)\psi_t^{\tau-1}\rangle \\ & \quad - \tfrac{1}{2}\langle(B\circ B(\tau - 1) + A_1(\tau - 1)\tilde{\gamma}(\tau - 1, t)A_1^*(\tau - 1))Z_2, Z_2\rangle \\ & \qquad + i\langle Z_2, A_1(\tau - 1)\varphi_t^{\tau-1}\theta_t\rangle],\end{aligned}$$

which together with (5.6) leads to the equality

$$\begin{aligned}(5.11)\qquad & \mathbf{M}\{\exp i[\langle Z_1, \theta_t\rangle + \langle Z_2, \xi_\tau\rangle] \mid \mathscr{F}^\xi_{\tau-1}\} \\ & = \exp[i\langle Z_2, A_0(\tau - 1) + A_1(\tau - 1)\psi_t^{\tau-1}\rangle \\ & \quad - \tfrac{1}{2}\langle(B\circ B(\tau - 1) + A_1(\tau - 1)\tilde{\gamma}(\tau - 1, t)A_1^*(\tau - 1))Z_2, Z_2\rangle] \\ & \qquad \cdot\mathbf{M}\{\exp i[\langle Z_1, \theta_t\rangle + \langle Z_2, A_1(\tau - 1)\varphi_t^{\tau-1}\theta_t\rangle] \mid \mathscr{F}^\xi_{\tau-1}\}.\end{aligned}$$

Let $\tau = t + 1$. Since $\Pi_\alpha(t, t) \sim \mathscr{N}(m(t), \gamma(t))$, it follows from (5.11) that the distribution $P\{\theta_t \leqq \alpha, \xi_{t+1} \leqq x \mid \mathscr{F}^\xi_t\}$ is normal. From this it is not hard to deduce the normality of the distribution $\Pi_\alpha(t, t + 1)$. The normality of the distribution $\Pi_\alpha(t, \tau)$ for all $\tau > t$ can be carried out, making use of (5.11), by induction.

REMARK 5.1. Suppose that $s < t \leqq \tau$. Then by the same methods which were used in the proof of Theorem 5.1, one can establish that the *a posteriori* distribution $P\{\theta_t, \leqq \alpha \mid \mathscr{F}^\xi_\tau, \theta_s = \delta\}$ is normal with probability 1.

According to Theorem 5.1 the *a posteriori* distribution $\Pi_\alpha(t, \tau) = P\{\theta_t \leqq \alpha \mid \mathscr{F}^\xi_\tau\}$ is normal, $\mathscr{N}(m(t, \tau), \gamma(t, \tau))$. We now turn to the derivation of direct equations (with respect to τ) for the parameters $m(t, \tau)$ and $\gamma(t, \tau)$ of this distribution.

THEOREM 5.2. *Suppose that the distribution $\Pi_\alpha(t)$ is Gaussian with probability 1. Then the parameters $m(t, \tau)$ and $\gamma(t, \tau)$ of the conditional distribution $\Pi_\alpha(t, \tau)$ satisfy the following equations with respect to τ, $\tau \geqq t + 1$:*

$$\begin{aligned}(5.12)\qquad & m(t, \tau + 1) \\ & = m(t, \tau) + \gamma(t, \tau)(\varphi_t^\tau)^*A_1^*(\tau)[B\circ B(\tau) + A_1(\tau)\gamma(\tau)A_1^*(\tau)]^+ \\ & \qquad \cdot[\xi_{\tau+1} - A_0(\tau) - A_1(\tau)m(\tau)],\end{aligned}$$

$$\begin{aligned}(5.13)\qquad & \gamma(t, \tau + 1) = \gamma(t, \tau) - \gamma(t, \tau)(\varphi_t^\tau)^*A_1^*(\tau) \\ & \qquad \cdot[B\circ B(\tau) + A_1(\tau)\gamma(\tau)A_1^*(\tau)]^+A_1(\tau)\varphi_t^\tau\gamma(t, \tau),\end{aligned}$$

where $m(t, t) = m(t)$, $\gamma(t, t) = \gamma(t)$, and the matrices φ_t^τ are defined in (5.3).

PROOF. As was established in the proof of Theorem 5.1, the conditional distribution $P\{\theta_t \leqq \alpha,\ \xi_\tau \leqq x \mid \mathscr{F}^\xi_{\tau-1}\}$ is normal. Its parameters can be determined from formula (5.11). However, they can be found more easily by using the theorem on normal correlation.

Since the conditional distribution $P\{\theta_t \leqq \alpha,\ \xi_\tau \leqq x \mid \mathscr{F}^\xi_{\tau-1}\}$ is normal, by the Remark 2.1 we have

$$\mathbf{M}(\theta_t \mid \xi_\tau, \mathscr{F}^\xi_{\tau-1}) = \mathbf{M}(\theta_t \mid \mathscr{F}^\xi_{\tau-1}) + d_{1,2}d^+_{2,2}\big(\xi_\tau - \mathbf{M}(\xi_\tau \mid \mathscr{F}^\xi_{\tau-1})\big), \tag{5.14}$$

where

$$\begin{aligned} d_{1,2} &= \operatorname{Cov}(\theta_t, \xi_\tau \mid \mathscr{F}^\xi_{\tau-1}), \\ d_{2,2} &= \operatorname{Cov}(\xi_\tau \mid \mathscr{F}^\xi_{\tau-1}) \\ &= \operatorname{Cov}\big[A_1(\tau-1)\big(\theta_{\tau-1} - m(\tau-1)\big) \\ &\qquad + B_1(\tau-1)\varepsilon_1(\tau) + B_2(\tau-1)\varepsilon_2(\tau) \mid \mathscr{F}^\xi_{\tau-1}\big] \\ &= A_1(\tau-1)\gamma(\tau-1)A_1^*(\tau-1) + B\circ B(\tau-1). \end{aligned} \tag{5.15}$$

We wish to simplify the expression for $d_{1,2}$. To this end, we note that by virtue of Lemma 5.1

$$\begin{aligned} m(\tau-1) &= \mathbf{M}(\theta_{\tau-1} \mid \mathscr{F}^\xi_{\tau-1}) = \mathbf{M}[\mathbf{M}(\theta_{\tau-1} \mid \mathscr{F}^\xi_{\tau-1}, \theta_t) \mid \mathscr{F}^\xi_{\tau-1}] \\ &= \mathbf{M}[\varphi_t^{\tau-1}\theta_t + \psi_t^{\tau-1} \mid \mathscr{F}^\xi_{\tau-1}] = \varphi_t^{\tau-1}m(t,\tau-1) + \psi_t^{\tau-1}. \end{aligned} \tag{5.16}$$

Further,

$$\begin{aligned} &\mathbf{M}\big(\theta_{\tau-1} - m(\tau-1) \mid \mathscr{F}^\xi_{\tau-1}, \theta_t\big) \\ &\quad = \varphi_t^{\tau-1}\theta_t + \psi_t^{\tau-1} - \big(\varphi_t^{\tau-1}m(t,\tau-1) + \psi_t^{\tau-1}\big) \\ &\quad = \varphi_t^{\tau-1}\big(\theta_t - m(t,\tau-1)\big), \end{aligned} \tag{5.17}$$

$$\mathbf{M}(\xi_\tau \mid \mathscr{F}^\xi_{\tau-1}) = A_0(\tau-1) + A_1(\tau-1)m(\tau-1) \tag{5.18}$$

and, once again by Lemma 5.1,

$$\begin{aligned} &\mathbf{M}\{[\xi_\tau - \mathbf{M}(\xi_\tau \mid \mathscr{F}^\xi_{\tau-1}]^* \mid \mathscr{F}^\xi_{\tau-1}, \theta_t\} \\ &\quad = \mathbf{M}\{[A_1(\tau-1)\big(\theta_{\tau-1} - m(\tau-1)\big) \\ &\qquad\qquad + B_1(\tau-1)\varepsilon_1(\tau) + B_2(\tau-1)\varepsilon_2(\tau)]^* \mid \mathscr{F}^\xi_{\tau-1}, \theta_t\} \\ &\quad = \mathbf{M}\{[A_1(\tau-1)\big(\theta_{\tau-1} - m(\tau-1)\big)]^* \mid \mathscr{F}^\xi_{\tau-1}, \theta_t\} \\ &\quad = [\theta_t - m(t,\tau-1)]^*(\varphi_t^{\tau-1})^*A_1^*(\tau-1). \end{aligned} \tag{5.19}$$

Thus,

$$\begin{aligned} d_{1,2} &= \operatorname{Cov}(\theta_t, \xi_\tau \mid \mathscr{F}^\xi_{\tau-1}) \\ &= \mathbf{M}\{[\theta_t - m(t,\tau-1)][\xi_\tau - \mathbf{M}(\xi_\tau \mid \mathscr{F}^\xi_{\tau-1})]^* \mid \mathscr{F}^\xi_{\tau-1}\} \\ &= \mathbf{M}\{[\theta_t - m(t,\tau-1)][\theta_t - m(t,\tau-1)]^* \\ &\qquad (\varphi_t^{\tau-1})^*A_1^*(\tau-1) \mid \mathscr{F}^\xi_{\tau-1}\} \\ &= \gamma(t,\tau-1)(\varphi_t^{\tau-1})^*A_1^*(\tau-1). \end{aligned} \tag{5.20}$$

From (5.14), (5.15), (5.18) and (5.20), we obtain (5.12).

Let us now establish equation (5.13). By the remark to Theorem 2.1,

$$\gamma(t, \tau) = \operatorname{Cov}(\theta_t \mid \xi_\tau, \mathscr{F}^{\xi}_{\tau-1}) = d_{1,1} - d_{1,2}d^{+}_{2,2}d^{*}_{1,2}, \tag{5.21}$$

where

$$d_{1,1} = \operatorname{Cov}(\theta_t \mid \mathscr{F}^{\xi}_{\tau-1}) = \gamma(t, \tau - 1). \tag{5.22}$$

We obtain the desired equation (5.13) from (5.21), (5.22), (5.20), and (5.15).

THEOREM 5.3. *If the matrices $B \circ B(s)$, $s = 0, 1, \cdots$, are nonsingular, then the solutions $m(t, \tau)$ and $\gamma(t, \tau)$ of equations* (5.12) *and* (5.13) *are given by*

$$\begin{aligned} m(t,\tau) &= \Big[E + \gamma(t) \sum_{s=t}^{\tau-1} (\varphi^s_t)^* A^*_1(s)\big(B\circ B(s) + A_1(s)\tilde{\gamma}(s, t)A^*_1(s)\big)^{-1}A_1(s)\varphi^s_t\Big]^{-1} \\ &\quad \cdot\Big[m(t) + \gamma(t) \sum_{s=t}^{\tau-1} (\varphi^s_t)^* A^*_1(s)\big(B\circ B(s) + A_1(s)\tilde{\gamma}(s, t)A^*_1(s)\big)^{-1} \\ &\qquad\qquad \cdot\big(\xi_{s+1} - A_0(s) - A_1(s)\psi^s_t\big)\Big], \end{aligned} \tag{5.23}$$

$$\begin{aligned} \gamma(t, \tau) = \Big[E + \gamma(t) \sum_{s=t}^{\tau-1} (\varphi^s_t)^* A^*_1(s)\big(B\circ B(s) \\ + A_1(s)\tilde{\gamma}(s, t)A^*_1(s)\big)^{-1}A_1(s)\varphi^s_t\Big]^{-1} \gamma(t), \end{aligned} \tag{5.24}$$

where φ^s_t, ψ^s_t and $\tilde{\gamma}(s, t)$ are defined by formulas (5.3), (5.4) *and* (5.5).

PROOF. We first show that for all $\tau > t$

$$\gamma(\tau - 1) = \tilde{\gamma}(\tau - 1, t) + \varphi^{\tau-1}_t \gamma(t, \tau - 1)(\varphi^{\tau-1}_t)^*. \tag{5.25}$$

In fact,

$$\begin{aligned} \gamma(\tau - 1) &= \operatorname{Cov}(\theta_{\tau-1} \mid \mathscr{F}^{\xi}_{\tau-1}) = \mathbf{M}\{[\theta_{\tau-1} - m(\tau - 1)][\theta_{\tau-1} - m(\tau - 1)]^* \mid \mathscr{F}^{\xi}_{\tau-1}\} \\ &= \mathbf{M}\{[\theta_{\tau-1} - m_{\theta_t}(\tau - 1, t) + m_{\theta_t}(\tau - 1, t) - m(\tau - 1)] \\ &\qquad \cdot[\theta_{\tau-1} - m_{\theta_t}(\tau - 1, t) + m_{\theta_t}(\tau - 1, t) - m(\tau - 1)]^* \mid \mathscr{F}^{\xi}_{\tau-1}\} \\ &= \mathbf{M}\{\mathbf{M}[(\theta_{\tau-1} - m_{\theta_t}(\tau - 1, t))(\theta_{\tau-1} - m_{\theta_t}(\tau - 1, t))^* \mid \mathscr{F}^{\xi}_{\tau-1}, \theta_t] \mid \mathscr{F}^{\xi}_{\tau-1}\} \\ &\quad + \mathbf{M}\{[m_{\theta_t}(\tau - 1, t) - m(\tau - 1)][m_{\theta_t}(\tau - 1, t) - m(\tau - 1)]^* \mid \mathscr{F}^{\xi}_{\tau-1}\} \\ &= \mathbf{M}\{\tilde{\gamma}(\tau - 1, t) \mid \mathscr{F}^{\xi}_{\tau-1}\} \\ &\qquad + \mathbf{M}\{\varphi^{\tau-1}_t(\theta_t - m(t, \tau - 1))(\theta_t - m(t, \tau - 1))^*(\varphi^{\tau-1}_t)^* \mid \mathscr{F}^{\xi}_{\tau-1}\} \\ &= \tilde{\gamma}(\tau - 1, t) + \varphi^{\tau-1}_t \gamma(t, \tau - 1)\varphi^{\tau-1}_t, \end{aligned} \tag{5.26}$$

where we have made use of the fact that according to (5.16)

$$m(\tau - 1) = \varphi^{\tau-1}_t m(t, \tau - 1) + \psi^{\tau-1}_t. \tag{5.27}$$

From (5.13) and (5.25) we have

$$\begin{aligned}(5.28)\quad \gamma(t,\tau) = \gamma(t,\tau-1) &- \gamma(t,\tau-1)(\varphi_t^{\tau-1})^* A_1^*(\tau-1)[B\circ B(\tau-1)\\ &+ A_1(\tau-1)\tilde{\gamma}(\tau-1,t)A_1^*(\tau-1)\\ &+ A_1(\tau-1)\varphi_t^{\tau-1}\gamma(t,\tau-1)(\varphi_t^{\tau-1})^* A_1^*(\tau-1)]^{-1}\\ &\cdot A_1(\tau-1)\varphi_t^{\tau-1}\gamma(t,\tau-1).\end{aligned}$$

Let us put

$$(5.29)\quad \begin{aligned}\tilde{A}_1(\tau-1) &= A_1(\tau-1)\varphi_t^{\tau-1},\\ \widetilde{B\circ B}(\tau-1) &= B\circ B(\tau-1) + A_1(\tau-1)\tilde{\gamma}(\tau-1,t)A_1^*(\tau-1).\end{aligned}$$

Then with respect to τ, the function $\gamma(t,\tau)$ will satisfy the equation

$$\begin{aligned}(5.30)\quad \gamma(t,\tau) = \gamma(t,\tau-1) &- \gamma(t,\tau-1)\tilde{A}_1^*(\tau-1)\\ &\cdot[\widetilde{B\circ B}(\tau-1) + \tilde{A}_1(\tau-1)\gamma(t,\tau-1)\tilde{A}_1^*(\tau-1)]^{-1}\\ &\cdot\tilde{A}_1(\tau-1)\gamma(t,\tau-1),\end{aligned}$$

which is similar to equation (5.4) for $\gamma(\tau)$. Taking account of the notation (5.29), we obtain the required representation (5.24) from (5.5).

Along with (5.29), let us put $\tilde{A}_0(\tau-1) = A_0(\tau-1) + A_1(\tau-1)\psi_t^{\tau-1}$. Then equation (5.12) assumes the following form (with τ replaced by $\tau-1$):

$$\begin{aligned}(5.31)\quad m(t,\tau) = m(t,\tau-1) &+ \gamma(t,\tau-1)\tilde{A}_1^*(\tau-1)[B\circ B(\tau-1)\\ &+ \tilde{A}_1(\tau-1)\gamma(t,\tau-1)\tilde{A}_1^*(\tau-1)]^{-1}\\ &\cdot[\xi_\tau - \tilde{A}_0(\tau-1) - \tilde{A}_1(\tau-1)m(t,\tau-1)].\end{aligned}$$

Comparing this equation with (5.3), whose solution is given by formula (5.6), we obtain the representation (5.23) for $m(t,\tau)$.

We now proceed to the consideration of another class of interpolation problems, consisting in the construction of the best (in the mean sense) estimates of the vector θ_t in terms of observations of $\xi^\tau = (\xi_0,\cdots,\xi_\tau)$ and the known value $\theta_\tau = \beta$, $\tau > t$. We put

$$(5.32)\quad \begin{aligned}\Pi_{\alpha,\beta}(t,\tau) &= P\{\theta_t \leqq \alpha \mid \mathscr{F}_\tau^\xi, \theta_\tau = \beta\},\\ m_\beta(t,\tau) &= \mathbf{M}(\theta_t \mid \mathscr{F}_\tau^\xi, \theta_\tau = \beta),\\ \gamma_\beta(t,\tau) &= \mathrm{Cov}\,(\theta_t \mid \mathscr{F}_\tau^\xi, \theta_\tau = \beta)\end{aligned}$$

THEOREM 5.4. *If the distribution* $\Pi_\alpha(t) = P\{\theta_t \leqq \alpha \mid \mathscr{F}_t^\xi\}$ *is normal, then the* a posteriori *distribution* $\Pi_{\alpha,\beta}(t,\tau)$ *is normal for all* $\tau \geqq t$.

PROOF. Let us compute the conditional characteristic function

$$\begin{aligned}(5.33)\quad &\mathbf{M}\{\exp i[\langle Z,\theta_t\rangle + \langle \bar{Z},\theta_\tau\rangle] \mid \mathscr{F}_\tau^\xi\}\\ &= \mathbf{M}\{\exp i[\langle Z,\theta_t\rangle]\mathbf{M}[\exp i\langle \bar{Z},\theta_\tau\rangle \mid \mathscr{F}_\tau^\xi,\theta_t] \mid \mathscr{F}_\tau^\xi\},\end{aligned}$$

where $Z = (Z_1,\cdots,Z_k)$, $\bar{Z} = (\bar{Z}_1,\cdots,\bar{Z}_k)$. According to Remark 3.1 to Theorem

3.2, the distribution $P\{\theta_\tau \leqq \beta \mid \theta_t, \mathscr{F}_\tau^\xi\}$ is normal, $\mathscr{N}(m_{\theta_t}(\tau, t), \gamma_{\theta_t}(\tau, t))$, where by (5.2) $m_{\theta_t}(\tau, t) = \varphi_t^\tau \theta_t + \psi_t^\tau$, and the covariance $\gamma_{\theta_t}(\tau, t)$ does not depend upon θ_t: $\gamma_{\theta_t}(\tau, t) = \tilde{\gamma}(\tau, t)$. Therefore,

$$\mathbf{M}\{\exp i\langle \bar{Z}, \theta_\tau\rangle \mid \theta_t, \mathscr{F}_\tau^\xi\} = \exp\left[i\langle \bar{Z}, \varphi_t^\tau \theta_t + \psi_t^\tau\rangle - \tfrac{1}{2}\langle \tilde{\gamma}(\tau, t)\bar{Z}, \bar{Z}\rangle\right], \tag{5.34}$$

and consequently,

$$\begin{aligned}\mathbf{M}\{\exp i[\langle Z, \theta_t\rangle + \langle \bar{Z}, \theta_\tau\rangle] \mid \mathscr{F}_\tau^\xi\} &= \exp i[\langle \bar{Z}, \psi_t^\tau\rangle - \tfrac{1}{2}\langle \tilde{\gamma}(\tau, t)\bar{Z}, \bar{Z}\rangle] \\ &\quad \cdot \mathbf{M}\{\exp i[\langle Z, \theta_t\rangle + \langle \bar{Z}, \varphi_t^\tau \theta_t\rangle] \mid \mathscr{F}_\tau^\xi\}.\end{aligned} \tag{5.35}$$

But the conditional distribution $\Pi_\alpha(t, \tau)$ is normal. Therefore, by virtue of (5.35) and Theorem 3.1, the distributions $P\{\theta_t \leqq \alpha, \theta_\tau \leqq \beta \mid \mathscr{F}_\tau^\xi\}$ and $P\{\theta_t \leqq \alpha \mid \theta_\tau = \beta, \mathscr{F}_\tau^\xi\}$ are normal.

The method used in the proof of Theorem 5.4 can be applied to establish the assertion to the effect that the process (θ_t, ξ_t), $t = 0, 1, \cdots$ is conditionally Gaussian (see Theorem 3.3).

PROOF OF THEOREM 3.3. Let us consider the conditional characteristic function:

$$\begin{aligned}&\mathbf{M}\left\{\exp i\left[\sum_{k=0}^{t} \langle Z_k, \theta_k\rangle\right] \Big| \mathscr{F}_t^\xi\right\} \\ &= \mathbf{M}\left\{\left[\exp i \sum_{k=0}^{t-1} \langle Z_k, \theta_k\rangle\right] \mathbf{M}[\exp i\langle Z_t, \theta_t\rangle \mid \theta_0, \cdots, \theta_t, \mathscr{F}_t^\xi] \mid \mathscr{F}_t^\xi\right\} \\ &= \mathbf{M}\left\{\left[\exp i \sum_{k=0}^{t-1} \langle Z_k, \theta_k\rangle\right] \mathbf{M}[\exp i\langle Z_t, \theta_t\rangle \mid \theta_{t-1}, \mathscr{F}_t^\xi] \mid \mathscr{F}_t^\xi\right\} \\ &= \mathbf{M}\left\{\exp i\left[\sum_{k=0}^{t-2} \langle Z_k, \theta_k\rangle + \langle Z_{t-1}, \theta_{t-1}\rangle\right.\right. \\ &\qquad \left.\left. + \langle Z_t, \varphi_{t-1}^t \theta_{t-1} + \psi_{t-1}^t\rangle\right] \Big| \mathscr{F}_t^\xi\right\} \cdot \exp\left[-\tfrac{1}{2}\langle \tilde{\gamma}(t, t-1) Z_t, Z_t\rangle\right] \\ &= \exp\left[-\tfrac{1}{2}\langle \tilde{\gamma}(t, t-1) Z_t, Z_t\rangle\right] \cdot \mathbf{M}\left\{\exp i\left(\sum_{k=0}^{t-2} \langle Z_k, \theta_k\rangle\right)\right. \\ &\qquad \left. \cdot \mathbf{M}[\exp i\langle Z_{t-1} + (\varphi_{t-1}^t)^* Z_t, \theta_{t-1}\rangle \mid \mathscr{F}_t^\xi, \theta_{t-2}] \cdot \exp\left[i\langle Z_t, \psi_{t-1}^t\rangle\right]\right\},\end{aligned} \tag{5.36}$$

where the conditional characteristic function of the "split off" variable θ_{t-1} is given by

$$\begin{aligned}&\mathbf{M}\{\exp[i(\langle Z_{t-1} + (\varphi_{t-1}^t)^* Z_t, \theta_{t-1}\rangle + \langle Z_{t-1}, \psi_{t-1}^t\rangle)] \mid \mathscr{F}_t^\xi, \theta_{t-2}\} \\ &= \mathbf{M}\{\exp[i(\langle Z_{t-1} + (\varphi_{t-1}^t)^* Z_t, \theta_{t-1}\rangle] \mid \mathscr{F}_t^\xi, \theta_{t-2}\} \exp(i\langle Z_t, \psi_{t-1}^t\rangle).\end{aligned} \tag{5.37}$$

The distribution $P\{\theta_{t-1} \leqq \beta \mid \mathscr{F}_t^\xi, \theta_{t-2}\}$ is normal, and by (5.12) and (5.13) its mean depends linearly upon θ_{t-2}, while its covariance does not depend upon θ_{t-2}.

Thus, we have in (5.37) (compare (5.34))

$$\begin{aligned}(5.38)\quad & \mathbf{M}\{\exp[i(\langle Z_{t-1} + (\varphi^t_{t-1})^* Z_t, \theta_{t-1}\rangle)] \mid \mathscr{F}^\xi_t, \theta_{t-2}\} \\ &= \exp[i\langle Z_{t-1} + (\varphi^t_{t-1})^* Z_t, a(t-1, t-2)\theta_{t-2} + b(t-1, t-2)\rangle \\ &\quad -\tfrac{1}{2}\langle c(t-1, t-2)(Z_{t-1} + (\varphi^t_{t-1})^* Z_t), (Z_{t-1} + (\varphi^t_{t-1})^* Z_t)\rangle],\end{aligned}$$

where the explicit form of the functions a, b, and c, which depend only upon ξ^t and time, is not of importance to us here. What is essential is that θ_{t-} enters linearly in the exponential, and Z_t, Z_{t-1} enter quadratically.

Thus,

$$\begin{aligned}(5.39)\quad & \mathbf{M}\left[\exp i \sum_{k=0}^{t} \langle Z_k, \theta_k\rangle \mid \mathscr{F}^\xi_t\right] = \exp[-\tfrac{1}{2}\langle \tilde{\gamma}(t, t-1) Z_t, Z_t\rangle \\ &\quad -\tfrac{1}{2}\langle c(t-1, t-2)(Z_{t-1} + (\varphi^t_{t-1})^* Z_t), (Z_{t-1} + (\varphi^t_{t-1})^* Z_t)\rangle] \\ &\cdot \mathbf{M}\left[\exp i\left(\sum_{k=0}^{t-2} \langle Z_k, \theta_k\rangle + \langle Z_{t-1} + (\varphi^t_{t-1})^* Z_t, a(t-1, t-2)\theta_{t-2}\right.\right. \\ &\qquad \left.\left. + b(t-1, t-2)\rangle\right) \mid \mathscr{F}^\xi_t\right].\end{aligned}$$

Continuing this method of "splitting off" the variables, we see that the characteristic function $\mathbf{M}(\exp\{i\, \Sigma^t_{k=0} \langle Z_k, \theta_k\rangle\} \mid \mathscr{F}^\xi_t)$ has the form of the exponential of a quadratic form in the variables $Z_0, \cdots, Z_k$, which proves that the process (θ_t, ξ_t), $t = 0, 1, \cdots$ is conditionally Gaussian.

THEOREM 5.5. *If the distribution $\Pi_\alpha(t)$ is normal, then the parameters $m_\beta(t, \tau)$ and $\gamma_\beta(t, \tau)$ of the normal distribution $\Pi_{\alpha,\beta}(t, \tau)$ are defined for all $\tau > t$ by*

$$(5.40)\qquad m_\beta(t, \tau) = m(t, \tau) + \gamma(t, \tau)(\varphi^\tau_t)^* \gamma^+(\tau)(\beta - m(\tau)),$$

$$(5.41)\qquad \gamma_\beta(t, \tau) = \gamma(t, \tau) - \gamma(t, \tau)(\varphi^\tau_t)^* \gamma^+(\tau) \varphi^\tau_t \gamma(t, \tau),$$

where $m_\beta(t, t) = \beta$, $\gamma_\beta(t, t) = 0$.

PROOF. The conditional distribution $P\{\theta_t \leqq \alpha, \theta_\tau \leqq \beta \mid \mathscr{F}^\xi_\tau\}$ is normal. Therefore, by the Remark 2.1,

$$(5.42)\qquad m_\beta(t, \tau) = \mathbf{M}(\theta_t \mid \mathscr{F}^\xi_\tau, \theta_\tau = \beta) = \mathbf{M}(\theta_t \mid \mathscr{F}^\xi_\tau) + d_{1,2} d^+_{2,2}(\beta - \mathbf{M}(\theta_\tau \mid \mathscr{F}^\xi_\tau))$$

and

$$(5.43)\qquad \gamma_\beta(t, \tau) = d_{1,1} - d_{1,2} d^+_{2,2} d_{1,2},$$

where

$$(5.44)\qquad \begin{aligned} d_{1,1} &= \operatorname{Cov}(\theta_t \mid \mathscr{F}^\xi_\tau) = \gamma(t, \tau), \\ d_{2,2} &= \operatorname{Cov}(\theta_\tau \mid \mathscr{F}^\xi_\tau) = \gamma(\tau), \\ d_{1,2} &= \operatorname{Cov}(\theta_t, \theta_\tau \mid \mathscr{F}^\xi_\tau). \end{aligned}$$

By Lemma 5.1 and formula (5.16),

$$\begin{aligned}(5.45)\qquad & \mathbf{M}[(\theta_\tau - m(\tau))^* \mid \mathscr{F}^\xi_\tau, \theta_t] \\ &= \theta^*_t(\varphi^\tau_t)^* + (\psi^\tau_t)^* - [m^*(t, \tau)(\varphi^\tau_t)^* + (\psi^\tau_t)^*] \\ &= (\theta_t - m(t, \tau))^*(\varphi^\tau_t)^*.\end{aligned}$$

Therefore,

$$(5.46)\quad \begin{aligned} d_{1,2} &= \mathrm{Cov}\,(\theta_t, \theta_\tau | \mathscr{F}_\tau^\xi) = \mathbf{M}[(\theta_t - m(t,\tau))(\theta_\tau - m(\tau))^* | \mathscr{F}_\tau^\xi] \\ &= \mathbf{M}\{(\theta_t - m(t,\tau))\mathbf{M}[(\theta_\tau - m(\tau))^* | \mathscr{F}_\tau^\xi, \theta_t] | \mathscr{F}_\tau^\xi\} \\ &= d_{1,1}(\varphi_t^\tau)^* = \gamma(t,\tau)(\varphi_t^\tau)^*. \end{aligned}$$

We obtain the required representations (5.40) and (5.41) from (5.42) through (5.46).

REMARK 5.2. It follows from (5.41) that the covariance $\gamma_\beta(t, \tau)$ does not depend upon β. We shall denote it by $\bar{\gamma}(t, \tau)$.

6. Inverse equations for interpolation

To derive inverse equations (with respect to t for fixed $\tau \geqq t$) for $m(t, \tau)$, $\gamma(t, \tau)$, $m_\beta(t, \tau)$ and $\bar{\gamma}(t, \tau)$, we need the following auxiliary results.

LEMMA 6.1. *For the process* (θ_t, ξ_t), $t = 0, 1, \cdots$, *defined by the equations* (3.1), *we have with probability* 1 *for all* $t < s \leqq \tau$

$$(6.1)\quad P\{\theta_t \leqq \alpha | \mathscr{F}_\tau^\xi, \theta_s, \theta_{s+1}, \cdots \theta_\tau\} = P\{\theta_t \leqq \alpha | \mathscr{F}_s^\xi, \theta_s\}.$$

The proof follows from the fact that for the process being considered, the right side of (6.1) satisfies the same relations as does the conditional probability $P\{\theta_t \leqq \alpha | \mathscr{F}_\tau^\xi, \theta_s, \theta_{s+1}, \cdots, \theta_\tau\}$.

LEMMA 6.2. *For the process* (θ_t, ξ_t), $t = 0, 1, \cdots$, *defined by the equations* (3.1), *we have for all* $t < s \leqq \tau$

$$(6.2)\quad \Pi_{\alpha,\beta}(t, \tau) = \mathbf{M}\{\Pi_{\alpha,\theta_s}(t, s) | \mathscr{F}_\tau^\xi, \theta_\tau = \beta\}.$$

The proof follows at once from (6.1).

THEOREM 6.1. *Suppose that assumptions* (a) *through* (d) *are fulfilled. Then the moments* $m_\beta(t, \tau)$ *and* $\bar{\gamma}(t, \tau)$ *of the conditional distribution* $\Pi_{\alpha,\beta}(t, \tau)$ *obey, for fixed* τ, *the inverse equations*

$$(6.3)\quad \begin{aligned} m_\beta(t, \tau) &= m(t, t+1) \\ &\quad + \gamma(t, t+1)(\varphi_t^{t+1})^* \gamma^+(t+1)[m_\beta(t+1, \tau) - m(t+1)], \end{aligned}$$

$$(6.4)\quad \begin{aligned} \bar{\gamma}(t, \tau) &= \bar{\gamma}(t, t+1) \\ &\quad + \gamma(t, t+1)(\varphi_t^{t+1})^* \gamma^+(t+1)\bar{\gamma}(t+1, \tau)\gamma^+(t+1)\varphi_t^{t+1}\gamma(t, t+1), \end{aligned}$$

with $m_\beta(\tau, \tau) = \beta$, $\bar{\gamma}(\tau, \tau) = 0$.

PROOF. From (5.40) and (5.41),

$$(6.5)\quad m_\beta(t, t+1) = m(t, t+1) + \gamma(t, t+1)(\varphi_t^{t+1})^* \gamma^+(t+1)(\beta - m(t+1)),$$

$$(6.6)\quad \bar{\gamma}(t, t+1) = \gamma(t, t+1) - \gamma(t, t+1)(\varphi_t^{t+1})^* \gamma^+(t+1)\varphi_t^{t+1}\gamma(t, t+1).$$

By Lemma 6.2, $m_\beta(t, \tau) = \mathbf{M}\{m_{\theta_{t+1}}(t, t+1) | \mathscr{F}_\tau^\xi, \theta_\tau = \beta\}$, which together with (6.5) leads to equation (6.3).

Further, according to (6.1) and (6.2),

$$(6.7)\quad \begin{aligned}\gamma_\beta(t,\tau) &= \mathrm{Cov}\,(\theta_t\,|\,\mathscr{F}_\tau^\xi, \theta_\tau = \beta)\\ &= \mathbf{M}[\mathrm{Cov}\,(\theta_t\,|\,\mathscr{F}_\tau^\xi, \theta_\tau = \beta, \theta_{t+1})\,|\,\mathscr{F}_\tau^\xi, \theta_\tau = \beta]\\ &\qquad + \mathrm{Cov}\,\{\mathbf{M}(\theta_t\,|\,\mathscr{F}_\tau^\xi, \theta_\tau = \beta, \theta_{t+1})\,|\,\mathscr{F}_\tau^\xi, \theta_\tau = \beta\}\\ &= \mathbf{M}[\mathrm{Cov}\,(\theta_t\,|\,\mathscr{F}_{t+1}^\xi, \theta_{t+1})\,|\,\mathscr{F}_\tau^\xi, \theta_\tau = \beta]\\ &\qquad + \mathrm{Cov}\,[\mathbf{M}(\theta_t\,|\,\mathscr{F}_{t+1}^\xi, \theta_{t+1})\,|\,\mathscr{F}_\tau^\xi, \theta_\tau = \beta]\\ &= \mathbf{M}[\gamma_{\theta_{t+1}}(t, t+1)\,|\,\mathscr{F}_\tau^\xi, \theta_\tau = \beta]\\ &\qquad + \mathrm{Cov}\,[m_{\theta_{t+1}}(t, t+1)\,|\,\mathscr{F}_\tau^\xi, \theta_\tau = \beta]\\ &= \bar{\gamma}(t, t+1) + \mathbf{M}[(m_{\theta_{t+1}}(t, t+1) - m_\beta(t,\tau))(m_{\theta_{t+1}}(t, t+1)\\ &\qquad\qquad - m_\beta(t,\tau))^*\,|\,\mathscr{F}_\tau^\xi, \theta_\tau = \beta].\end{aligned}$$

But from (6.3) and (6.5), we have

$$(6.8)\quad \begin{aligned}&m_{\theta_{t+1}}(t, t+1) - m_\beta(t,\tau)\\ &\qquad = \gamma(t, t+1)(\varphi_t^{t+1})^*\gamma^+(t+1)(\theta_{t+1} - m_\beta(t+1,\tau)).\end{aligned}$$

Therefore,

$$(6.9)\quad \begin{aligned}\gamma_\beta(t,\tau) = \bar{\gamma}(t, t+1) + \gamma(t, t+1)(\varphi_t^{t+1})^*\gamma^+(t+1)\gamma_\beta(t+1,\tau)&\\ \cdot\gamma^+(t+1)\varphi_t^{t+1}\gamma(t, t+1)&,\end{aligned}$$

which establishes equation (6.4), since $\gamma_\beta(t,\tau)$ does not depend upon $\beta(\gamma_\beta(t,\tau) = \bar{\gamma}(t,\tau))$.

THEOREM 6.2. *Suppose that assumptions* (a) *through* (d) *are fulfilled. Then for fixed* τ *the moments* $m(t,\tau)$ *and* $\gamma(t,\tau)$ *of the conditional distribution* $\Pi_\alpha(t,\tau)$ *satisfy the inverse equations*

$$(6.10)\quad \begin{aligned}m(t,\tau) &= m(t, t+1)\\ &\qquad + \gamma(t, t+1)(\varphi_t^{t+1})^*\gamma^+(t+1)[m(t+1,\tau) - m(t+1)],\end{aligned}$$

$$(6.11)\quad \begin{aligned}\gamma(t,\tau) &= \bar{\gamma}(t, t+1)\\ &\qquad + \gamma(t, t+1)(\varphi_t^{t+1})^*\gamma^+(t+1)\gamma(t+1,\tau)\\ &\qquad\qquad \cdot\gamma^+(t+1)\varphi_t^{t+1}\gamma(t, t+1),\end{aligned}$$

with $m(\tau,\tau) = m(\tau)$, $\gamma(\tau,\tau) = \gamma(\tau)$.

PROOF. Equation (6.10) follows from (6.3). To derive equation (6.11), we note that, by (6.1) and (6.2),

$$(6.12)\quad \begin{aligned}\gamma(t,\tau) &= \mathrm{Cov}\,(\theta_t\,|\,\mathscr{F}_\tau^\xi)\\ &= \mathbf{M}[\mathrm{Cov}\,(\theta_t\,|\,\mathscr{F}_\tau^\xi, \theta_{t+1})\,|\,\mathscr{F}_\tau^\xi] + \mathrm{Cov}\,[\mathbf{M}(\theta_t\,|\,\mathscr{F}_\tau^\xi, \theta_{t+1})\,|\,\mathscr{F}_\tau^\xi]\\ &= \mathbf{M}[\mathrm{Cov}\,(\theta_t\,|\,\mathscr{F}_{t+1}^\xi, \theta_{t+1})\,|\,\mathscr{F}_\tau^\xi] + \mathrm{Cov}\,[\mathbf{M}(\theta_t\,|\,\mathscr{F}_{t+1}^\xi, \theta_{t+1})\,|\,\mathscr{F}_\tau^\xi]\\ &= \bar{\gamma}(t, t+1) + \mathbf{M}\{[m_{\theta_{t+1}}(t, t+1) - m(t,\tau)]\\ &\qquad\qquad \cdot[m_{\theta_{t+1}}(t, t+1) - m(t,\tau)]^*\,|\,\mathscr{F}_\tau^\xi\}.\end{aligned}$$

From (6.3) and (6.10), we have

$$(6.13)\quad \begin{aligned} &m_{\theta_{t+1}}(t, t+1) - m(t, \tau) \\ &\qquad = \gamma(t, t+1)(\varphi_t^{t+1})^* \gamma^+(t+1)(\theta_{t+1} - m(t+1, \tau)). \end{aligned}$$

Therefore, by virtue of Lemma 6.2,

$$(6.14)\quad \begin{aligned} &\mathbf{M}\{[m_{\theta_{t+1}}(t, t+1) - m(t, \tau)][m_{\theta_{t+1}}(t, t+1) - m(t, \tau)]^* \mid \mathscr{F}_\tau^\xi\} \\ &\quad = \gamma(t, t+1)(\varphi_t^{t+1})^* \gamma^+(t+1)\gamma(t+1, \tau)\gamma^+(t+1)\varphi_t^{t+1}\gamma(t, t+1). \end{aligned}$$

We obtain equation (6.11) from (6.12) and (6.14).

7. Examples of the use of the interpolation equations

EXAMPLE 7.1. *Interpolating the values of a Gaussian Markov chain.* Let $\theta_t = (\theta_1(t), \cdots, \theta_k(t))$, $t = 0, 1, \cdots$, be a Markov chain governed by the recursive equations

$$(7.1)\qquad \theta_{t+1} = a_0(t) + a_1(t)\theta_t + b(t)\varepsilon(t+1),$$

where the Gaussian vectors $\varepsilon(t+1)$, as well as the vector function $a_0(t)$ and the matrices $a_1(t)$, $b(t)$, which depend only upon t, are the same as in system (3.1). The random vector θ_0 is assumed normal with $\mathbf{M}\theta_0 = m$, $\operatorname{Cov}\theta_0 = \gamma$.

We consider the problem of finding the best (in the mean square sense) estimate of the quantity θ_t under the assumption that $\theta_\tau = \beta$, $t < \tau$.

Equation (6.14) can be regarded as a special case of the system (3.32), consisting of two equations with $A_0 = 0, A_1 = 0, B_1 = 0, B_2 = 0, b_2 = 0$. Therefore, the quantities $m(t, \tau)$, $\gamma(t, \tau)$, $m_\beta(t, \tau)$ and $\gamma_\beta(t, \tau)$ introduced above maintain their meaning. For the case (7.1),

$$(7.2)\quad m(t, \tau) = m(t) = \mathbf{M}\theta_t,\ \gamma(t, \tau) = \gamma(t) = \mathbf{M}[(\theta_t - m(t))(\theta_t - m(t))^*],$$

and

$$(7.3)\quad \begin{aligned} m_\beta(t, \tau) &= \mathbf{M}(\theta_t \mid \theta_\tau = \beta), \\ \gamma_\beta(t, \tau) &= \mathbf{M}[(\theta_t - m_\beta(t, \tau))(\theta_t - m_\beta(t, \tau))^* \mid \theta_\tau = \beta]. \end{aligned}$$

By Theorem 3.2 the quantities $m(t)$ and $\gamma(t)$ can be determined from the linear equations

$$(7.4)\qquad m(t+1) = a_0(t) + a_1(t)m(t),$$

$$(7.5)\qquad \gamma(t+1) = a_1(t)\gamma(t)a_1^*(t) + b(t)b^*(t).$$

On the basis of Theorem 6.1,

$$(7.6)\qquad m_\beta(t, \tau) = m(t) + \gamma(t)(\varphi_t^\tau)^*\gamma^+(\tau)(\beta - m(\tau)),$$

$$(7.7)\qquad \bar{\gamma}(t, \tau) = \gamma(t) - \gamma(t)(\varphi_t^\tau)^*\gamma^+(\tau)\varphi_t^\tau\gamma(t),$$

where

$$\varphi_t^\tau = \prod_{s=t}^{\tau-1} a_1(s) = a_1(\tau - 1) \cdots a_1(t), \tag{7.8}$$

$$\varphi_t^t = E, \qquad \bar{\gamma}(t, \tau) = \gamma_\beta(t, \tau).$$

Inverse equations for $m_\beta(t, \tau)$ and $\bar{\gamma}(t, \tau)$ can be determined from (6.3) and (6.4):

$$m_\beta(t, \tau) = m(t) + \gamma(t)a_1^*(t)\gamma^+(t + 1), \tag{7.9}$$

$$\bar{\gamma}(t, \tau) = \bar{\gamma}(t, t + 1) + \gamma(t)a_1^*(t)\gamma^+(t + 1)\bar{\gamma}(t + 1, \tau)\gamma^+(t + 1)a_1(t)\gamma(t), \tag{7.10}$$

where $t + 1 \leqq \tau$ and $m_\beta(\tau, \tau) = \beta$, $\bar{\gamma}(\tau, \tau) = 0$.

In particular, if

$$\theta_{t+1} = \theta_t + \varepsilon(t + 1), \tag{7.11}$$

$k = 1$, $m(0) = m$, $\gamma(0) = \gamma$, then from (7.4) and (7.5) we find that $m(t) = m$, $\gamma(t) = t + \gamma$, while from (7.6) and (7.7) we obtain

$$\begin{aligned} m_\beta(t, \tau) &= m + \frac{t + \gamma}{\tau + \gamma}(\beta - m), \\ \bar{\gamma}(t, \tau) &= (t + \gamma) - \frac{(t + \gamma)^2}{\tau + \gamma}. \end{aligned} \tag{7.12}$$

These explicit formulas are also not difficult to obtain from the inverse equations (7.9) and (7.10).

EXAMPLE 7.2. *Interpolation with a fixed lag.* Let $\tau = t + h$, where h is some fixed number. The direct and inverse equations for $m(t, \tau)$ and $\gamma(t, \tau)$ obtained above enable us to find equations (with respect to t) for $m(t, t + h) = \mathbf{M}(\theta_t | \mathscr{F}_{t+h}^\xi)$ and $\gamma(t, t + h) = \mathrm{Cov}\,(\theta_t | \mathscr{F}_{t+h}^\xi)$. The quantity $m(t, t + h)$ determines the optimal interpolation of the quantity θ_t in the presence of a fixed lag h, that is in terms of observations of ξ^{t+h}.

Let us put $m_h(t) = m(t, t + h)$ and $\gamma_h(t) = \gamma(t, t + h)$. Suppose that for all $t = 0, 1, \cdots$, the matrix $\gamma(t, t + 1)(\varphi_t^{t+1})^*\gamma^{-1}(t + 1)$ is nonsingular.

From the direct equation (5.12), we have

$$\begin{aligned} m_h(t + 1) = m(t + 1, t + h) &+ \gamma(t + 1, t + h)(\varphi_{t+1}^{t+h})^* A_1^*(t + h) \\ &\cdot [B \circ B(t + h) + A_1(t + h)\gamma(t + h)A_1^*(t + h)]^+ \\ &\cdot [\xi_{t+h+1} - A_0(t + h) - A_1(t + h)m(t + h)]. \end{aligned} \tag{7.13}$$

From the inverse equation (7.4), under the assumption that the matrix $\gamma(t, t + 1)(\varphi_t^{t+h})^*\gamma^{-1}(t + h)$ is nonsingular, we find that

$$\begin{aligned} m(t + 1, t + h) = m(t + 1) &+ [\gamma(t, t + 1)(\varphi_t^{t+h})^*\gamma^{-1}(t + 1)]^{-1} \\ &\cdot [m_h(t) - m(t, t + 1)]. \end{aligned} \tag{7.14}$$

Thus, from (7.13) and (7.14), we obtain

$$(7.15)\quad \begin{aligned} m_h(t+1) &= m(t+1) \\ &+ [\gamma(t, t+1)(\varphi_t^{t+1})^* \gamma^{-1}(t+1)]^{-1}[m_h(t) - m(t, t+1)] \\ &+ \gamma(t+1, t+h)(\varphi_t^{t+1})^* A_1^*(t+h) \\ &\cdot [B \circ B(t+h) + A_1(t+h)\gamma(t+h)A_1^*(t+h)]^+ \\ &\cdot [\xi_{t+h+1} - A_0(t+h) - A_1(t+h)m(t+h)]. \end{aligned}$$

Let us now consider equation (5.13) for $\gamma_h(t+1) = \gamma(t+1, t+1+h)$:

$$(7.16)\quad \begin{aligned} \gamma_h(t+1) &= \gamma(t+1, t+h) - \gamma(t+1, t+h)(\varphi_{t+1}^{t+h})^* A_1(t+h) \\ &\cdot [B \circ B(t+h) + A_1(t+h)\gamma(t+h)A_1^*(t+h)]^+ \\ &\cdot A_1(t+h)\varphi_{t+h}^{t+h}\gamma(t+1, t+h). \end{aligned}$$

From the inverse equation (7.5),

$$(7.17)\quad \begin{aligned} \gamma(t+1, t+h) &= [\gamma(t, t+1)(\varphi_t^{t+1})^* \gamma^{-1}(t+1)]^{-1}[\gamma_h(t) - \bar{\gamma}(t, t+1)] \\ &\cdot [\gamma^{-1}(t+1)\varphi_t^{t+1}\gamma(t, t+1)]^{-1}. \end{aligned}$$

Inserting the expression (7.17) for $\gamma(t+1, t+h)$ into (7.15) and (7.16), we obtain direct equations for $m_h(t)$ and $\gamma_h(t)$, where $m_h(0) = m(0, h)$ and $\gamma_h(0) = \gamma(0, h)$ can be found from equations (5.12) and (5.13).

For the special case $h = 1$, we find from (7.15) that

$$(7.18)\quad \begin{aligned} m_1(t+1) &= m(t+1) + \gamma(t+1)A_1^*(t+1) \\ &\cdot [B \circ B(t+1) + A_1(t+1)\gamma(t+1)A_1^*(t+1)]^+ \\ &\cdot [\xi_{t+2} - A_0(t+1) - A_1(t+1)m(t+1)]. \end{aligned}$$

8. Extrapolation equations

By extrapolation is understood the problem of estimating the vectors θ_τ, ξ_τ by means of observations of $\xi^t = (\xi_0, \cdots, \xi_t)$, where $\tau > t$. As in the case of interpolation, one can consider equations for the optimal estimates in terms of τ as well as of t. The equations in terms of τ make it possible to understand how the prediction (the extrapolation) deteriorates as τ increases, while the equations in terms of t enable us to ascertain the degree to which the quality of the prediction improves with the amount of data, that is, as $t \uparrow \tau$.

In contrast with $\Pi_\beta(t)$, $\Pi_\alpha(t, \tau)$, and $\Pi_{\alpha,\beta}(t, \tau)$, the distributions

$$(8.1)\quad \mathfrak{P}_{\beta,x}(\tau, t) = P\{\theta_\tau \leqq \beta, \xi_\tau \leqq x \mid \mathscr{F}_t^\xi\}, \qquad \tau > t$$

for $\tau > t + 1$, generally speaking, are not Gaussian. (The distribution $\mathfrak{P}_x(t+1, t)$ is Gaussian, as was established in the proof of Theorem 3.1.) This circumstance complicates the construction of optimal estimates for extrapolation for the general case of system (3.1.). We shall consider below two special cases of system (3.1) in which we are able, nevertheless, to obtain equations for the optimal estimates of the vectors θ_τ and ξ_τ.

Before proceeding to the statement of the theorem, let us clarify in what way one can distinguish those cases in which we are able to construct estimates for extrapolation.

By virtue of (3.18) and (3.56),

$$m(\tau+1) = [a_0(\tau) + a_1(\tau)m(\tau)] + [b\circ b(\tau) + a_1(\tau)\gamma(\tau)A_1^*(\tau)] \cdot [B\circ B(\tau) + A_1(\tau)\gamma(\tau)A_1^*(\tau)]^+ \cdot [B\circ B(\tau) + A_1(\tau)\gamma(\tau)A_1^*(\tau)]^{1/2}\bar{\varepsilon}(\tau+1) \tag{8.2}$$

$$\xi_{\tau+1} = [A_0(\tau) + A_1(\tau)m(\tau)] + [B\circ B(\tau) + A_1(\tau)\gamma(\tau)A_1^*(\tau)]^{1/2}\bar{\varepsilon}(\tau+1). \tag{8.3}$$

We put $n_1(\tau, t) = \mathbf{M}(\theta_\tau \mid \mathscr{F}_t^\xi)$, $n_2(\tau, t) = \mathbf{M}(\xi_\tau \mid \mathscr{F}_t^\xi)$, the optimal (in the mean square sense) estimates for extrapolation. Then since

$$n_1(\tau, t) = \mathbf{M}\{\mathbf{M}(\theta_\tau \mid \mathscr{F}_\tau^\xi) \mid \mathscr{F}_t^\xi\} = \mathbf{M}(m(\tau) \mid \mathscr{F}_t^\xi), \tag{8.4}$$

and $\mathbf{M}(\bar{\varepsilon}(\tau+1) \mid \mathscr{F}_t^\xi) = 0$, it follows that we can attempt to determine $n_1(\tau, t)$ and $n_2(\tau, t)$ by applying $\mathbf{M}(\cdot \mid \mathscr{F}_t^\xi)$ to both sides of (8.2) and (8.3).

It is not hard to see that joint determination of $n_1(\tau, t)$ and $n_2(\tau, t)$ is possible, if we assume that

$$a_0(t, \omega) = a_0(t) + a_2(t)\xi_t, \qquad a_1(t, \omega) = a_1(t), \tag{8.5}$$

$$A_0(t, \omega) = A_0(t) + A_2(t)\xi_t, \qquad A_1(t, \omega) = A_1(t), \tag{8.6}$$

where a_2 is a $k \times \ell$, and A_2 an $\ell \times \ell$ matrix.

If we are only interested in extrapolating θ_τ, then determination of $n_1(\tau, t)$ becomes possible if we require that (8.5) be fulfilled with $a_2(t) \equiv 0$.

THEOREM 8.1. (i) *Suppose that* assumptions (a) *through* (d), and (6.3), (6.4) *are fulfilled. Then the moments* $n_1(\tau, t)$ *and* $n_2(\tau, t)$ *satisfy the systems of equations*

$$n_1(\tau+1, t) = a_0(\tau) + a_1(\tau)n_1(\tau, t) + a_2(\tau)n_2(\tau, t), \tag{8.7}$$

$$n_2(\tau+1, t) = A_0(\tau) + A_1(\tau)n_1(\tau, t) + A_2(\tau)n_2(\tau, t) \tag{8.8}$$

with $n_1(t, t) = m(t)$, $n_2(t, t) = \xi_t$. (ii) *Suppose that assumptions* (a) *through* (d), *and* (8.5) *with* $a_2(t) \equiv 0$ *are fulfilled. Then*

$$n_1(\tau+1, t) = a_0(\tau) + a_1(\tau)n_2(\tau, t), \qquad n_1(t, t) = m(t). \tag{8.9}$$

The proof follows at once by averaging both sides of equations (8.2) and (8.3).

In the theorem just below, we give inverse equations (with respect to t) for $n_1(\tau, t)$ and $n_2(\tau, t)$.

THEOREM 8.2. (i) *Suppose that assumptions* (a) *through* (d), *as well as* (8.5) *and* (8.6), *are fulfilled. Then the following equations are valid.*

$$\begin{pmatrix} n_1(\tau, t+1) \\ n_2(\tau, t+1) \end{pmatrix} = \begin{pmatrix} n_1(\tau, t) \\ n_2(\tau, t) \end{pmatrix} + \Phi_{t+1}^{\tau-1} \begin{pmatrix} u_\tau v_\tau^+ \\ E \end{pmatrix} \cdot [\xi_{t+1} - A_0(t) - A_1(t)m(t) - A_2(t)\xi_t], \tag{8.10}$$

In these relations $E = E(\ell \times \ell)$. *The pair* (u_τ, v_τ) *is given by* $u_\tau = [b \circ B(\tau) + a_1(\tau)\gamma(\tau)A_1^*(\tau)]$ *and* $v_\tau = [B \circ B(\tau) + A_1(\tau)\gamma(\tau)A_1^*(\tau)]$. *The matrix* Φ_t^τ *is determined from the recursive equation*

$$\Phi_t^\tau = \begin{pmatrix} a_1(\tau - 1) & a_2(\tau - 1) \\ A_1(\tau - 1) & A_2(\tau - 1) \end{pmatrix} \Phi_t^{\tau - 1}, \qquad \Phi_t^t = E((k + \ell) \times (k + \ell)), \tag{8.11}$$

and

$$\begin{pmatrix} n_1(\tau, 0) \\ n_2(\tau, 0) \end{pmatrix} = \Phi_0^\tau \begin{pmatrix} m(0) \\ \xi_0 \end{pmatrix} + \sum_{s=0}^{\tau-1} \Phi_s^{\tau-1} \begin{pmatrix} a_0(s) \\ A_0(s) \end{pmatrix}. \tag{8.12}$$

(ii) *Suppose that assumptions* (a) *through* (d), *as well as* (8.5) *with* $a_2(t) \equiv 0$, *are fulfilled. Then*

$$n_1(\tau, t + 1) = n_1(\tau, t) + \psi_{t+1}^{\tau-1} u_t v_t^+ [\xi_{t+1} - A_0(t) - A_1(t)m(t)], \tag{8.13}$$

where $u_t = [b \circ B(t) + a_1(t)\gamma(t)A_1^*(t)]$ *and* $v_t = [B \circ B(t) + A_1(t)\gamma(t)A_1^*(t)]$. *The matrix* ψ_t^τ *is determined from the equations* $\psi_t^\tau = a_1(\tau - 1)\psi_t^{\tau-1}$, $\psi_t^t = E(k \times k)$ *and*

$$n_1(\tau, 0) = \psi_0^\tau m(0) + \sum_{s=0}^{\tau-1} \psi_s^{\tau-1} a_0(s). \tag{8.14}$$

PROOF. By induction, we obtain from (8.2) and (8.3),

$$\begin{pmatrix} m(\tau) \\ \xi_\tau \end{pmatrix} = \Phi_0^\tau \begin{pmatrix} m(0) \\ \xi_0 \end{pmatrix} + \sum_{s=0}^{\tau-1} \Phi_s^{\tau-1} \begin{pmatrix} a_0(s) \\ A_0(s) \end{pmatrix} + \sum_{s=0}^{\tau-1} \Phi_s^{\tau-1} \begin{pmatrix} u_s v_s^+ v_s^{1/2} \\ v_s^{1/2} \end{pmatrix} \bar{\varepsilon}(s + 1), \tag{8.15}$$

where $u_s = [b \circ B(s) + a_1(s)\gamma(s)A_1^*(s)]$ and $v_s = [B \circ B(s) + A_1(s)\gamma(s)A_1^*(s)]$.

Applying $\mathbf{M}(\cdot | \mathscr{F}_{t+1}^\xi)$ to both sides of (8.15), we find that

$$\begin{pmatrix} n_1(\tau, t + 1) \\ n_2(\tau, t + 1) \end{pmatrix} = \Phi_0^\tau \begin{pmatrix} m(0) \\ \xi_0 \end{pmatrix} + \sum_{s=0}^{\tau-1} \Phi_s^{\tau-1} \begin{pmatrix} a_0(s) \\ A_0(s) \end{pmatrix} + \sum_{s=0}^{t} \Phi_s^{\tau-1} \begin{pmatrix} u_s v_s^+ v_s^{1/2} \\ v_s^{1/2} \end{pmatrix} \bar{\varepsilon}(s + 1), \tag{8.16}$$

where u_s and v_s are as in (8.15), since the vector $v_s^{1/2}\bar{\varepsilon}(s + 1)$ for $s \leqq t$ is $\mathscr{F}_{t+1}^\xi$ measurable, $\mathbf{M}(\bar{\varepsilon}(s + 1) | \mathscr{F}_{t+1}^\xi) = 0$, $s > t$, and $\Phi_s^{\tau-1}$ does not depend upon "chance".

From (8.10), we obtain

$$\begin{pmatrix} n_1(\tau, t + 1) \\ n_2(\tau, t + 1) \end{pmatrix} = \begin{pmatrix} n_1(\tau, t) \\ n_2(\tau, t) \end{pmatrix} + \Phi_{t+1}^\tau \begin{pmatrix} u_t v_t^+ v_t^{1/2} \\ v_t^{1/2} \end{pmatrix} \bar{\varepsilon}(t + 1), \tag{8.17}$$

which together with (3.56) leads to equation (8.10).

To derive equation (8.13), we note that by (8.2),

$$m(\tau + 1) = [a_0(\tau) + a_1(\tau)m(\tau)] + u_\tau v_\tau^+ v_\tau^{1/2}\bar{\varepsilon}(t + 1), \tag{8.18}$$

Hence, by induction,

$$m(\tau) = \psi_0^\tau m(0) + \sum_{s=0}^{\tau-1} \psi_s^{\tau-1} a_0(s) + \sum_{s=0}^{\tau-1} \psi_s^{\tau-1} u_s v_s^+ v_s^{1/2}\bar{\varepsilon}(s + 1). \tag{8.19}$$

Applying $\mathbf{M}(\cdot \mid \mathscr{F}_{t+1}^\xi)$ to both sides of (8.19), as was done to (8.16), we obtain

$$n_1(\tau, t + 1) = \psi_0^\tau m(0) + \sum_{s=0}^{\tau-1} \psi_s^{\tau-1} a_0(s) + \sum_{s=0}^{t} \psi_s^{\tau-1} u_s v_s^+ v_s^{1/2}\bar{\varepsilon}(s + 1). \tag{8.20}$$

Taking account of formula (3.56), we obtain equation (8.13).

We now show how the theory discussed above can be applied to the construction of the optimal linear prediction of a stationary sequence.

Let ξ_t, $t = 0, \pm 1, \pm 2, \cdots$ be a wide sense stationary process with $\mathbf{M}\xi_t = 0$ and spectral density

$$f(\lambda) = \frac{|e^{i\lambda} + 1|^2}{|e^{2i\lambda} + \frac{1}{2}e^{i\lambda} + \frac{1}{2}|^2} \tag{8.21}$$

The problem consists of constructing the optimal (in the mean square sense) linear estimates of the random variable ξ_τ in terms of the results of observations on $\xi^t = (\xi_0, \cdots, \xi_t)$, where the process ξ_t admits the spectral representation

$$\xi_t = \int_{-\pi}^{\pi} e^{i\lambda t} \frac{e^{i\lambda} + 1}{e^{2i\lambda} + \frac{1}{2}e^{i\lambda} + \frac{1}{2}} \Phi(d\lambda), \tag{8.22}$$

in which $\Phi(\cdot)$ is an orthogonal random measure with $\mathbf{M}\Phi(d\lambda) = 0$, $\mathbf{M}|\Phi(d\lambda)|^2 = d\lambda/2\pi$.

We construct a Gaussian random process ξ_t for $t = 0, \pm 1, \pm 2, \cdots$ with spectral representation (8.22). Such a process can be obtained as the solution of the equation

$$\xi_{t+2} + \tfrac{1}{2}(\xi_{t+1} + \xi_t) = \varepsilon(t + 2) + \varepsilon(t + 1), \tag{8.23}$$

where $\varepsilon(t)$, $t = 0, \pm 1, \pm 2, \cdots$ is a sequence of independent Gaussian random variables with $\mathbf{M}\varepsilon(t) = 0$, $\mathbf{M}\varepsilon^2(t) = 1$, $t = 0, \pm 1, \pm 2, \cdots$.

Put $\theta_t = \xi_{t+1} - \varepsilon(t + 1)$. Then we obtain for (θ_t, ξ_t), $t = 0, \pm 1, \pm 2, \cdots$, the system of equations

$$\begin{aligned} \theta_{t+1} &= -\tfrac{1}{2}\theta_t - \tfrac{1}{2}\xi_t + \tfrac{1}{2}\varepsilon(t + 1), \\ \xi_{t+1} &= \theta_t + \varepsilon(t + 1). \end{aligned} \tag{8.24}$$

By Theorem 8.1, $n_1(\tau, t) = \mathbf{M}(\theta_\tau \mid \mathscr{F}_t^\xi)$ and $n_2(\tau, t) = \mathbf{M}(\xi_\tau \mid \mathscr{F}_t^\xi)$ can be determined from the equations

$$\begin{aligned} n_1(\tau + 1, t) &= -\tfrac{1}{2}n_1(\tau, t) - \tfrac{1}{2}n_2(\tau, t), \\ n_2(\tau + 1, t) &= n_1(\tau, t), \end{aligned} \tag{8.25}$$

where $n_1(t, t) = m(t)$, $n_2(t, t) = \xi_t$.

The initial condition $m(t) = \mathbf{M}(\theta_t | \mathscr{F}_t^\xi)$ which enters in (8.25), together with $\gamma(t)$, can be determined from the equations

$$
\begin{aligned}
m(t+1) &= -\tfrac{1}{2}m(t) - \tfrac{1}{2}\xi_t + \frac{1-\gamma(t)}{2(1+\gamma(t))}[\xi_{t+1} - m(t)], \\
\gamma(t+1) &= \frac{\gamma(t)}{1+\gamma(t)}
\end{aligned} \tag{8.26}
$$

(see (3.18) and (3.19)).

We shall show that $m(0) = 0$, $\gamma(0) = 1$ in the system (8.26). Indeed, by virtue of the stationarity of the process (θ_t, ξ_t), the parameters $d_{1,1} = \mathbf{M}\theta_t^2$, $d_{1,2} = \mathbf{M}\theta_t\xi_t$, and $d_{2,2} = \mathbf{M}\xi_t^2$ can be found from the following system of equations, which is easily obtained from (8.24),

$$
\begin{aligned}
d_{1,1} &= \tfrac{1}{4}d_{1,1} + \tfrac{1}{4}d_{2,2} + \tfrac{1}{2}d_{1,2} + \tfrac{1}{4}, \\
d_{1,2} &= -\tfrac{1}{2}d_{1,1} - \tfrac{1}{2}d_{1,2} + \tfrac{1}{2}, \\
d_{2,2} &= d_{1,1} + 1.
\end{aligned} \tag{8.27}
$$

Hence, $d_{1,1} = 1$, $d_{1,2} = 0$, $d_{2,2} = 2$, and by virtue of the theorem on normal correlation $m(0) = 0$, $\gamma(0) = 1$.

REFERENCES

[1] F. R. GANTMACHER, *Theory of Matrices*, Moscow, "Nauka," 1966.

[2] O. A. GLONTI, "Sequential filtering and interpolation of the components of a Markov chain," *Litovsk. Mat. Sb.*, Vol. 9 (1969), pp. 263–279.

[3] ———, "Extrapolation of the components of a Markov chain," *Litovsk. Mat. Sb.*, Vol. 9 (1969), pp. 741–754.

[4] T. KAILATH, "An innovations approach to least-squares estimation," *IEEE Trans. Automatic Control*, Vol. 13 (1968), pp. 646–660.

[5] G. KALLIANPUR and C. STRIEBEL, "Stochastic differential equations occurring in the estimation of continuous parameter stochastic processes," *Theor. Probability Appl.*, Vol. 14 (1969), pp. 567–595.

[6] R. E. KALMAN, "A new approach to linear filtering and prediction problems," *Trans. ASME Ser. D. J. Basic Engrg.*, Vol. 82 (1960), pp. 35–45.

[7] R. E. KALMAN and R. S. BUCY, "New results in linear filtering and prediction theory," *Trans. ASME Ser. D. J. Basic Engrg.*, Vol. 83 (1961), pp. 95–107.

[8] R. SH. LIPTSER and A. N. SHIRYAYEV, "Nonlinear filtration of diffusion Markov processes," *Trudy Mat. Inst. Steklov*, Vol. 104 (1968), pp. 135–180.

[9] ———, "Non-linear interpolation of components of Markov diffusion processes (direct equations, effective formulas)," *Theor. Probability Appl.*, Vol. 13 (1968), pp. 564–583.

[10] ———, "The extrapolation of multidimensional Markov processes from incomplete data," *Theor. Probability Appl.*, Vol. 13 (1968), pp. 15–38.

[11] N. E. NAHI, *Estimation Theory and Applications*, New York, Wiley, 1969.

[12] *Proceedings of the Ninth Joint Conference on Automatic Control*, Ann Arbor, 1968.

[13] R. L. STRATONOVICH, *Conditional Markov Processes and Their Application to the Theory of Optimal Control*, Modern Analytic and Computational Methods in Science and Mathematics, No. 7, New York, Elsevier, 1968.