INT. J. CONTROL, 1976, VOL. 24, NO. 6, 821–842

Some reduced-order non-Riccati equations for linear least-squares estimation : the stationary, single-output case†

ANDERS LINDQUIST‡

The problem of determining the Kalman–Bucy filter for an n-dimensional single-output model is the topic of this paper. Both the discrete-time case and continuous-time case are considered. The model processes are assumed to be stationary. It is shown that, under certain regularity conditions, only n first-order difference or differential equations are required for determining the error covariance function, and hence also the filter gain, rather than $\frac{1}{2}n(n+1)$ equations as with the Riccati approach or $2n$ as in the previous non-Riccati algorithm. This reduction is achieved by constructing a system of simple integrals for the $2n$ non-Riccati equations. The reduced-order algorithms have non-trivial steady-state versions, which are equivalent to the algebraic equations obtained by spectral factorization. The stationary and single-output assumptions are for convenience. In fact, the basic method works also in a more general setting.

1. Introduction

This paper is concerned with recursive linear least-squares filtering of lumped, stationary stochastic processes in both discrete and continuous time. Such problems have traditionally been approached by means of Kalman–Bucy filtering techniques, which require the solution of an $n \times n$ matrix Riccati equation, n being the order of the model. Due to symmetry this amounts to solving $\frac{1}{2}n(n+1)$ first-order difference or differential equations to determine the $n \times m$ matrix gain function, where m is the number of outputs. Recently, however, a new type of algorithm has been developed which, whenever $m \ll n$, requires much fewer dynamic relations. In fact, the discrete-time version (Lindquist 1974 a) contains $2mn + \frac{1}{2}m(m+1)$ first-order difference equations, whereas only $2mn$ first-order differential equations are required in continuous time (Kailath 1973). The continuous time version could be regarded as an extension of the Chandrasekhar-type results of Casti *et al.* (1972) and Casti and Tse (1972). A similar discrete-time algorithm for ARMA models, containing $2(n+1)m^2$ equations, is due to Rissanen (1973). The reader is referred to Lindquist (1975 a) for a discussion of the relation between these results.

It is possible to reduce the order of these equations even further to obtain an algorithm with only mn first-order equations, both in continuous and discrete time. This reduction is the topic of the present paper. However, to make our presentation more transparent, we shall only consider the single-output case ($m = 1$) here. The generalization to $m > 1$ is decidedly non-trivial, and we shall present it elsewhere together with modifications for non-stationary models.

Received 24 February 1975 ; revised 10 November 1975.

† This work was supported by the National Science Foundation under grant MPS75–07028.

‡ Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506.

The reduction of the non-Riccati algorithm will be performed in several steps. First, we shall develop a parameter-free version of the algorithm with the property that the dynamic equations do not depend on the system parameters other than through the initial conditions. This is precisely the property of Rissanen's algorithm. In fact, the discrete-time version of the parameter-free algorithm is essentially identical to Rissanen's scalar-output equations. (In the multi-output case things are not so simple, in that Rissanen's algorithm contains more equations.) Secondly, we shall derive a system of time-independent *integrals* for the dynamic equations. In continuous time this provides us with n static relations between the $2n$ differential equations ; in discrete time $n+1$ static relations between $2n+1$ difference equations. Thirdly, after applying a suitable transformation, we shall use the integrals to reduce the number of dynamic relations, leaving us with only n. This amounts to inverting a Hurwitz matrix in continuous time and a sum of a Toeplitz and a Hankel matrix in discrete time. The discrete-time equations can be reduced even further if either the measurement noise is zero or the system matrix is singular. (This phenomenon is also observed in the theory of constant direction due to Bucy *et al.* 1970.) Finally, we demonstrate that the reduced system of n dynamic equations is sufficient for determining not only the filter gain but the complete Riccati solution.

It is well established (Willis and Brockett 1965, Rugh and Murphy 1969, Buelens and Hellinckx 1974) that the algebraic Riccati equation can be reduced to the n algebraic equations of the spectral factorization. What we have done is to construct the ' transients ' of these equations. Indeed, the n algebraic equations are the steady-state version of our reduced-order system. In this context we may note that the original non-Riccati algorithm (Kailath 1973, Lindquist 1974 a) has no non-trivial steady-state solution.

So far we have focused our attention on the number of first-order dynamic equations contained in each algorithm. This question has a theoretical interest but has little practical relevance. Indeed, in any computational procedure the number of arithmetic operations is more interesting. In this context we must recognize the fact that the reduction presented in this paper is bought at the expense of greater algebraic complexity. Therefore, the original non-Riccati algorithm may well be preferred from a computational point of view. If so, the integrals could nevertheless be used to check the solutions ' on-line ', and re-initiate the algorithm in the case of numerical divergence.

The paper consists of two parts. In § 2 the continuous-time equations are developed, whereas § 3 is devoted to the discrete-time case.

2. The continuous-time case

2.1. *Problem formulation*

Consider the usual continuous-time, lumped, stochastic single-output model

$$\dot{x} = Ax + Bv \; ; \quad x(0) = x_0 \tag{2.1}$$

$$y = c'x + w \tag{2.2}$$

where x is the n-dimensional state process and y the scalar observation process. Prime (′) denotes transpose. The stochastics of the model is provided by the stochastic vector x_0 and the white noise processes v and w ($v(t) \in R^p$, $w(t) \in R$), all of which are mutually uncorrelated, have zero mean and the following second-order properties :

$$E\{x_0 x_0'\} = P_0 \tag{2.3}$$

$$E\{v(s)v(t)'\} = I\delta(s-t) \tag{2.4}$$

$$E\{w(s)w(t)\} = \delta(s-t) \tag{2.5}$$

where δ is the Dirac (generalized) function. The $n \times n$ system matrix A, the $n \times p$ matrix B and the n vector c are constant. Moreover, we assume that (x, y) is (wide sense) stationary. That is, A is a stability matrix, i.e. all the zeros of the characteristic polynomial

$$\det(sI - A) = z^n + a_1 z^{n-1} + a_2 z^{n-2} + \ldots + a_n \tag{2.6}$$

have negative real parts, and the Liapunov equation

$$AP_0 + P_0 A' + BB' = 0 \tag{2.7}$$

holds.

It is well known that the linear least-squares estimate $\hat{x}(t)$ of $x(t)$ given the data $\{y(s)\,;\; 0 \leqslant s \leqslant t\}$ is generated by the Kalman–Bucy filter

$$\left.\begin{aligned} \frac{d\hat{x}}{dt} &= A\hat{x}(t) + k(t)[y(t) - c'\hat{x}(t)] \\ \hat{x}(0) &= 0 \end{aligned}\right\} \tag{2.8}$$

whereby the problem is reduced to determining the n-dimensional gain vector function k. This is usually done by solving the matrix Riccati equation

$$\left.\begin{aligned} \dot{P} &= AP + PA' - Pcc'P + BB' \\ P(0) &= P_0 \end{aligned}\right\} \tag{2.9}$$

for the $n \times n$ error covariance matrix P, in terms of which the gain function

$$k(t) = P(t)c \tag{2.10}$$

is formed. Since P is symmetric, $\frac{1}{2}n(n+1)$ coupled scalar equations need to be solved.

Recently another procedure to determine the gain k has been proposed which only requires $2n$ scalar first-order equations (Kailath 1973, Casti 1974, Lindquist 1974 b). This non-Riccati algorithm consists of the two n-vector equations

$$\dot{k} = -(c'q^*)q^*\,; \quad k(0) = P_0 c \tag{2.11}$$

$$\dot{q}^* = (A - kc')q^*\,; \quad q^*(0) = P_0 c \tag{2.12}$$

In this paper we shall demonstrate that these equations can be reduced even further, leaving us with only n first-order differential equations. In our analysis, the characteristic polynomial

$$Q(t, z) :\; = \det(zI - A + k(t)c') = z^n + q_1(t)z^{n-1} + \ldots + q_n(t) \tag{2.13}$$

of the *feedback matrix*

$$A - k(t)c' \tag{2.14}$$

will play an important role. Note that (2.14) is the coefficient matrix in both the filter eqn. (2.8) and in (2.12).

In the sequel we shall assume that (A, c) is *observable*, i.e. that the $n \times n$ matrix

$$M = \begin{bmatrix} c'A^{n-1} \\ c'A^{n-2} \\ \vdots \\ c' \end{bmatrix} \tag{2.15}$$

is non-singular. If, in addition, (A, B) is *controllable*, i.e.

$$N = (B, AB, A^2B, \ldots, A^{n-1}B) \tag{2.16}$$

has full rank, (A, B, c) is *a minimal realization* (see, e.g. Brockett 1970). Later, in § 2.4, we shall need to impose this rather natural condition.

2.2. *A parameter-free version of the non-Riccati algorithm*

For the moment, let A be the companion matrix

$$\Gamma(a) = \begin{bmatrix} -a_1 & 1 & 0 & \ldots & 0 \\ -a_2 & 0 & 1 & \ldots & 0 \\ \hdashline -a_{n-1} & 0 & 0 & \ldots & 1 \\ -a_n & 0 & 0 & \ldots & 0 \end{bmatrix} \tag{2.17}$$

where a is the vector $(a_1, a_2, \ldots, a_n)'$ of coefficients of (2.6), and let c be the unit vector

$$h = (1, 0, 0, \ldots, 0)' \tag{2.18}$$

This is no restriction. Indeed, if (A, c) is an arbitrary observable pair, there is a non-singular matrix

$$T = (A^{n-1}M^{-1}h, A^{n-2}M^{-1}h, \ldots, M^{-1}h) \tag{2.19}$$

such that

$$\Gamma(a) = T^{-1}AT \tag{2.20}$$

$$h' = c'T \tag{2.21}$$

and therefore all relations in § 2.1 remain valid with (A, c) exchanged for $(\Gamma(a), h)$, if only at the same time P_0 and P are transformed as

$$P \mapsto T^{-1}P(T')^{-1} \tag{2.22}$$

and B together with all vectors as

$$B \mapsto T^{-1}B \tag{2.23}$$

First observing that

$$\Gamma(z) = J - zh' \tag{2.24}$$

where J is the shift matrix

$$J = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \hline 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} \tag{2.25}$$

we see that the feedback matrix (2.14) can be written

$$J - ah' - kh' = \Gamma(k+a) \tag{2.26}$$

Hence we have the well-known relation

$$k = q - a \tag{2.27}$$

where the components $q_1, q_2, \dots, q_n$ of the vector q are defined by (2.13). Then we can rewrite (2.11) and (2.12) as

$$\dot{q} = -q_1{}^* q^* \; ; \quad q(0) = P_0 h + a \tag{2.28}$$

$$\dot{q}^* = Jq^* - q_1{}^* q \; ; \quad q^*(0) = P_0 h \tag{2.29}$$

To see this, first observe that $\Gamma(q)$ is the coefficient matrix of (2.12). Then use (2.24) to obtain (2.29). To derive (2.28) differentiate (2.27). In the following lemma we shall collect these observations, reformulated in the general setting where (A, c) is not necessarily of the form (2.17) and (2.18).

Lemma 2.1

Let (A, c) be observable. Then the gain vector function (2.10) is

$$k(t) = T[q(t) - a] \tag{2.30}$$

where T is the transformation (2.19) and $q_1(t), q_2(t), \dots, q_n(t)$ are the coefficients of the characteristic polynomial (2.13). The vector function q satisfies the system of differential equations

$$\dot{q} = -q_1{}^* q^* \; ; \quad q(0) = T^{-1} P_0 c + a \tag{2.31}$$

$$\dot{q}^* = Jq^* - q_1{}^* q \; ; \quad q^*(0) = T^{-1} P_0 c \tag{2.32}$$

Hence we have obtained a parameter-free version of (2.11) and (2.12) which exhibits its dependence on the systems parameters only in the initial conditions, the differential equations themselves being ‘ universal ’. Each system, characterized by (A, B, c) provides a set of initial conditions for these differential equations. As we shall see in the next section, there are n *simple* functions $\phi_1, \phi_2, \dots, \phi_n$ such that $\phi_i(q(t), q^*(t)) \equiv \text{constant}$ for $i = 1, 2, \dots, n$. In the classical theory of differential equations (e.g. Moulton 1930) such relations are called *integrals*.

2.3. *Integrals of the non-Riccati system*

To facilitate the formulation of the integrals, define the auxiliary polynomial

$$Q^*(t, z) = q_1{}^*(t) z^{n-1} + q_2{}^*(t) z^{n-2} + \dots + q_n{}^*(t) \tag{2.33}$$

Equations (2.31) and (2.32) can be reformulated in terms of the polynomials Q and Q^*. In fact, it is easily seen that

$$\dot{Q}(t, z) = -q_1{}^*Q^*(t, z) \tag{2.34}$$

$$\dot{Q}^*(t, z) = zQ^*(t, z) - q_1{}^*Q(t, z) \tag{2.35}$$

where the dot denotes differentiation with respect to t. We may note that these differential equations are identical to certain continuous analogues (Krein 1955) of Szego's polynomials orthogonal on the unit circle. However, the initial conditions are quite different in that $Q(0, z)$ and $Q^*(0, z)$ are polynomials. We have previously (Lindquist 1975 a) used Krein's equations to derive the non-Riccati algorithm (2.11) and (2.12).

Lemma 2.2

The polynomials (2.13) and (2.33) satisfy the following equation for all t

$$Q(t, z)Q(t, -z) - Q^*(t, z)Q^*(t, -z) = D(z^2) \tag{2.36}$$

where $D(z) = \sum_{i=0}^{n} d_i z^{n-i}$ is a constant polynomial.

The proof of this lemma is immediate; just differentiate the left member of (2.36) and use (2.34) and (2.35) to see that all terms cancel. Clearly (2.36) must be a polynomial in z^2. To determine D we may just take $t=0$ in (3.36) and insert the initial conditions. However, the formula thus received will contain P_0, a quantity which we shall later eliminate from the algorithm. Therefore, in § 2.5, we shall derive an alternative expression for D which depends explicitly on (A, B, c) only.

Identifying coefficients of z in (2.36), we obtain n integrals for the system (2.26) and (2.27). We shall use these to reduce the order of the algorithm. Taking $t=\infty$ in (2.36), we obtain the spectral factorization formula

$$Q(\infty, z)Q(\infty, -z) = D(z^2) \tag{2.37}$$

for the steady-state filter (Willis and Brockett 1965, Brockett 1970). In fact, $Q^*(t, z) \to 0$ as $t \to \infty$. (See the end of § 2.6.)

2.4. *Solution of the integrals*

Let

$$U(t, z) = z^n + u_1(t)z^{n-1} + \ldots + u_n(t) \tag{2.38}$$

$$V(t, z) = z^n + v_1(t)z^{n-1} + \ldots + v_n(t) \tag{2.39}$$

be the monic polynomial functions defined by

$$U(t, z) = Q(t, z) - Q^*(t, z) \tag{2.40}$$

$$V(t, z) = Q(t, z) + Q^*(t, z) \tag{2.41}$$

and let $u = (u_1, u_2, \ldots, u_n)'$ and $v = (v_1, v_2, \ldots, v_n)'$.

Lemma 2.3

Let (A, B, c) be a minimal realization. Then $U(t, z)$ and $V(t, z)$ are stability polynomials for all $t \geq 0$.

Proof

Since (A, c) is observable, it is no restriction to assume that $A = \Gamma(a)$ and $c = h$. It can be shown (Kailath 1973, Casti 1974, Lindquist 1974 b) that

$$\dot{P} = -q^* q^{*\prime} \tag{2.42}$$

and therefore the Riccati eqn. (2.9) can be written

$$-q^* q^{*\prime} = \Gamma(a)P + P\Gamma(a)' - kk' + BB' \tag{2.43}$$

Since

$$\Gamma(a) = \Gamma(u) + (u-a)h' \tag{2.44}$$

and $q^* = k - (u-a)$, we may reformulate (2.43) to obtain

$$\Gamma(u)P + P\Gamma(u)' + (u-a)(u-a)' + BB' = 0 \tag{2.45}$$

Hence for each fixed y and $u = u(t)$, $P(t)$ satisfies the Liapunov eqn. (2.44). We shall now proceed much along the same lines as in Brockett (1970, p. 148) to show that for each such u, which is kept *constant*,

$$\dot{x} = \Gamma(u)'x$$

is asymptotically stable. To this end, use (2.45) to see that

$$-\frac{d}{dt}[x(t)'Px(t)] = [(u-a)'x(t)]^2 + x(t)'BB'x(t) \geqslant 0 \tag{2.46}$$

where P is the *constant* non-negative definite solution of (2.45). Now (2.41) cannot be identically zero on any interval $(i, i+1)$ unless $x(i) = 0$, for if this were so, in view of (2.44) we would have

$$\dot{x} = Ax\ ; \quad B'x = 0$$

on $(i, i+1)$, which contradicts the assumption of controllability. Therefore,

$$x(i)'Px(i) - x(i+1)'Px(i+1) \geqslant \epsilon \|x(i)\|^2$$

for some $\epsilon > 0$, which does not depend on i. This gives us

$$\epsilon \sum_{i=0}^{\infty} \|x(i)\|^2 \leqslant x(0)'Px(0) \tag{2.47}$$

and hence $x(t) \to 0$ as $t \to \infty$. Moreover, from (2.46), we see that the function $x(t)'Px(t)$ is non-increasing, and from (2.47) that P is positive definite. Hence $\|x(t)\| \leqslant \lambda \|x(0)\|$ for some λ. This establishes the required asymptotic stability. Consequently, all eigenvalues of $\Gamma(u(t))$ have negative real parts, i.e. $U(t, z)$ is a stability polynomial. The stability of $V(t, z)$ is proved in the same way. ■

Remark 2.4

For $t = 0$ the Liapunov eqn. (2.45) is the same as (2.7); for $t = \infty$ it becomes the algebraic Riccati equation. Clearly $U(\infty, z) = V(\infty, z) = Q(\infty, z)$ is also a stability polynomial, for the proof is valid for $t = \infty$, too. ■

Now reformulate (2.36) in terms of U and V to obtain

$$U(t, z)V(t, -z) + V(t, z)U(t, -z) = 2D(z^2) \tag{2.48}$$

Then, identifying coefficients of like power in z, we have

$$\sum_{j=1}^{n} (-1)^j u_{2i-j} v_j + u_{2i} = d_i \; ; \quad i = 1, 2, \ldots, n \tag{2.49}$$

or in matrix form

$$H(u)v = d - H(0)u \tag{2.50}$$

where $H(u)$ is the *Hurwitz matrix*

$$H(u) = \begin{bmatrix} -u_1 & 1 & 0 & 0 & \ldots & 0 \\ -u_3 & u_2 & -u_1 & 1 & \ldots & 0 \\ -u_5 & u_4 & -u_3 & u_2 & \ldots & 0 \\ \hdotsfor{6} \\ 0 & 0 & 0 & 0 & \ldots & (-1)^n u_n \end{bmatrix} \tag{2.51}$$

of the polynomial $U(t, -z)$. Since $U(t, z)$ is stable (Lemma 2.3), $U(t, -z)$ has no opposite roots. Therefore, Orlando's formula (Gantmacher 1959) implies that $H(u(t))$ is non-singular. Consequently, we can solve (2.50) for v:

$$v(t) = H(u(t))^{-1}[d - H(0)u(t)] \tag{2.52}$$

In § 2.6 we shall use this relation to reduce the order of the non-Riccati algorithm. First, however, we shall investigate similar redundancies in the matrix Riccati equation.

A fast procedure for solving (2.48) may be designed by means of the Euclidean algorithm (see § 4).

2.5. *Formulae for the error covariance*

The (symmetric) matrix Riccati eqn. (2.9) and the non-Riccati eqns. (2.31) and (2.32) together constitute a system of $\frac{1}{2}n(n+1)+2n$ first-order differential equations. We have already found n integrals for this system (Lemma 2.2). The following lemma provides us with another $\frac{1}{2}n(n+1)$.

Lemma 2.5

Let (A, c) be in companion form $(\Gamma(a), h)$. Let $q_0 \equiv 1$, $q_0^* \equiv 0$, $a_0 = 1$, and let P_{ij} and B_{ij} be zero whenever some index is 0 or $n+1$. Then the solutions of (2.9), (2.31) and (2.32) satisfy

$$P_{i+1,\, j}(t) + P_{i,\, j+1}(t) = q_i(t)q_j(t) - q_i^*(t)q_j^*(t) - a_i a_j - (BB')_{ij} \tag{2.53}$$

for all $t \geqslant 0$ and all $i, j = 0, 1, 2, \ldots, n$.

Proof

In view of (2.10) and (2.27)

$$Ph = q - a \tag{2.54}$$

which gives us (2.53) for the case that either i or j is zero. To see that (2.53) is true for $i, j = 1, 2, \ldots, n$, we note that we can write the Riccati equation

$$JP + PJ' = qq' - q^*q^{*\prime} - BB' - aa' \tag{2.55}$$

which is the same as (2.53). To see this, just insert (2.24), (2.10) and (2.54) into (2.43). ■

We shall solve the integrals (2.53) for P. For later reference we shall express the solutions in terms of u and v.

Lemma 2.6

Given all the assumptions of Lemma 2.5, the solution P of the Riccati eqn. (2.9) is given by

$$P(t)=\Pi(u(t), v(t), B) \tag{2.56}$$

where

$$\Pi_{ij}(u, v, B)=\sum_{k=j}^{n}(-1)^{k-j}[\tfrac{1}{2}u_{i+j-1-k}v_k+\tfrac{1}{2}v_{i+j-1-k}u_k - a_{i+j-1-k}a_k-(BB')_{i+j-1-k\ \ k}] \tag{2.57}$$

in which we have taken $u_0=v_0=1$ and $v_i=u_i=0$ for $i<0$ and $i>n$. Moreover, the coefficient vector $d=(d_1, d_2, \ldots, d_n)'$ of the polynomial D, defined in Lemma 2.2, is $\delta(B)$, where

$$\delta_i(B)=\sum_{j=0}^{n}(-1)^j[a_{2i-j}a_j+(BB')_{2i-j,\ j}] \tag{2.58}$$

Proof

Relations (2.56) and (2.57) follow immediately from (2.53) upon exchanging $q_iq_j-q_i{}^*q_j{}^*$ for $\frac{1}{2}u_iv_j+\frac{1}{2}v_iu_j$ and performing the appropriate summation. It is easy to see that (2.53) holds for *all* (i, j) provided that we define P_{ij} and B_{ij} to be zero whenever an index is less than one or greater than n. Hence P_{ij} equals the right members of (2.57) for all (i, j). In particular, by noting that $P_{2i+1,\,0}\equiv 0$ for $i=1, 2, \ldots, n$, we obtain an alternative derivation of (2.49) and the expression (2.58) for d. ■

2.6. *The reduced-order non-Riccati system*

We shall now apply the linear transformation

$$u=q-q^* \tag{2.59}$$

$$v=q+q^* \tag{2.60}$$

defined by (2.40) and (2.41), to the system (2.31) and (2.32).

Lemma 2.7

The vector functions u and v satisfy the system of first-order differential equations

$$\dot{u}=\tfrac{1}{2}\Gamma(u)(u-v)\ ; \quad u(0)=a \tag{2.61}$$

$$\dot{v}=\tfrac{1}{2}\Gamma(v)(v-u)\ ; \quad v(0)=2T^{-1}P_0c+a \tag{2.62}$$

where Γ is defined by (2.24).

Proof

Substracting (2.32) from (2.31), we obtain

$$\dot{u}=-\Gamma(u)q^*\ ; \quad u(0)=a \tag{2.63}$$

Then insert $q^* = -\frac{1}{2}(u-v)$ to obtain (2.61). Equation (2.62) is derived in the same way. ■

The function v can be eliminated from (2.61) by means of (2.52). This will leave us with n first-order differential equations to determine k *and* the error-covariance P. We collect this result in the following theorem.

Theorem 2.8

Let (A, B, c) be a minimal realization. Then the gain function (2.10) is given by

$$k(t) = \tfrac{1}{2}T[u(t) + H(u(t))^{-1}(d - H(0)u(t)) - 2a] \tag{2.64}$$

where u is the unique solution of the system

$$\dot{u} = \tfrac{1}{2}\Gamma(u)[u - H(u)^{-1}(d - H(0)u)] \tag{2.65 a}$$

$$u(0) = a \tag{2.65 b}$$

of n first-order differential equations, T is the constant matrix (2.19), and $d = \delta(T^{-1}B)$, δ being defined by (2.58). The matrix functions Γ are H are given by (2.24) and (2.51) respectively. Moreover, the solution P of the Riccati eqn. (2.9) is

$$P(t) = T\Pi(u(t), H(u(t))^{-1}(d - H(0)u(t)), T^{-1}B)T' \tag{2.66}$$

where Π is the matrix function (2.57). Everything above remains true if the initial condition (2.55 *b*) is exchanged for

$$u(0) = 2T^{-1}P_0c + a \tag{2.67}$$

Proof

Insert (2.52) into

$$k = \tfrac{1}{2}T(u + v - 2a) \tag{2.68}$$

[which is a consequence of (2.30) and (2.59) and (2.60)], (2.61) and (2.57) (and perform transformations (2.22) and (2.23) where necessary) to obtain (2.64), (2.65) and (2.66) respectively. For each fixed $t \geqslant 0$, the inverse of $H(u)$ exists for $u = u(t)$ and, by continuity, in some neighbourhood thereof (Lemma 2.3). Therefore the right member of (2.65 *a*) is differentiable and hence locally Lipschitz. Consequently, (2.65) has a unique solution. Since (2.48) is symmetric with respect to u and v, we have

$$u = H(v)^{-1}[d - H(0)v] \tag{2.69}$$

formed in analogy with (2.52). From (2.62) and (2.69) it follows that v, too, satisfies the differential eqn. (2.65 *a*) ; however, with the initial condition (2.67). Since (2.57) and (2.68) are symmetric in u and v, we may exchange u for v in (2.64) and (2.66). ■

The initial condition (2.55 *a*) is particularly simple ; we do not need P_0 as with eqns. (2.11) and (2.12). This should be an advantage, since usually (A, B, c) is given, whereas P_0 has to be solved from the Liapunov eqn. (2.7). Since d is also expressed in terms of (A, B, c), we have completely removed P_0 from the algorithm. The initial condition (2.67) is supplied merely for completeness.

Note that like the Riccati equation, but unlike the original non-Riccati algorithm (2.11) and (2.12), the reduced-order system (2.65) has a steady-state version obtained by putting $\dot{u}=0$, namely

$$u=H^{-1}(u)[d-H(0)u] \tag{2.70}$$

In fact, (2.42) being non-positive ensures the convergence of $\dot{P}(t)$ to zero as $t\to\infty$.† Hence, in view of (2.42), $\dot{q}^*(t)$ will also tend to zero, as will $\dot{q}(t)$. Consequently, $u(t)$ will tend to $q(\infty)$ and the right member of (2.65 *a*) to a constant, which must be zero. Since, $u_n(\infty)=U(\infty, 0)\neq 0$ (Remark 2.4), $\Gamma(u(\infty))$ is non-singular. Therefore, $u(\infty)$ satisfies (2.70). The steady-state gain $k(\infty)$ and the solution $P(\infty)$ of the algebraic Riccati equation are obtained by inserting $u(\infty)$ into (2.64) and (2.66). Finally, we may note that the steady-state eqn. (2.70) is equivalent to the spectral factorization formula (2.37).

3. The discrete-time case

3.1. *Problem formulation*

In the discrete-time setting the n-dimensional system process x and the scalar observation process y are generated by

$$x(t+1)=Ax(t)+Bv(t)\ ; \quad x(0)=x_0 \tag{3.1}$$

$$y(t)=c'x(t)+w(t) \tag{3.2}$$

($t=0, 1, 2, \ldots$), where x_0, v and w have zero mean, are mutually uncorrelated, and have the following second-order statistics :

$$E\{x_0\,x_0'\}=P_0 \tag{3.3}$$

$$E\{v(s)v(t)'\}=I\delta_{st} \tag{3.4}$$

$$E\{w(s)w(t)\}=\alpha\delta_{st} \tag{3.5}$$

Here δ_{st} is the Kronecker symbol (taking the value one if $s=t$ and zero otherwise) and α is a non-negative constant. The system parameters (A, B, c) are constant and all matrices have the same dimensions as in § 2.

Again we assume that (x, y) is (wide sense) stationary, i.e. the characteristic polynomial (2.6) has all its zeros inside the unit circle, and P_0 satisfies the discrete Liapunov equation

$$P_0=AP_0A'+BB' \tag{3.6}$$

Let $T(t)$ be the $t\times t$ Toeplitz matrix

$$T_{ij}(t)=c'A^{|i-j|}P_0c+\alpha\delta_{i-j} \tag{3.7}$$

To ensure that the random sequence y has full rank even if $\alpha=0$, we assume that

$$\det T(t)>0\ ; \quad t=1, 2, 3, \ldots \tag{3.8}$$

Condition (3.8) is always fulfilled when $\alpha>0$.

† In fact, since $P(t)$ is monotone non-increasing and bounded from below by zero, it tends to a limit as $t\to\infty$. Then $\dot{P}(t)$ tends to a limit also, which must be zero.

The linear least-squares estimate $\hat{x}(t)$ of $x(t)$ given the data $\{y(0), y(1), \ldots, y(t-1)\}$ is generated by the Kalman filter

$$\left.\begin{aligned} \hat{x}(t+1) &= A\hat{x}(t) + \frac{1}{q_0(t)} k(t)[y(t) - c'\hat{x}(t)] \\ \hat{x}(0) &= 0 \end{aligned}\right\} \tag{3.9}$$

Therefore it remains to determine the n-dimensional gain vector sequence k and the scalar sequence q_0. This is usually done by solving the matrix Riccati equation

$$\left.\begin{aligned} P(t+1) &= A[P(t) - P(t)c(c'P(t)c + \alpha)^{-1}c'P(t)]A' + BB' \\ P(0) &= P_0 \end{aligned}\right\} \tag{3.10}$$

for the $n \times n$ matrix sequence P, in terms of which we have

$$k(t) = AP(t)c \tag{3.11}$$

$$q_0(t) = c'P(t)c + \alpha \tag{3.12}$$

[The matrix $P(t)$ is the covariance of the estimation error $x(t) - \hat{x}(t)$, and $q_0(t)$ is the variance of the innovation process $y(t) - c'\hat{x}(t)$. Condition (3.8) ensures that $q_0(t) > 0$ for all t. (See below).]

The discrete-time analogue of the non-Riccati algorithm (2.11) and (2.12) can be written (Lindquist 1974 *a*)

$$k(t+1) = k(t) - \gamma_t Aq^*(t)\ ; \quad k(0) = AP_0c \tag{3.13}$$

$$q^*(t+1) = Aq^*(t) - \gamma_t k(t)\ ; \quad q^*(0) = AP_0c \tag{3.14}$$

$$q_0(t+1) = (1 - \gamma_t^2)q_0(t)\ ; \quad q_0(0) = c'P_0c + \alpha \tag{3.15}$$

where

$$\gamma_t = \frac{1}{q_0(t)} c'q^*(t) \tag{3.16}$$

(also, see Kailath *et al.* (1973) which contains certain extensions of the results of Lindquist (1974 *a*)). It can be shown that (3.8) is equivalent to the condition

$$|\gamma_t| < 1, \quad t = 0, 1, 2, \ldots \tag{3.17}$$

In fact, given the connections with Szego's polynomials orthogonal on the unit circle explained in Lindquist (1974 a, 1975 a), this equivalence is an immediate consequence of Theorem 8.1 in Geronimus (1961). Therefore, since $q_0(0) = \det T(1) > 0$, $q_0(t) > 0$ for all t.

In § 3.3 we shall demonstrate that the system (3.13)–(3.15) of $2n+1$ first-order difference equations has $n+1$ easily solvable 'integrals', which can be used to reduce the order of the system to n. To this end, we shall first derive a parameter-free version of the non-Riccati algorithm (3.13)–(3.15).

3.2. *A parameter-free version of the non-Riccati algorithm*

The discrete-time version of (2.31) and (2.32) consists of the $2n+1$ recursions

$$q_0(t+1)=q_0(t)-\gamma_t q_1^*(t)\ ; \quad q_0(0)=c'P_0c+\alpha \tag{3.18}$$

$$q(t+1)=q(t)-\gamma_t Jq^*(t)\ ; \quad q(0)=T^{-1}AP_0c+q_0(0)a \tag{3.19}$$

$$q^*(t+1)=Jq^*(t)-\gamma_t q(t)\ ; \quad q^*(0)=T^{-1}AP_0c \tag{3.20}$$

where $q(t)$ and $q^*(t)$ are n-dimensional vectors, and

$$\gamma_t=q_1^*(t)/q_0(t) \tag{3.21}$$

Lemma 3.1

Let (A, c) be observable. Then the gain parameters q_0 and k in the Kalman filter (3.9) are given by (3.18) and

$$k(t)=T[q(t)-q_0(t)a] \tag{3.22}$$

respectively, where T is defined by (2.19) and q by (3.19). The last component of q is constant, i.e.

$$q_n=\alpha a_n \tag{3.23}$$

Proof

Assume for the moment that $A=\Gamma(a)$ and $c=h$ as defined in § 2.2. Then (3.16) becomes (3.21), and consequently, (3.18) is the same as (3.15). Note that $T=I$, and let q be defined by (3.22). Then (3.13) yields

$$q(t+1)-q_0(t+1)a=q(t)-q_0(t)a-\gamma_t Jq^*(t)+\gamma_t q_1^*(t)a$$

which, in view of (3.18), is the same as the recursion (3.19). Inserting (3.22) into (3.14), we have

$$q^*(t+1)=Jq^*(t)-q_1^*(t)a-\gamma_t q(t)+\gamma_t q_0(t)a$$

which by (3.21), gives us recursion (3.20). Next, let (A, c) be an arbitrary observable pair. Then transformations (2.20)–(2.23) must be performed. This leaves all recursions intact, but changes the initial conditions to those exhibited in (3.18)–(3.20). Also (3.22) is obtained. It is clear from (3.19) that q_n is constant. To determine this constant, observe that the initial condition of (3.19) can be written

$$\Gamma(a)T^{-1}P_0c+q_0(0)a$$

the nth component of which is αa_n. ∎

Remark 3.2

If $(A, c)=(\Gamma(a), h)$, the filter recursion (3.9) may be written

$$\hat{x}(t+1)=J\hat{x}(t)+\frac{1}{q_0(t)}\,q(t)[y(t)-\hat{y}(t)]-ay(t)$$

where $\hat{y}(t)=h'\hat{x}(t)$, and therefore the standard procedure yields the following recursion in y:

$$\hat{y}(t+1)-\sum_{i=0}^{n-1}\eta_{i+1}(t+1)[y(t-i)-\hat{y}(t-i)]=-\sum_{i=0}^{n-1}a_{i+1}y(t-i) \tag{3.24}$$

where

$$\eta_i(t+1)=\frac{1}{q_0(t-i)}q_i(t-i) \tag{3.25}$$

The predictor (3.24) has been studied by Rissanen (1973) who derived a fast algorithm for the coefficients $\eta_i(t)$. In fact, by introducing the new vector sequences

$$z_{ij}=q_{j-i}(i-j)$$
$$z_{ij}^*=q_{j-i}^*(i)$$

the parameter-free version (3.18)–(3.20) of the non-Riccati algorithm (3.13)–(3.15) is seen to be essentially identical to Rissanen's algorithm. This should clarify the connection between the results of Rissanen (1973) and Lindquist (1974 *a*) as far as single-output processes are concerned. In the multi-output case ($m>1$), things are more complicated in that the numbers of equations in the two algorithms do not match. (Also see Rissanen 1975.) ∎

3.3. *Integrals for the non-Riccati system*

Let Q_t and Q_t^*, $t=0, 1, 2, \ldots$, be the polynomials

$$Q_t(z)=q_0(t)z^n+q_1(t)z^{n-1}+\ldots+q_n(t) \tag{3.26}$$

$$Q_t^*(z)=q_1^*(t)z^{n-1}+q_2^*(t)z^{n-2}+\ldots+q_n^*(t) \tag{3.27}$$

Then eqns. (3.18)–(3.20) can be written in the following form:

$$Q_{t+1}(z)=Q_t(z)-\gamma_t zQ_t^*(z) \tag{3.28}$$

$$Q_{t+1}^*(z)=zQ_t^*(z)-\gamma_t Q_t(z) \tag{3.29}$$

These recursions are identical to those found in the theory of polynomials orthogonal on the unit circle (Geronimus 1961, Akiezer 1965). However, whereas the orthogonal polynomials have initial values 1 and increase in degree with t, Q_t and Q_t^* have polynomial initial conditions and constant degree.

Proposition 3.3

Let (A, c) be observable. Then the characteristic polynomial of the feedback matrix

$$A-\frac{1}{q_0(t)}k(t)c' \tag{3.30}$$

equals $\dfrac{1}{q_0(t)}Q_t(z)$, i.e.

$$Q_t(z)=q_0(t)\det\left[zI-A+\frac{1}{q_0(t)}k(t)c'\right] \tag{3.31}$$

Proof

Since the determinant is invariant under the similarity transformation (2.20), we may exchange (A, c) for $(\Gamma(a), h)$ in (3.31). Hence, the right member of (3.31) equals

$$q_0(t) \det \left[zI - \Gamma\left(\frac{1}{q_0(t)}\, q(t)\right)\right]$$

which is the same as $Q_t(z)$. ■

We shall now proceed to construct an 'integral' for the system (3.28) and (3.29).

Lemma 3.4

Let $Q_t(z)$ and $Q_t^*(z)$ be defined by (3.26) and (3.27). Then there is a constant polynomial

$$D(z) = d_n z^n + d_{n-1} z^{n-1} + \ldots + d_0 \tag{3.32}$$

such that

$$\frac{1}{q_0(t)}\, [Q_t(z)Q_t(1/z) - Q_t^*(z)Q_t^*(1/z)] = \tfrac{1}{2}[D(z) + D(1/z) - D(0)] \tag{3.33}$$

for all t.

Proof

By using the recursions (3.28) and (3.29) we see that

$$Q_{t+1}(z)Q_{t+1}(1/z) - Q_{t+1}^*(z)Q_{t+1}^*(1/z) = (1-\gamma_t^2)[Q_t(z)Q_t(1/z) - Q_t^*(z)Q_t^*(1/z)]$$

which, in view of (3.15), proves the lemma. ■

By identifying coefficients of z^i in (3.31), we obtain the following $n+1$ integrals for the system (3.18)–(3.20).

$$\sum_{j=0}^{n-i} [q_{j+i}(t)q_j(t) - q_{j+i}^*(t)q_j^*(t)] = \tfrac{1}{2} d_i q_0(t)\,; \quad i = 0, 1, 2, \ldots, n \tag{3.34}$$

where we have defined q_0^* to be identically zero.

Unlike the situation in continuous time, here the degree of the polynomial D determines how far the order of the system (3.18)–(3.20) may be reduced. In fact, if the degree of D is less than n, certain components of q and q^* will become zero after a few time-steps.

Lemma 3.5

The degree of D is n if and only if $\alpha > 0$ and A is non-singular. If D has degree $p < n$ (i.e. $d_i = 0$ for $i > p$ and $d_p \neq 0$), $q_i(t) = 0$ for $i > p$ and $t \geq n - i$, $q_i^*(t) = 0$ for $i > p$ and $t \geq n - i + 1$, and $q_p(t) = \tfrac{1}{2} d_p$ for $t \geq n - p$.

Proof

Reformulate (3.34) to obtain

$$q_n(t) = \tfrac{1}{2} d_n \tag{3.35}$$

$$q_{n-i}(t) = \tfrac{1}{2} d_{n-i} - \frac{1}{q_0(t)} \sum_{j=1}^{i} [q_{j+n-i}(t)q_j(t) - q_{j+n-i}^*(t)q_j^*(t)] \quad \text{for } i = 1, 2, \ldots, n \tag{3.36}$$

Moreover, from (3.20) and (3.21) we have

$$q_i^*(t+1) = q_{i+1}^*(t) - \frac{q_1^*(t)}{q_0^*(t)} q_i(t) \tag{3.37}$$

where $q_{n+1}^* = 0$. In view of (3.23) and (3.35), $d_n = 2\alpha a_n$, and therefore the first statement of the lemma is true, for det $A = a_n$. If $d_n = 0$, $q_n \equiv 0$, and consequently, by (3.37), $q_n^*(t) = 0$ for $t \geqslant 1$. Then, by (3.34), $q_{n-1}(t) = \frac{1}{2} d_{n-1}$ for $t \geqslant 1$. Likewise, if $d_{n-1} = 0$, $q_{n-1}(t) = 0$ for $t \geqslant 1$, and $q_{n-1}^*(t) = 0$ and $q_{n-2}(t) = \frac{1}{2} d_{n-2}$ for $t \geqslant 2$. This procedure carried out in $n-p$ steps proves the lemma. ■

Lemma 3.5 implies that only $2p$ of the recursions (3.18)–(3.20) are needed. We should however remember that $p < n$ only if either there is no measurement noise or A is singular. Such models can be reduced to remove these properties, and therefore we can safely ignore them in our subsequent analysis.

Since $Q_t^*(z) \to 0$ as $t \to \infty$ (Cf. the discussion in § 2), (3.31) has the steady-state version

$$Q_t(z) Q_t(1/z) = \tfrac{1}{2}[D(z) + D(1/z) - D(0)] \tag{3.38}$$

This is the spectral factorization equation corresponding to the discrete-time, algebraic Riccati equation. (Cf. the discussion at the end of § 3.6.)

3.4. *Solution of the integrals*

To facilitate the solution of the integrals we introduce, in analogy with the continuous-time case, the polynomial sequences

$$U_t(z) = \frac{1}{q_0(t)} [Q_t(z) - Q_t^*(z)] \tag{3.39}$$

$$V_t(z) = Q_t(z) - Q_t^*(z) \tag{3.40}$$

Note that U_t is a monic polynomial with coefficients as in (2.38), whereas V_t has leading coefficient

$$v_0(t) = q_0(t) \tag{3.41}$$

Moreover, let $u = (u_1, u_2, \ldots, u_n)'$ and $v = (v_1, v_2, \ldots, v_n)'$.

Lemma 3.6

Let (A, B, c) be a minimal realization, and let $\alpha > 0$. Then for each $t = 0, 1, 2, \ldots, \infty$, U_t is a stability polynomial, i.e. all its zeros have moduli less than one.

Proof

It can be shown (see, e.g. Lindquist 1975 a) that

$$P(t+1) = P(t) - \frac{1}{q_0(t)} q^*(t) q^*(t)' \tag{3.42}$$

Now, since (A, c) is observable, we may without restriction take $A = \Gamma(a)$ and $c = h$. Then the Riccati eqn. (3.10) together with (3.42) gives us

$$P - \frac{1}{q_0} q^* q^{*\prime} = \Gamma(a) P \Gamma(a)' - \frac{1}{q_0} kk' + BB' \tag{3.43}$$

which, in view of (2.44), (3.11) and (3.12), and

$$q^* = k - q_0(u-a)$$

can be reformulated as

$$P = \Gamma(u)P\Gamma(u)' + \alpha(u-a)(u-a)' + BB' \tag{3.44}$$

Therefore, for each fixed t (which, since the limits exist, may be ∞) and $u = u(t)$, $P(t)$ satisfies the Liapunov eqn. (3.42). We shall show that, as a consequence of this, $\Gamma(u)$ is a stability matrix, or, which is an equivalent statement,

$$x_{i+1} = \Gamma(u)'x_i \tag{3.45}$$

is asymptotically stable. Then U_t is a stability polynomial. To this end, use (3.42) to see that

$$x_i'Px_i - x_{i+n}'Px_{i+n} = \sum_{j=1}^{i+n-1} [\alpha((u-a)'x_j)^2 + x_j'BB'x_j] \geqslant 0 \tag{3.46}$$

However, (3.46) cannot be zero. For it if were, in view of (2.44) and (3.45), we would have

$$x_{j+1} = A'x_j\,; \quad B'x_j = 0 \quad \text{for } j = i, i+1, \ldots, i+n-1$$

which violates the assumption of controllability. Therefore, (3.46) must be greater or equal to $\epsilon\|x_i\|^2$ for some $\epsilon > 0$ which does not depend on i. Hence,

$$\epsilon \sum_{j=0}^{\infty} \|x_{jn}\|^2 \leqslant x_0'Px_0 \tag{3.47}$$

and therefore $x_i \to 0$ as $i \to \infty$. The asymptotic stability is now established in the same way as in Lemma 2.3. ∎

Now apply to the transformation (3.39) and (3.40) to (3.33) to obtain

$$U_t(z)V_t(1/z) + U_t(1/z)V_t(z) = D(z) + D(1/z) - D(0) \tag{3.48}$$

which gives us the following $n+1$ equations in the coefficients of U_t and V_t:

$$\sum_{j=0}^{n-i} u_{j+i}(t)v_j(t) + \sum_{j=i}^{n} u_{j-i}(t)v_j(t) = d_i\,; \quad i = 0, 1, \ldots, n \tag{3.49}$$

that is

$$S(u(t))\tilde{v}(t) = d \tag{3.50}$$

where $\tilde{v} = (v_0, v')'$, $d = (d_0, d_1, \ldots, d_n)'$ and

$$S(u) = \begin{bmatrix} 1 & u_1 & u_2 & \ldots & u_n \\ u_1 & u_2 & u_3 & \ldots & 0 \\ u_2 & u_3 & u_4 & \ldots & 0 \\ \hdashline u_n & 0 & 0 & \ldots & 0 \end{bmatrix} + \begin{bmatrix} 1 & u_1 & u_2 & \ldots & u_n \\ 0 & 1 & u_1 & \ldots & u_{n-1} \\ 0 & 0 & 1 & \ldots & u_{n-2} \\ \hdashline 0 & 0 & 0 & \ldots & 1 \end{bmatrix} \tag{3.51}$$

The two matrices in (3.51) are the *Hankel* and the *Toeplitz* matrix respectively for the sequence

$$\ldots, 0, 0, 1, u_1, u_2, \ldots, u_n, 0, 0, \ldots$$

To proceed we shall need the following lemma, the proof of which is postponed to § 3.5.

Lemma 3.7

The Liapunov equation

$$P = \Gamma(u)P\Gamma(u)' + G \tag{3.52}$$

has a unique solution if and only if $S(u)$ is non-singular.

Now, since $\Gamma(u(t))$ is a stability matrix (Lemma 3.6), (3.52) with $u = u(t)$ does have a unique solution, and consequently the inverse of $S(u(t))$ exists. Hence (3.50) can be solved for $\tilde{v}$ to yield

$$\tilde{v}(t) = S(u(t))^{-1}d \tag{3.53}$$

We may note that there are more effective ways to solve the polynomial eqn. (3.48) than to convert it to a system of linear equations. (See § 4.)

3.5. *Formulae for the error covariance*

The discrete-time version of Lemma 2.4 reads :

Lemma 3.8

Let (A, c) be in the companion form $(\Gamma(a), h)$. Then the system of first-order difference equations consisting of (3.10), (3.18), (3.19) and (3.20) has the following set of integrals :

$$P_{i+1,\, j+1} - P_{ij} = \frac{1}{q_0}[q_i q_j - q_i{}^* q_j{}^*] - \alpha a_i a_j - (BB')_{ij} \tag{3.54}$$

which holds for $t = 0, 1, 2, \ldots$ and $i, j = 0, 1, 2, \ldots, n$, provided that we extend the definitions of P, B and a as in Lemma 2.5.

Proof

In view of (3.11), (3.12) and (2.24), we may write (3.22) (with $T = I$) as

$$JPh = q - \alpha a \tag{3.55}$$

which gives us (3.54) for the case that either i or j is zero. To see that (3.54) is also true for $i, j = 1, 2, \ldots, n$, insert (2.24) and (3.22) into (3.43) and use (3.55) to obtain

$$JPJ' - P = \frac{1}{q_0}[qq' - q^* q^{*\prime}] - \alpha aa' - BB' \;\blacksquare$$

Lemma 3.9

Given all the assumptions of Lemma 3.8, we have

$$P(t) = \Pi(u(t), v(t), B) \tag{3.56}$$

where

$$\Pi_{ij}(u, v, B) = \sum_{k=i}^{n+i-j} [\alpha a_k a_{k+j-i} + (BB')_{k,\, k+j-i} - \tfrac{1}{2} u_k v_{k+j-i} - \tfrac{1}{2} u_{k+j-i} v_k] \tag{3.57}$$

Here u_i and v_i are taken to be zero whenever $i < 0$ or $i > n$. Moreover, the coefficients $d = (d_0, d_1, \ldots, d_n)'$ of (3.32) are

$$d = \delta(B) \tag{3.58}$$

where

$$\delta_i(B) = 2 \sum_{j=0}^{n-i} [\alpha a_j a_{j+i} + (BB')_{j,\, j+i}] \tag{3.59}$$

Proof

To see that (3.56) is true, just exchange $(1/q_0)\ [q_i q_j - q_i{}^* q_j{}^*]$ for $\tfrac{1}{2} u_i v_j + \tfrac{1}{2} v_i u_j$ in (3.54) and perform the appropriate summation. By taking $i = 0$ in (3.57), which then equals zero, we obtain an alternative derivation of (3.49), and therefore (3.58) follows. ■

Proof of Lemma 3.7

Apply the method of Lemmas 3.8 and 3.9 to (3.52) to obtain

$$P_{ij} = \sum_{k=i}^{n+i-j} [a_k P_{1,\, k+j-i+1} + P_{1,\, k+1} a_{k+j-i} - P_{11} a_k a_{k+j-i} - G_{k,\, k+j-i}] \tag{3.60}$$

for $i, j = 0, 1, \ldots, n$, where $P_{ij} = G_{ij} = 0$ whenever some index is zero or $n+1$, and $a_0 = 1$. Now taking $i = 0$ in (3.60), we have

$$S(a)z = b \tag{3.61}$$

where $b = (b_0, b_1, \ldots, b_n)'$ is given by

$$b_i = \sum_{j=0}^{n-i} G_{j,\, j+i} \tag{3.62}$$

and $z = (z_0, z_1, \ldots, z_n)'$ is related to P through the relation

$$z_i = P_{1,\, i+1} - \tfrac{1}{2} a_i P_{11} \tag{3.63}$$

Once z is known, we can also determine P (and vice versa). Indeed, the first column of P can be determined from (3.63), whereupon the rest of P is given by (3.60). Therefore (3.52) has a unique solution if and only if (3.61) has a unique solution. ■

This proof is constructive and it provides us with an algorithm to solve a Liapunov equation. An analogous algorithm for the continuous-time case (involving the inversion of the Hurwitz matrix $H(a)$) can be obtained using the methods of Lemmas 2.4 and 2.5. This algorithm proves the continuous-time counterpart of Lemma 3.7, which states that the existence of $H(a)^{-1}$ is necessary and sufficient for the Liapunov equation to have a unique solution. (Cf. Lehnigk 1966, p. 44.)

6 G 2

3.6. *The reduced-order non-Riccati system*

Let $e_0, e_1, \ldots, e_n$ be the unit vectors in R^{n+1}, and let $f: R^n \to R^n$ be the function

$$f_i(u) = \frac{e_i' S(u)^{-1} \delta(T^{-1}B)}{e_0' S(u)^{-1} \delta(T^{-1}B)}, \quad i = 1, 2, \ldots, n \tag{3.64}$$

where S and δ are defined by (3.51) and (3.59) respectively.

Theorem 3.10

Let (A, B, c) be a minimal realization, and let $\alpha > 0$. Then the gain vector sequence in the filter eqn. (3.9) is given by

$$\frac{1}{q_0(t)} k(t) = \tfrac{1}{2} T[u(t) + f(u(t)) - 2a] \tag{3.65}$$

where u is the solution of the system of n first-order equations

$$\left.\begin{aligned} &u(t+1) = [2 + u_1(t) - f_1(u(t))]^{-1}[(I+J)u(t) + (I-J)f(u(t))] \\ &u(0) = a \end{aligned}\right\} \tag{3.66}$$

The error variance q_0 is given by

$$q_0(t) = e_0' S(u(t)) \delta(T^{-1}B) \tag{3.67}$$

and the solution P of the matrix Riccati eqn. (3.10) by

$$P(t) = T\Pi(u(t), q_0(t) f(u(t)), T^{-1}B)T' \tag{3.68}$$

where Π is defined by (3.53).

Proof

Apply the transformations (3.39) and (3.40) and (3.15) to the recursions (3.28) and (3.29) to obtain

$$U_{t+1} = \tfrac{1}{2}(1 - \gamma_t)^{-1}[(1+z)U_t + \frac{1}{q_0}(1-z)V_t]$$

which can also be written

$$u(t+1) = \tfrac{1}{2}(1 - \gamma_t)^{-1}[(I+J)u(t) + \frac{1}{q_0(t)}(I-J)v(t)] \tag{3.69}$$

Since Lemmas 3.6 and 3.7 apply, we can use (3.53), which together with (3.58) and (3.41), yields (3.67) and

$$v(t) = q_0(t) f(u(t)) \tag{3.70}$$

where the argument of δ has been adjusted to account for the fact that Lemma 3.9 must be applied to the transformed triplet $(\Gamma(a), h, T^{-1}B)$. Therefore, in view of the fact that $\gamma = \tfrac{1}{2}(1/q_0)v_1 - \tfrac{1}{2}u_1$, (3.66) holds. It is clear from (3.19) and (3.20) that $u(0) = a$. Equation (3.65) follows from (3.22) and (3.70). Finally, (3.68) is obtained from Lemma 3.9 after applying the transformations (2.20)–(2.23). ■

Equation (3.66) has the *steady-state version*

$$u = f(u) \tag{3.71}$$

To see this, just put $u(t+1) = u(t) = u$ in (3.66) to obtain

$$\Gamma(u)[u - f(u)] = u - f(u) \tag{3.72}$$

With an argument analogous to the one in the continuous-time case, we can show that P being non-increasing ensures the convergence of $u(t)$ to $u(\infty) = q_0(\infty)q(\infty)$ as $t \to \infty$, and that $u(\infty)$ must satisfy (3.72). However, by Lemma 3.6, $\Gamma(u(\infty))$ is a stability matrix, and hence it has no eigenvalue equal to one. Consequently, $u(\infty)$ satisfies (3.71).

If either A is singular ($a_n = 0$) or there is no measurement noise ($\alpha = 0$), according to Lemma 3.5, $u_n(t)$ and $v_n(t)$ will be zero for $t \geqslant 1$. In fact, if D has degree p, for $i = p+1, p+2, \ldots, n$, $u_i(t) = v_i(t) = 0$ whenever $t \geqslant n - i + 1$. Therefore it is possible to construct a reduced-order algorithm with even fewer equations. However, this will be the topic of a future study.

4. Concluding remarks

We have attempted to provide some further theoretical insight into the structure of the non-Riccati algorithms (2.11) and (2.12) and (3.13)–(3.15). To this end we have derived a set of integrals for each of them. The integrals have been used to reduce the number of first-order dynamic equations. Unlike the original non-Riccati algorithms, these reduced-order systems have non-trivial steady-state versions, which are equivalent to the algebraic equations of the spectral factorization. Therefore, in a certain sense, these equations provide a link between Wiener and Kalman–Bucy filtering techniques.

The reduced-order systems contain an inverse of a Hurwitz matrix (in continuous time) or the sum of a Toeplitz and a Hankel matrix (in discrete time). Fast algorithms for these inversions can be constructed by applying the Euclidean algorithm as it appears in Berlekamp (1968) to the polynomial relations (2.48) and (3.48) respectively. (The author is indebted to Professors Eakin and Sathaye for pointing out this.) Nevertheless, in the end we may find that computational considerations will cause us to retain the original non-Riccati system and instead use the integrals to control the convergence. In any case, we think that the results of this paper will prove valuable in studying the numerical properties of the algorithms.

In our analysis, the assumption of stationarity is used essentially only to secure that the decompositions (2.42) and (3.42) hold. The appropriate modifications for the non-stationary case will be discussed elsewhere. The relations to the inverse problem of stationary covariance generation (Anderson 1969) is presently studied by G. Picci and the author.

Acknowledgments

The author would like to thank Professors W. T. Cashman and Z. Schuss with whom he had some stimulating discussions concerning the continuous-time case. The final version of the paper has profited from discussions with Professors R. W. Brockett, G. Picci and L. E. Zachrisson.

Note added in proof.—It can be shown (Lindquist 1976) that Theorems 2.8 and 3.10 still hold when the controllability assumption is removed.

References

AKIEZER, N. I., 1965, *The Classical Moment Problem* (Hafner).

ANDERSON, B. D. O., 1969, *J. Statist. Phys.*, **1,** 133.

BERLEKAMP, E. R., 1968, *Algebraic Coding Theory* (McGraw-Hill).

BROCKETT, R. W., 1970, *Finite Dimensional Linear Systems* (Wiley).

BUCY, R. S., RAPPAPORT, D., and SILVERMAN, L. M., 1970, *I.E.E.E. Trans. autom. Control*, **15,** 535.

BUELENS, P. F., and HELLINCKX, L. J., 1974, *Int. J. Control*, **20,** 113.

CASTI, J., 1974, *Utilitas Mathematica*, **6,** 95.

CASTI, J., KALABA, R., and MURTHY, V. K., 1972, *I.E.E.E. Trans. Inform. Theory*, **4,** 515.

CASTI, J., and TSE, E., 1972, *J. math. Analysis Applic.*, **40,** 45.

GANTMACHER, F. R., 1959, *Theory of Matrices*, Vol. II (Chelsea).

GERONIMUS, L. YA., 1961, *Orthogonal Polynomials* (Consultant Bureau).

KAILATH, T., 1973, *I.E.E.E. Trans. Inf. Theory*, **19,** 750.

KAILATH, T., MORF, M., and SIDHU, S., 1973, *Proc. Seventh Princeton Symp. on Inform. and System Sciences.*

KREIN, M., 1955, *Dokl. Akad. Nauk SSSR*, **105,** 637.

LENIGK, S. H., 1966, *Stability Theorems for Linear Motions* (Prentice-Hall).

LINDQUIST, A., 1974 a, *J. SIAM Control*, **12,** 736 ; 1974 b, *Ibid.*, **12,** 747 ; 1975 a, *Int. J. appl. Math. Optim.*, **1,** 355 ; 1975 b, *Control Theory, Numerical Methods and Computer Systems Modelling*, Lecture Notes in Economics and Maths Systems, Vol. 107, p. 44 ; 1976, *Appl. Math. Optim.* (to appear).

MOULTON, F. J., 1930, *Differential Equations* (Mcmillan).

RISSANEN, J., 1973, *I.E.E.E. Trans. autom. Control*, **18,** 555 ; 1975, *Proc. Sixth IFAC Congr.*, Boston, Mass.

RUGH, W. J., and MURPHY, G. J., 1969, *J. bas. Engng*, D, **91,** 149.

WILLIS, B. H., and BROCKETT, R. W., 1965, *I.E.E.E. Trans. autom. Control*, **10,** 262.