KYBERNETIKA — VOLUME 12 (1976), NUMBER 3

Algebraic Methods in Discrete Linear Estimation

VLADIMÍR KUČERA

In a series of recent papers the author has developed a new algebraic theory of discrete linear control. It is based exclusively on algebraic properties of polynomials and the synthesis procedure reduces to solving a linear Diophantine equation in polynomials. In this paper an application of these algebraic methods to the minimum variance filtering, prediction and smoothing of scalar random sequences is considered.

INTRODUCTION

In his pioneering work, Wiener [36] formulated the problems of prediction of random signals and separation of random signals from random noise. He showed that these problems lead to the so-called Wiener-Hopf integral equation and gave a method of solution in the important case of stationary statistics and rational spectra. It is the method of spectral factorization in the complex domain [3; 6; 28; 32; 35].

Kalman [13; 14] generalized his results to the nonstationary case using the time-domain approach. He obtained the optimal filter and predictor through the solution of nonlinear matrix Riccati differential equation [1; 5; 11; 12; 16; 30; 34]. The optimal smoothing was subsequently solved in [4; 8; 12; 30; 33]. For the stationary Wiener problem the Riccati equation reduces to a matrix quadratic algebraic equation [23; 24; 25; 26].

The Wiener's solution is conceptually simple but the numerical computations are quite involved and poorly suited to machine processing. The Kalman's solution of the stationary filtering can be easily algorithmized but solving matrix quadratic equations is not a simple task [23; 24; 25; 26].

The algebraic approach presented here is applicable to stationary random sequences with rational spectra and leads to solving a linear Diophantine equation in polynomials. This is not only conceptually simpler but also yields a very efficient computational algorithm. The method was originally developed to solve optimal control problems [17; 18; 19; 20; 21; 22] and its application to optimal filtering was inspired by the duality principle of Kalman [13; 15; 16].

The problems considered in the paper can be loosely described as follows. Let S be an observed mixture of a random signal W and some random noise N. Find a linear system $\mathscr{F}$ whose output at time k recovers in an optimal way the value of W at time l, where l may be less than, equal to, or greater than k; see Fig. 1. If $l < k$, this is a *smoothing* problem. If $l = k$, this is called *filtering*. If $l > k$, we have a ***prediction*** problem. A collective term for the three problems is *estimation*.

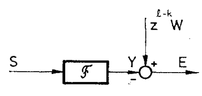

Fig. 1. The optimal estimation.

POLYNOMIALS AND RATIONAL FUNCTIONS

For detail discussion of algebraic concepts used below, the reader is referred to [16; 17; 20; 27; 38].

Let $\mathfrak{R}$ be the field of reals and $\mathfrak{R}(z^{-1})$ the field of *rational functions* over $\mathfrak{R}$. An element $A \in \mathfrak{R}(z^{-1})$ can be written as

$$A = \alpha_n z^{-n} + \alpha_{n+1} z^{-(n+1)} + \dots, \qquad \alpha_k \in \mathfrak{R}$$

for some integer n. If $\alpha_n \neq 0$, then n is the order of A. We denote $\langle A \rangle = \alpha_0$, the coefficient of A at z^0.

The elements of $\mathfrak{R}(z^{-1})$ with nonnegative order form the ring $\mathfrak{R}\{z^{-1}\}$ of *realizable rational functions.* They have the form

$$A = \alpha_0 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots, \qquad \alpha_k \in \mathfrak{R}. \tag{1}$$

Elements (1) for which the sequence $\{\alpha_0, \alpha_1, \alpha_2, \dots\}$ converges to zero constitute the ring of *stable realizable rational functions* denoted by $\mathfrak{R}^+\{z^{-1}\}$. Let

$$A = \alpha_0 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots \in \mathfrak{R}^+\{z^{-1}\}.$$

Viewing $\mathfrak{R}^+\{z^{-1}\}$ as a vector space over $\mathfrak{R}$, the norm $\|A\|$ of $A \in \mathfrak{R}^+\{z^{-1}\}$ can be defined by

$$\|A\|^2 = \sum_{k=0}^{\infty} |\alpha_k|^2. \tag{2}$$

With the notation

$$\bar{A} = \alpha_0 + \alpha_1 z + \alpha_2 z^2 + \dots$$

the norm (2) can be expressed as the inner product

$$\|A\|^2 = \langle \bar{A}A \rangle .$$

Elements (1) with only a finite number of nonzero coefficients form the ring $\mathfrak{R}[z^{-1}]$ of *polynomials* over $\mathfrak{R}$. If

$$(3) \qquad a = \alpha_0 + \alpha_1 z^{-1} + \ldots + \alpha_n z^{-n} \in \mathfrak{R}[z^{-1}]$$

and $\alpha_n \neq 0$, then n is the degree of a, denoted by ∂a. By convention, $\partial 0 = -\infty$. If $a, b \in \mathfrak{R}[z^{-1}]$, we write $b \mid a$ to denote that b divides a and (a, b) for the greatest common divisor of a and b. A polynomial $a \in \mathfrak{R}[z^{-1}]$ is said to be stable if $1/a \in \in \mathfrak{R}^+\{z^{-1}\}$.

Given a nonzero polynomial $a \in \mathfrak{R}[z^{-1}]$, we define the factorization [17, 20]

$$(4) \qquad a = a^+ a^- ,$$

where a^+ is the stable factor of a having highest degree and belonging to $\mathfrak{R}[z^{-1}]$.

Given a polynomial (3) of degree $n \geqq 0$, we define the polynomials [17, 20]

$$(5) \qquad \tilde{a} = z^{-n}\bar{a} = \alpha_0 z^{-n} + \alpha_1 z^{-(n-1)} + \ldots + \alpha_n$$

and

$$(6) \qquad a^* = a^+ \tilde{a}^- .\dagger$$

Note that $\bar{a}a = \bar{a}^* a^*$.

DIOPHANTINE EQUATIONS

Consider the equation

$$(7) \qquad ax + by = c$$

for unknown polynomials x, y and given polynomials a, b, $c \in \mathfrak{R}[z^{-1}]$. This equation has been called a linear Diophantine equation in polynomials, see [10; 31].

It is shown in [17; 31] that equation (7) has a solution if and only if $(a, b) \mid c$. If x_1, y_1 is a solution of (7), then

$$x = x_1 + b_0 t$$
$$y = y_1 - a_0 t$$

is also a solution of (7), where

$$a_0 = \frac{a}{(a, b)}, \quad b_0 = \frac{b}{(a, b)}$$

and t is an arbitrary polynomial of $\mathfrak{R}[z^{-1}]$.

† For typographical reasons, the symbols $\tilde{a}^-$, $\tilde{a}^*$, $\bar{a}^*$, etc. are used in place of $\widetilde{a^-}$, $\widetilde{a^*}$, $\overline{a^*}$, etc. throughout.

A particular solution of equation (7) can be effectively found via the Euclidean algorithm, see [10; 17;, 31]. This algorithm is very simple, fast, and well-adapted for machine processing.

In applications we often seek for a particular solution x_0, y_0 of equation (7) that satisfies $\partial y_0 < \partial a$. This solution is not unique unless $(a, b) = 1$ and can be found by applying the division algorithm.

SYSTEMS AND RANDOM SEQUENCES

Throughout the paper we shall consider finite-dimensional discrete linear constant single-input single-output systems defined over the field $\mathfrak{R}$. They are described by the equations

$$x_{k+1} = Ax_k + Bu_k \qquad (8)$$
$$y_k = Cx_k + Du_k\,, \quad k = 0, 1, \ldots.$$

for the input $u \in \mathfrak{R}$, output $y \in \mathfrak{R}$ and the state vector $x \in \mathfrak{R}^n$. System (8) is said to be stable if $A^k \to 0$ for $k \to \infty$.

The sequence

$$S = D + z^{-1}C(I_n - z^{-1}A)^{-1}\,B \in \mathfrak{R}\{z^{-1}\} \qquad (9)$$

is called the impulse response of the system. Conversely, any quadruple $\{A, B, C, D\}$ satisfying (9) is a realization of S; if A is of least possible size, the realization is minimal [16, 20].

The S can be written as the ratio of two polynomials [17; 20; 22]

$$S = \frac{b}{a}$$

which satisfy $(a, b) = 1$, $(a, z^{-1}) = 1$.

For convenience, we review here some elementary facts about random sequences. The details could be found in [7; 9; 12; 28; 29] or [2; 6; 32; 35].

A random variable over $\mathfrak{R}$ is a function whose values belong to $\mathfrak{R}$ and depend on the outcome of a chance event. A random variable α can be defined by stating the probability that α is less than or equal to some constant $\alpha \in \mathfrak{R}$,

$$\mathsf{P}(\alpha \leqq \alpha) = F_\alpha(\alpha)\,.$$

The expectation of a random variable α is defined by the integral

$$\mathsf{E}\alpha = \int_{-\infty}^{\infty} \alpha dF_\alpha(\alpha)\,.$$

A sequence of random variables

$$A = \{\ldots, \alpha_0, \alpha_1, \alpha_2, \ldots\} \tag{10}$$

is called a random sequence over $\mathfrak{R}$. The k function with values $\mathsf{E}\alpha_k$ is called the mean value function of A. The l, m function whose values are $\mathsf{E}\alpha_l\alpha_m$ is the correlation function of A. If

$$B = \{\ldots, \beta_0, \beta_1, \beta_2, \ldots\}$$

is another random sequence, the l, m function with values $\mathsf{E}\alpha_l\beta_m$ is the cross-correlation function of A and B (in this order).

A random sequence (10) is said to be (weakly) stationary if $\mathsf{E}\alpha_k$ is independent of k and $\mathsf{E}\alpha_l\alpha_m$ depends only on $l-m$ and is bounded. A random sequence (10) is called white (uncorrelated) if

$$\begin{aligned} \mathsf{E}(\alpha_l - \mathsf{E}\alpha_l)(\alpha_m - \mathsf{E}\alpha_m) &= 1\,, \quad l = m\,, \\ &= 0\,, \quad l \neq m\,. \end{aligned}$$

In most cases, observed random sequences are not white. The correlation between random sequences observed at different times is usually explained by the presence

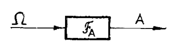

Fig. 2. Random sequence model.

of a dynamic system between the primary random source and the observer. Thus a random sequence A may be thought of as the output of a system $\mathscr{F}_A$ excited by a white random sequence Ω, see Fig. 2. The $\mathscr{F}_A$ is usually called the shaping filter of A.

Throughout the paper we shall restrict ourselves to random sequences which may be observed at times $k = 0, 1, \ldots$ at the output of a shaping filter governed by equations (8) and having a zero-mean white random sequence applied at its input at time $k = -\infty$.

It follows from the definition that such a random sequence

$$A = \{\ldots, \alpha_0, \alpha_1, \alpha_2, \ldots\}\,, \quad \mathsf{E}\alpha_k = 0$$

is stationary if and only if the shaping filter $\mathscr{F}_A$ is stable. Then the correlation function of A can be written as

$$\Phi_{AA} = \ldots + \varphi_{-1}z + \varphi_0 + \varphi_1 z^{-1} + \ldots\,,$$

where

$$\varphi_k = \mathsf{E}\alpha_{k+l}\alpha_l = \varphi_{-k}\,,$$

and $\Phi_{AA} = \bar{F}_A F_A$. The φ_0 is called the variance of A and, since

$$\varphi_0 = \langle \Phi_{AA} \rangle = \langle \bar{F}_A F_A \rangle = \|F_A\|^2 ,$$

it can be interpretted as the squared norm of the impulse response of its shaping filter.

If

$$B = \{\ldots, \beta_0, \beta_1, \beta_2, \ldots\} , \quad \mathsf{E}\beta_k = 0$$

is another zero-mean stationary random sequence, the cross-correlation function of A and B can be written as

$$\Phi_{AB} = \ldots + \psi_{-1} z + \psi_0 + \psi_1 z^{-1} + \ldots ,$$

where

$$\psi_k = \mathsf{E}\alpha_{k+l}\,\beta_l ,$$

and $\Phi_{AB} = \bar{F}_L F_R$ for some $F_L, F_R \in \mathfrak{R}^+\{z^{-1}\}$. Evidently, $\Phi_{BA} = \Phi_{AB}$.

Now consider a random sequences $S = A \pm B$ obtained as the sum (difference) of two stationary random sequences A and B. Then [6, 28; 29; 32; 35]

$$(11) \qquad \begin{aligned} \Phi_{SS} &= \Phi_{AA} \pm \Phi_{AB} \pm \Phi_{BA} + \Phi_{BB} , \\ \Phi_{AS} &= \Phi_{AA} \pm \Phi_{AB} . \end{aligned}$$

If a stationary random sequence E passes through a stable system $\mathscr{G}$ with impulse response G to yield a random sequence A, i.e. $A = GE$, then [6; 28; 29; 32; 35]

$$(12) \qquad \begin{aligned} \Phi_{AA} &= G\Phi_{EE}G , \\ \Phi_{EA} &= \Phi_{EE}G \end{aligned}$$

in the steady state.

MINIMUM VARIANCE ESTIMATION

As stated in the Introduction, the observed random sequence S is a mixture of a random signal W to be recovered and a random noise N. Since the noise is not explicitely known, the most natural initial data to solve the estimation problems would be the four impulse responses $F_W, F_S, F_L, F_R \in \mathfrak{R}^+\{z^{-1}\}$, which define the random characteristics

$$(13) \qquad \begin{aligned} \Phi_{WW} &= \bar{F}_W F_W , \quad & \Phi_{WS} &= \bar{F}_R F_L \\ \Phi_{SW} &= \bar{F}_L F_R , \quad & \Phi_{SS} &= \bar{F}_S F_S \end{aligned}$$

of S and W. The relations between F_W, F_S, F_L and F_R depend on the type of interaction between the signal W and noise N. In the case of additive interaction, $S = W + N$, a necessary and sufficient condition for the F_W, F_S, F_L and F_R to be meaningfull is the existence of a matrix F with elements from $\mathfrak{R}^+\{z^{-1}\}$ such that

(14)
$$\begin{bmatrix} \Phi_{WW} & \Phi_{WN} \\ \Phi_{NW} & \Phi_{NN} \end{bmatrix} = \bar{F}F' ,$$

where

$$\Phi_{WN} = \Phi_{WS} - \Phi_{WW} ,$$
$$\Phi_{NW} = \Phi_{SW} - \Phi_{WW} ,$$
$$\Phi_{NN} = \Phi_{SS} - \Phi_{SW} - \Phi_{WS} + \Phi_{WW}$$

and F' is the transpose of F. If

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix},$$

we have the model of random sequences W and N shown in Fig. 3.

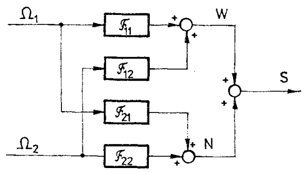

Fig. 3. Additive model of signal and noise.

If the F_W, F_S, F_L and F_R happen not to satisfy condition (14), they correspond to some non-additive interaction between W and N. However, this has no effect on the solution of estimation problems.

For our estimation to be optimal, we have to define some optimality criterion. Since we want the optimal estimator to be linear, our object will be to minimize the steady-state variance of the estimation error E, which is to say, to minimize the norm $\|F_E\|^2$ of the impulse response of the shaping filter for E.

Hence we have the following formal formulations:

(15) *Minimum variance filtering*

Given the configuration of Fig. 1, where the zero-mean stationary random sequences W and S are characterized by

$$F_W = \frac{q}{p} \in \mathfrak{R}^+\{z^{-1}\} ,$$

$$F_S = \frac{b}{a} \in \mathfrak{R}^+\{z^{-1}\} , \quad b \neq 0 ,$$

$$F_L = \frac{v}{u} \in \mathfrak{R}^+\{z^{-1}\},$$

$$F_R = \frac{s}{r} \in \mathfrak{R}^+\{z^{-1}\}$$

and the filtering error is defined as $E = W - Y$.

Find a stable filter $\mathscr{F}$ which is a (not necessarily minimal) realization of some $F \in \mathfrak{R}^+\{z^{-1}\}$ such that the norm $\|F_E\|^2$ of the filtering error is minimized. □

(16) *Minimum variance prediction*

Given the configuration of Fig. 1, where the zero-mean stationary random sequences W and S are characterized by

$$F_W = \frac{q}{p} \in \mathfrak{R}^+\{z^{-1}\},$$

$$F_S = \frac{b}{a} \in \mathfrak{R}^+\{z^{-1}\}, \quad b \neq 0,$$

$$F_L = \frac{v}{u} \in \mathfrak{R}^+\{z^{-1}\},$$

$$F_R = \frac{s}{r} \in \mathfrak{R}^+\{z^{-1}\}$$

and the prediction error for $\lambda = l - k$ stage prediction is defined as $E = (1/z^{-\lambda}) W - Y$.

Find a stable predictor $\mathscr{F}$ which is a (not necessarily minimal) realization of some $F \in \mathfrak{R}^+\{z^{-1}\}$ such that the norm $\|F_E\|^2$ of the prediction error is minimized. □

(17) *Minimum variance smoothing*

Given the configuration of Fig. 1, where the zero-mean stationary random sequences W and S are characterized by

$$F_W = \frac{q}{p} \in \mathfrak{R}^+\{z^{-1}\},$$

$$F_S = \frac{b}{a} \in \mathfrak{R}^+\{z^{-1}\}, \quad b \neq 0,$$

$$F_L = \frac{v}{u} \in \mathfrak{R}^+\{z^{-1}\},$$

$$F_R = \frac{s}{r} \in \mathfrak{R}^+\{z^{-1}\}$$

and the smoothing error for $\mu = k - l$ stage smoothing is defined as $E = z^{-\mu}W - Y$.

Find a stable smoother $\mathscr{F}$ which is a (not necessary minimal) realization of some $F \in \mathfrak{R}^+\{z^{-1}\}$ such that the norm $\|F_E\|^2$ of the smoothing error is minimized. $\square$

The estimation problem could be generalized in the sense that some *function* of the random signal W is to be recovered from S. The estimation error is then defined as $E = KW - Y$, where $K \in \mathfrak{R}(z^{-1})$ represents the desired function of W.

THE FILTERING PROBLEM

Denoting

$$m = \partial av\,, \quad n = \partial b^*u\,, \tag{18}$$

we are ready to prove the following result.

Theorem 1. Problem (15) has a solution if and only if the linear Diophantine equation

$$z^{-m}\tilde{b}^*\tilde{u}x + ry = z^{-n}\tilde{a}\tilde{v}s \tag{19}$$

has a solution x_0, y_0 such that $\partial y_0 < \partial z^{-m}\tilde{b}^*\tilde{u}$ and

$$F = \frac{ax_0}{rb^*} \tag{20}$$

belongs to $\mathfrak{R}^+\{z^{-1}\}$.

The optimal filter $\mathscr{F}$ is given as a stable realization of (20) and it is essentially unique when minimally realized. Moreover,

$$\|F_E\|^2_{\min} = \left\langle \overline{\left(\frac{y_0}{b^*u}\right)}\left(\frac{y_0}{b^*u}\right)\right\rangle + \langle \Phi_{WW} - \Phi_{WS}\Phi_{SS}^{-1}\Phi_{SW}\rangle\,. \tag{21}$$

Proof. In order to minimize the norm $\|F_E\|^2$ of the filtering error E, we shall assume that E is a stationary random sequence. Then $F_E \in \mathfrak{R}^+\{z^{-1}\}$ and

$$\|F_E\|^2 = \langle \bar{F}_E F_E\rangle\,. \tag{22}$$

We will manipulate the inner product $\langle \bar{F}_E F_E\rangle$ so as to make the minimizing choice of F obvious.

Write $E = W - FS$ and, by (11), (12) and (13),

$$\begin{aligned}\bar{F}_E F_E &= \bar{F}_W F_W - \bar{F}_R F_L F - \bar{F}\bar{F}_L F_R + \bar{F}\bar{F}_S F_S F = \\ &= \overline{\left(\frac{\bar{a}}{\bar{b}^*}\frac{\bar{v}s}{\bar{u}r} - \frac{b^*}{a}F\right)}\left(\frac{\bar{a}}{\bar{b}^*}\frac{\bar{v}s}{\bar{u}r} - \frac{b^*}{a}F\right) + \\ &\quad + \Phi_{WW} - \Phi_{WS}\Phi_{SS}^{-1}\Phi_{SW}\,.\end{aligned} \tag{23}$$

Using (5) and (18), we obtain

$$\frac{\bar{a}\bar{v}}{\tilde{b}^*\bar{u}} = \frac{z^{-n}\tilde{a}\tilde{v}}{z^{-m}\tilde{b}^*\tilde{u}}.$$

Since the last two terms on the right-hand side of (23) are independent of F, the $\langle \bar{F}_E F_E \rangle$ attains its minimum for the same F as the $\langle \bar{F}_{E0} F_{E0} \rangle$ does, where

$$F_{E0} = \frac{z^{-n}\tilde{a}\tilde{v}s}{z^{-m}\tilde{b}^*\tilde{u}r} - \frac{b^*}{a}F. \tag{24}$$

Now take the decomposition

$$\frac{z^{-n}\tilde{a}\tilde{v}s}{z^{-m}\tilde{b}^*\tilde{u}r} = \frac{y}{z^{-m}\tilde{b}^*\tilde{u}} + \frac{x}{r}.$$

It follows that the polynomials x and y are coupled by equation (19).

Collecting the terms gives us

$$F_{E0} = \frac{y}{z^{-m}\tilde{b}^*\tilde{u}} + A,$$

where

$$A = \frac{x}{r} - \frac{b^*}{a}F. \tag{25}$$

Hence

$$\langle \bar{F}_{E0} F_{E0} \rangle = \left\langle \overline{\left(\frac{y}{z^{-m}\tilde{b}^*\tilde{u}}\right)} \left(\frac{y}{z^{-m}\tilde{b}^*\tilde{u}}\right) \right\rangle + \tag{26}$$
$$+ \left\langle \overline{\left(\frac{y}{z^{-m}\tilde{b}^*\tilde{u}}\right)} A \right\rangle + \left\langle \bar{A} \left(\frac{y}{z^{-m}\tilde{b}^*\tilde{u}}\right) \right\rangle + \langle \bar{A}A \rangle.$$

Any solution of equation (19) can be written in the form

$$x = x_0 + \frac{r}{(z^{-m}\tilde{b}^*\tilde{u}, r)}\, t \tag{27}$$

$$y = y_0 - \frac{z^{-m}\tilde{b}^*\tilde{u}}{(z^{-m}\tilde{b}^*\tilde{u}, r)}\, t \tag{28}$$

for arbitrary $t \in \mathfrak{R}[z^{-1}]$ and

$$\partial y_0 < \partial z^{-m}\tilde{b}^*\tilde{u}. \tag{29}$$

The key observation is that

$$\overline{\left(\frac{y_0}{z^{-m}\tilde{b}^*\tilde{u}}\right)} = \frac{\tilde{y}_0}{b^*}\, z^{-(\partial z^{-m}\tilde{b}^*\tilde{u} - \partial y_0)}$$

is divisible by z^{-1} due to (29). Therefore

$$(30) \qquad \left\langle \overline{\left(\frac{y_0}{z^{-m}\tilde{b}^*\tilde{u}}\right)} A \right\rangle = 0\,, \quad \left\langle \overline{\left(\frac{y_0}{z^{-m}\tilde{b}^*\tilde{u}}\right)} \frac{t}{(z^{-m}\tilde{b}^*\tilde{u}, r)} \right\rangle = 0$$

and the substitution of (28) into (26) yields

$$(31) \qquad \langle \bar{F}_{E0} F_{E0} \rangle = \left\langle \overline{\left(\frac{y_0}{z^{-m}\tilde{b}^*\tilde{u}}\right)} \left(\frac{y_0}{z^{-m}\tilde{b}^*\tilde{u}}\right) + \right.$$
$$\left. + \left\langle \overline{\left(A - \frac{t}{(z^{-m}\tilde{b}^*\tilde{u}, r)}\right)} \left(A - \frac{t}{(z^{-m}\tilde{b}^*\tilde{u}, r)}\right) \right\rangle \right\rangle$$

since the terms shown in (30) are zero.

The first term on the right-hand side of (31) cannot be affected by any choice of F. The best we can do to minimize (31) is to set

$$A - \frac{t}{(z^{-m}\tilde{b}^*\tilde{u}^*, r)} = 0\,,$$

i.e.

$$\frac{x}{r} - \frac{b^*}{a} F - \frac{t}{(z^{-m}\tilde{b}^*\tilde{u}^*, r)} = 0$$

by (25). But

$$\frac{x}{r} - \frac{t}{(z^{-m}\tilde{b}^*\tilde{u}^*, r)} = \frac{x_0}{r}$$

by virtue of (27). Hence (31) and, in turn, the inner product $\langle \bar{F}_E F_E \rangle$ is minimized by the F given in (20).

If this F belongs to $\mathfrak{R}^+\{z^{-1}\}$, the E is indeed a stationary random sequence and our assumption (22) holds true. Then the optimal filter $\mathcal{F}$ is a stable realization of (20). Since r and u are stable polynomials by assumption, we have $(z^{-m}\tilde{b}^*u^*, r) = 1$ and the solution x_0, y_0 is always unique. As a result, the F is also unique.

Expression (21) for the minimized norm is a direct consequence of (23), (31) and the identity

$$\overline{\left(\frac{y_0}{z^{-m}\tilde{b}^*u^*}\right)} \left(\frac{y_0}{z^{-m}\tilde{b}^*\tilde{u}^*}\right) = \overline{\left(\frac{y_0}{b^*u}\right)} \left(\frac{y_0}{b^*u}\right). \qquad \square$$

Remark 1. When the mixture S is given by its correlation function Φ_{SS} rather than by F_S, the shaping filter F_S is obtained through the decomposition $\Phi_{SS} = \bar{F}_S F_S$. This decomposition is unique [35; 37] up to an element $\omega \in \mathfrak{R}$ satisfying $\omega^2 = 1$. We shall show, however, that the optimal filter $\mathcal{F}$ is independent of ω. Indeed, let

$$F_S = \frac{b}{a}\,, \quad F_{S1} = \frac{b_1}{a}\,,$$

where $b = \omega b_1$. Then $b^* = \omega b_1^*$ and we have to solve the equation

$$z^{-m}\omega\tilde{b}_1^*\tilde{u}x + ry = z^{-n}\tilde{a}\tilde{v}s \tag{32}$$

instead of (19). The solution x_{10}, y_{10} of (32) is coupled with the solution x_0, y_0 of (19) as

$$\omega x_{10} = x_0\,, \quad y_{10} = y_0$$

and, therefore,

$$F_1 = \frac{ax_{10}}{rb_1^*} = \frac{1}{\omega^2}\frac{ax_0}{rb^*} = F\,.$$ □

Remark 2. If the random sequences W and N are not cross-correlated, we have $\Phi_{WN} = \Phi_{NW} = 0$ and hence

$$\Phi_{SW} = \Phi_{WS} = \Phi_{WW}\,,$$

i.e.

$$F_L = F_R = F_W\,.$$

Equation (19) then reads

$$z^{-m}\tilde{b}^*\tilde{p}x + py = z^{-n}\tilde{a}\tilde{q}q \tag{33}$$

and

$$F = \frac{ax_0}{pb^*}\,.$$ □

Remark 3. In case there is no corrupting noise, we have

$$\Phi_{SS} = \Phi_{SW} = \Phi_{WS} = \Phi_{WW}\,,$$

i.e.

$$F_S = F_L = F_R = F_W\,.$$

Equation (19) then reads

$$z^{-m}\tilde{q}^*\tilde{p}x + py = z^{-n}\tilde{p}\tilde{q}q \tag{34}$$

and $m = \partial pq$, $n = \partial q^*p$. In view of the identity

$$\tilde{q}q = z^{-(m-n)}\tilde{q}^*q^*$$

equation (34) can be put into the form

$$x + py_2 = q^*\,,$$

where $y = z^{-m}\tilde{q}^*\tilde{p}y_2$. It follows that $x_0 = q^*$, $y_0 = 0$ satisfies both equation (34) and $\partial y_0 < \partial z^{-m}\tilde{q}^*\tilde{p}$. Thus the optimal filter is simply given by

$$F = \frac{px_0}{pq^*} = 1$$

and $\|F_E\|^2_{\min} = 0$, an expected result. □

Example 1. Let

$$F_W = F_L = F_R = \sqrt{\frac{7}{8}}\,\frac{1}{1 - 0{\cdot}5z^{-1}}\,, \quad F_S = \frac{1}{\sqrt{2}}\,\frac{2 - 0{\cdot}5z^{-1}}{1 - 0{\cdot}5z^{-1}}\,.$$

This means that the S originated as the sum of W and a white noise N not cross-correlated with W. Solve problem (15).

We compute

$$m = 1\,, \quad n = 2\,,$$

$$b^* = \frac{1}{\sqrt{2}}\,(2 - 0{\cdot}5z^{-1})\,, \quad \tilde{b}^* = \frac{1}{\sqrt{2}}(2z^{-1} - 0{\cdot}5)\,,$$

$$p = 1 - 0{\cdot}5z^{-1}\,, \quad q = \sqrt{\frac{7}{8}}\,.$$

Equation (33) becomes

$$z^{-1}\,\frac{1}{\sqrt{2}}(2z^{-1} - 0{\cdot}5)\,(z^{-1} - 0{\cdot}5)\,x + (1 - 0{\cdot}5z^{-1})\,y = z^{-2}(z^{-1} - 0{\cdot}5)\,\frac{7}{8}$$

and using algorithm described in [17] we obtain the solution

$$x_0 = \frac{1}{\sqrt{2}}\,, \quad y_0 = \frac{1}{4}\,z^{-2} - \frac{1}{8}\,z^{-1}$$

satisfying $\partial y_0 < 3$.

Therefore, the optimal filter is a stable realization of

$$F = \frac{1}{2 - 0{\cdot}5z^{-1}}$$

and yields the minimized steady-state variance of the filtering error

$$\|F_E\|^2_{\min} = \frac{1}{30} + \frac{7}{6} - \frac{7}{10} = \frac{1}{2}$$

by virtue of (21). □

Example 2. Consider

$$F_W = \frac{1}{2}\,\frac{1}{1 - 0{\cdot}5z^{-1}}\,, \quad F_S = \frac{\sqrt{3}}{2}\,\frac{1 - z^{-1}}{1 - 0{\cdot}5z^{-1}}\,,$$

$$F_L = \frac{1}{2}\,\frac{1}{1 - 0{\cdot}5z^{-1}}\,, \quad F_R = \frac{1}{2}\,\frac{1 - z^{-1}}{1 - 0{\cdot}5z^{-1}}\,.$$

It means that the signal W is contamined by an additive white noise N, the cross-correlation between W and N being

$$\Phi_{WN} = \frac{-0{\cdot}25z}{(1 - 0{\cdot}5z)(1 - 0{\cdot}5z^{-1})}.$$

Solve problem (15).

Since

$$m = 1, \quad n = 2,$$

$$b^* = \frac{\sqrt{3}}{2}(z^{-1} - 1), \quad \tilde{b}^* = \frac{\sqrt{3}}{2}(1 - z^{-1}),$$

$$\tilde{v} = \tfrac{1}{2}, \quad s = \tfrac{1}{2}(1 - z^{-1}),$$

equation (19) becomes

$$z^{-1}\frac{\sqrt{3}}{2}(1 - z^{-1})(z^{-1} - 0{\cdot}5)\,x + (1 - 0{\cdot}5z^{-1})\,y =$$
$$= z^{-2}(z^{-1} - 0{\cdot}5)\tfrac{1}{2}\tfrac{1}{2}(1 - z^{-1})$$

and yields

$$x_0 = \frac{1}{2\sqrt{3}}z^{-1}, \quad y_0 = 0.$$

Then

$$F = -\frac{1}{3}\frac{z^{-1}}{1 - z^{-1}}$$

is not a stable realizable rational function and hence the problem has no solution. Indeed, the filtering error variance will never reach a steady state. □

Example 3. Let

$$F_W = \frac{1}{2}\frac{1}{1 - 0{\cdot}5z^{-1}}, \quad F_S = \frac{\sqrt{3}}{2}\frac{1 - z^{-1}}{1 - 0.5z^{-1}},$$

$$F_L = \frac{1}{4}\frac{1 - z^{-1}}{1 - 0{\cdot}5z^{-1}}, \quad F_R = \frac{1}{2}\frac{1 - z^{-1}}{1 - 0{\cdot}5z^{-1}}.$$

These data represent a signal W corrupted by an additive white noise N with correlation functions

$$\Phi_{WW} = \frac{0{\cdot}25}{(1 - 0{\cdot}5z)(1 - 0{\cdot}5z^{-1})}, \quad \Phi_{WN} = \frac{-0{\cdot}125z - 0{\cdot}125z^{-1}}{(1 - 0{\cdot}5z)(1 - 0{\cdot}5z^{-1})},$$

$$\Phi_{NW} = \frac{-0{\cdot}125z - 0{\cdot}125z^{-1}}{(1 - 0{\cdot}5z)(1 - 0{\cdot}5z^{-1})}, \quad \Phi_{NN} = 1.$$

$$m = 2\,, \quad n = 2\,,$$

$$b^* = \frac{\sqrt{3}}{2}(z^{-1} - 1)\,, \quad \tilde{b}^* = \frac{\sqrt{3}}{2}(1 - z^{-1})\,,$$

$$\tilde{v} = \tfrac{1}{4}(z^{-1} - 1)\,, \quad s = \tfrac{1}{2}(1 - z^{-1})\,,$$

equation (19) reads

$$z^{-2}\frac{\sqrt{3}}{2}(1 - z^{-1})(z^{-1} - 0{\cdot}5)\,x + (1 - 0{\cdot}5z^{-1})\,y =$$

$$= z^{-2}(z^{-1} - 0{\cdot}5)\,\tfrac{1}{4}(z^{-1} - 1)\,\tfrac{1}{2}(1 - z^{-1})$$

and gives

$$x_0 = \frac{1}{4\sqrt{3}}(z^{-1} - 1)\,, \quad y_0 = 0\,.$$

The optimal filter is then a stable realization of (20)

$$F = \tfrac{1}{6}$$

and yields

$$\|F_E\|^2_{\min} = 0 + \frac{1}{3} - \frac{1}{9} = \frac{2}{9}$$

by virtue of (21).

It is interesting to note that the problem has a solution even though b is divisible by the unstable polynomial $1 - z^{-1}$. The solvability of a given problem thus never can be inferred until the Diophantine equation (19) is solved. □

THE PREDICTION PROBLEM

It is often of interest not only to separate a signal from noise but to predict future values of the separated signal. A sequence W predicted for λ stages forward can be described as $(1/z^{-\lambda})\,W$ and, therefore, the prediction problem (16) can be solved analogously to the filtering problem (15).

Theorem 2. Problem (16) has a solution if and only if the linear Diophantine equation

$$z^{-\lambda-m}\tilde{b}^*\tilde{u}x + ry = z^{-n}\tilde{a}\tilde{v}s \tag{35}$$

has a solution x_0, y_0 such that $\partial y_0 < \partial z^{-\lambda-m}\tilde{b}^*\tilde{u}$ and

$$F = \frac{ax_0}{rb^*} \tag{36}$$

belongs to $\mathfrak{R}^+\{z^{-1}\}$.

The optimal λ-stage predictor $\mathscr{F}$ is given as a stable realization of (36) and it is essentially unique when minimally realized. Moreover,

$$\|F_E\|^2_{\min} = \left\langle \overline{\left(\frac{y_0}{b^*u}\right)} \left(\frac{y_0}{b^*u}\right) \right\rangle + \langle \Phi_{WW} - \Phi_{WS}\Phi_{SS}^{-1}\Phi_{SW} \rangle . \tag{37}$$

Proof. The proof is quite analogous to the proof of Theorem 1. We just write

$$E = \frac{1}{z^{-\lambda}} W - FS$$

and hence (24) reads

$$F_{E0} = \frac{z^{-n}\tilde{a}\tilde{v}s}{z^{-\lambda-m}\tilde{b}^*\tilde{u}r} - \frac{b^*}{a} F .$$

Taking the decomposition

$$\frac{z^{-n}\tilde{a}\tilde{v}s}{z^{-\lambda-m}\tilde{b}^*\tilde{u}r} = \frac{y}{z^{-\lambda-m}\tilde{b}^*\tilde{u}} + \frac{x}{r} ,$$

we obtain equation (35). The optimal predictor (36) and expression (37) follow by repeating the arguments. □

Example 4. Consider again

$$F_W = F_L = F_R = \sqrt{\frac{7}{8}} \frac{1}{1 - 0{\cdot}5z^{-1}} , \quad F_S = \frac{1}{\sqrt{2}} \frac{2 - 0{\cdot}5z^{-1}}{1 - 0{\cdot}5z^{-1}}$$

and solve problem (16) for $\lambda = 1$, i.e. single-stage prediction.

Equation (35) now becomes

$$z^{-2} \frac{1}{\sqrt{2}} (2z^{-1} - 0{\cdot}5)(z^{-1} - 0{\cdot}5)\, x + (1 - 0{\cdot}5z^{-1})\, y = z^{-2}(z^{-1} - 0{\cdot}5)\, \tfrac{7}{8}$$

and has the solution

$$x_0 = \frac{1}{2\sqrt{2}} , \quad y_0 = z^{-3} - 0{\cdot}5z^{-2}$$

satisfying $\partial y_0 < 4$.

Thus the optimal single-stage predictor is a stable a realization of (36)

$$F = \frac{1}{2} \frac{1}{2 - 0{\cdot}5z^{-1}}$$

and it yields

$$\|F_E\|^2_{\min} = \frac{8}{15} + \frac{7}{6} + \frac{7}{10} = 1\,.$$

Naturally, the prediction variance is greater than the filtering variance, see Example 1. □

Example 5. Given a noise-free stationary random sequence W by

$$F_W = \frac{1}{1 - 0{\cdot}5z^{-1}}\,,$$

solve problem (16) for all $\lambda = 1, 2, \ldots$.

In this particular case we have

$$F_S = F_L = F_R = F_W$$

and hence the equation

$$z^{-\lambda-1}(z^{-1} - 0{\cdot}5)\,x + (1 - 0{\cdot}5z^{-1})\,y = z^{-1}(z^{-1} - 0{\cdot}5)$$

is to be solved. We obtain

$$x_0 = \frac{1}{2^\lambda}$$

$$y_0 = (z^{-2} - 0{\cdot}5z^{-1})\left(1 + \frac{1}{2}z^{-1} + \ldots + \frac{1}{2^k}z^{-k} + \ldots + \frac{1}{2^{\lambda-1}}z^{-(\lambda-1)}\right)$$

and, therefore, the optimal λ-stage predictor is a stable realization of

$$F = \frac{1}{2^\lambda}\,.$$

The resulting prediction error variance can be written as

$$\|F_E\|^2_{\min} = \tfrac{4}{3}(1 - 0{\cdot}5^{2\lambda})\,.$$

It increases with λ but it remains bounded for large λ. □

THE SMOOTHING PROBLEM

The filtering and prediction problems are usually associated with real-time operations, in which estimates are required on the basis of observations or data available now. In a post mortem analysis, it is possible to wait for more observations to accumulate. In that case the estimate can be improved by smoothing. A sequence W smoothed for μ stages backward can be described as $z^{-\mu}W$ and, therefore, the smoothing problem (17) can again be solved analogously to the filtering problem (15).

Theorem 3. Problem (17) has a solution if and only if the linear Diophantine equation

$$z^{-m}\tilde{b}^*\tilde{u}x + ry = z^{-\mu-n}\tilde{a}\tilde{v}s \tag{38}$$

has a solution x_0, y_0 such that $\partial y_0 < \partial z^{-m}\tilde{b}^*\tilde{u}$ and

$$F = \frac{ax_0}{rb^*} \tag{39}$$

belongs to $\mathfrak{R}^+\{z^{-1}\}$.

The optimal μ-stage smoother $\mathscr{F}$ is given as a stable realization of (39) and it is essentially unique when minimally realized. Moreover,

$$\|F_E\|^2_{\min} = \left\langle\left(\frac{y_0}{b^*u}\right)\left(\frac{y_0}{b^*u}\right)\right\rangle + \langle\Phi_{WW} - \Phi_{WS}\Phi_{SS}^{-1}\Phi_{SW}\rangle\,. \tag{40}$$

Proof. The proof is again analogous to the proof of Theorem 1. We just write

$$E = z^{-\mu}W - FS$$

and hence (24) becomes

$$F_{E0} = \frac{z^{-\mu-n}\tilde{a}\tilde{v}s}{z^{-m}\tilde{b}^*\tilde{u}r} - \frac{b^*}{a}F\,.$$

Taking the decomposition

$$\frac{z^{-\mu-n}\tilde{a}\tilde{v}s}{z^{-m}\tilde{b}^*\tilde{u}r} = \frac{y}{z^{-m}\tilde{b}^*\tilde{u}} + \frac{x}{r}\,,$$

we obtain equation (38). The optimal smoother (39) and expression (40) follow by repeating the arguments. □

Example 6. Consider again

$$F_W = F_L = F_R = \sqrt{\frac{7}{8}}\frac{1}{1 - 0{\cdot}5z^{-1}}\,, \quad F_S = \frac{1}{\sqrt{2}}\frac{2 - 0{\cdot}5z^{-1}}{1 - 0{\cdot}5z^{-1}}$$

and solve problem (17) for $\mu = 1$, i.e. single-stage smoothing.

Equation (38) now reads

$$z^{-1}\frac{1}{\sqrt{2}}(2z^{-1} - 0{\cdot}5)(z^{-1} - 0{\cdot}5)\,x + (1 - 0{\cdot}5z^{-1})\,y = z^{-3}(z^{-1} - 0{\cdot}5)\tfrac{7}{8}$$

and has the solution

$$x_0 = \frac{1}{\sqrt{178}}(2 + 7z^{-1})$$

$$y_0 = \tfrac{1}{16}(z^{-2} - 0{\cdot}5z^{-1})$$

satisfying $\partial y_0 < 3$.

Thus the optimal single-stage smoother is a stable realization of

$$F = \frac{1}{8}\frac{2 + 7z^{-1}}{2 - 0{\cdot}5z^{-1}}$$

and it yields

$$\|F_E\|^2_{\min} = \frac{1}{480} + \frac{7}{6} - \frac{7}{10} = \frac{15}{32}\,.$$

Naturally, the smoothing variance is less than the filtering variance, see Example 1. □

Example 7. Given a noise-free stationary random sequence W by

$$F_W = \frac{1 - z^{-1}}{1 - 0{\cdot}5z^{-1}}\,,$$

solve problem (17) for all $\mu = 1, 2, \ldots$.

In this particular case we have

$$F_S = F_L = F_R = F_W$$

and hence the equation

$$z^{-2}(1 - z^{-1})(z^{-1} - 0{\cdot}5)\,x + (1 - 0{\cdot}5z^{-1})\,y =$$
$$= z^{-\mu-2}(z^{-1} - 0{\cdot}5)(z^{-1} - 1)(1 - z^{-1})$$

is to be solved. We obtain

$$x_0 = z^{-\mu}(z^{-1} - 1)\,, \quad y_0 = 0$$

and, therefore, the optimal μ-stage smoother is a stable realization of

$$F = z^{-\mu}\,.$$

The minimized variance of the smoothing error results

$$\|F_E\|^2_{\min} = 0$$

irrespective of μ.

CONCLUDING REMARKS

The algebraic approach, originally developed for the synthesis of discrete optimal control, has been applied to the problem of minimum variance filtering, prediction and smoothing. The synthesis procedure is very simple and, in contrast to traditional methods, reduces to solving a linear Diophantine equation in polynomials. Thus a complete algorithmization and machine processing is possible.

The reader's attention is drawn to a formal similarity between the deterministic problem of open-loop least squares control [18, 22] and the stochastic problem of minimum variance filtering. Even if conceptually very different, the two problems lead to solving essentially the same polynomial equation.

(Received November 13, 1975.)

REFERENCES

[1] Åström K. J.: Introduction to Stochastic Control Theory. Academic Press. New York 1970.
[2] Beneš J.: Statistická dynamika regulačních obvodů. SNTL, Praha 1961.
[3] Bode H. J., Shannon C. E.: A simplified derivation of linear least-squares smoothing and prediction theory. Proc. IRE 38 (1950), 417—425.
[4] Bryson A. E., Frazier M.: Smoothing for linear and nonlinear systems. TDR-63-119, Aeronautical Systems Division, Wright-Patterson AFB, Dayton, Oh. 1963.
[5] Bucy R. S., Joseph P. D.: Filtering for Stochastic Processes with Applications to Guidance. Interscience, New York 1958.
[6] Chang S. S. L.: Synthesis of Optimum Control Systems. McGraw-Hill, New York 1961.
[7] Doob J. L.: Stochastic Processes. Wiley, New York 1955.
[8] Fraser D. C.: A new technique for the optimal smoothing of data. D.Sc. Thesis, Instrumentation Lab., MIT, Cambridge, Mass. 1967.
[9] Гнеденко Б. В.: Курс теории вероятностей. Физматгиз, Москва 1961.
[10] Hardy G. H., Wright E. M.: An Introduction to the Theory of Numbers. Clarendon Press, Oxford 1960.
[11] Ho Y. C., Lee R. C. K.: A Bayesian approach to problems in stochastic estimation and control. IEEE Trans. Automatic Control AC-19 (1964), 4, 333—339.
[12] Jazwinski A. H.: Stochastic Processes and Filtering Theory. Academic Press, New York 1970.
[13] Kalman R. E.: A new approach to linear filtering and prediction problems. Trans. ASME, J. Basic Eng. 82D (1960), 35—45.
[14] Kalman R. E., Bucy R. S.: New results in linear filtering and prediction theory. Trans. ASME, J. Basic Eng. 83D (1961), 95—108.
[15] Kalman R. E.: On the general theory of control systems. Proc. 1st IFAC Congress, Moscow 1960.
[16] Kalman R. E., Falb P. L., Arbib M. A.: Topics in Mathematical System Theory. McGraw-Hill. New York 1969.
[17] Kučera V.: Algebraic theory of discrete optimal control for single-variable systems I — Preliminaries. Kybernetika 9 (1973), 2, 94—107.
[18] Kučera V.: Algebraic theory of discrete optimal control for single-variable systems II — Open-loop control. Kybernetika 9 (1973), 3, 206—221.
[19] Kučera V.: Algebraic theory of discrete optimal control for single-variable systems III — Closed-loop control. Kybernetika 9 (1973), 4, 291—312.
[20] Kučera V.: Closed-loop stability of discrete linear single-variable systems. Kybernetika 10 (1974), 2, 146—171.
[21] Kučera V.: Algebraic approach to discrete stochastic control. Kybernetika 11 (1975), 2, 114—147.
[22] Kučera V.: Algebraic approach to discrete linear control. IEE Trans. Automatic Control AC-20 (1975), 1, 116—120.
[23] Kučera V.: A contribution to matrix quadratic equation. IEEE Trans. Automatic Control AC-17 (1972), 3, 344—347.

[24] Kučera V.: The discrete Riccati equation of optimal control. Kybernetika 8 (1972), 5, 430—447.
[25] Kučera V.: On nonnegative definite solutions to matrix quadratic equations. Automatica 8 (1972), 4, 413—423.
[26] Kučera V.: A review of the matrix Riccati equation. Kybernetika 9 (1973), 1, 42—61.
[27] Kuroš A. G.: Kapitoly z obecné algebry. Academia, Praha 1968.
[28] Laning J. H., Battin R. H.: Random Processes in Automatic Control. McGraw-Hill, New York 1956.
[29] Levin B. R.: Teorie náhodných procesů a její aplikace v radiotechnice. SNTL, Praha 1965.
[30] Meditch J. S.: Stochastic Optimal Linear Estimation and Control. McGraw-Hill, New York 1969.
[31] Михелович Ш. Х.: Теоря чисел. Высшая школа, Москва 1967.
[32] Пугачев В. С.: Теория случайных функций и ее применение к задачам автоматического управления. Физматгиз, Москва 1962.
[33] Rauch H. E.: Solutions to the smoothing problem. IEEE Trans, Automatic Control AC-8 (1963), 371—372.
[34] Sage A. P.: Optimum Systems Control. Prentice-Hall, Englewood Cliffs, N. J. 1968.
[35] Strejc V. et al.: Syntéza regulačních obvodů s číslicovým počítačem. NČSAV, Praha 1965.
[36] Wiener N.: The Extrapolation, Interpolation and Smoothing of Stationary Time Series. Wiley, New York 1949.
[37] Youla D. C.: On the factorization of rational matrices. IRE Trans. Information Theory IT-7 (1961), 3, 172—189.
[38] Zariski O., Samuel P.: Commutative Algebra. Van Nostrand, Princeton, N. J. 1958.

Ing. Vladimír Kučera, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 180 76 Praha 8. Czechoslovakia.