Time-Varying Polynomial Systems Approach to Multichannel Optimal Linear Filtering

M.J. Grimble

Industrial Control Centre University of Strathclyde
Graham Hills Building 50 George Street GLASGOW G1 1QE

ABSTRACT

A new approach to linear estimation in time-varying discrete multivariable systems is described. The signal model is taken to be a time-varying vector difference equation which can be expressed in ARMA polynomial system form. An optimal linear filter and predictor is derived in terms of time-dependent polynomial operators and this can also be implemented as a recursive algorithm using difference equations. The system model and filter are particuarly relevant in self-tuning filtering applications.

1. INTRODUCTION

The polynomial systems approach to linear estimation and control problems has mainly been restricted to consideration of time-invariant stationary systems (*Kucera* [1]). This may often be adequate for fixed systems but is not so appropriate in situations where adaptive techniques are required which are by definition time variable. In such cases a slowly time-varying model for the signal source or plant is often needed. Since identification methods can provide estimates of the polynomial system model, expressions are needed to enable time-varying filters to be computed in this polynomial form.

The Kalman filter [2] does of course provide a solution to the time-varying non-stationary estimation problem but its state-space structure is not really suitable for adaptive estimation problems. Thus, the objective here is to solve this estimation problem using a time-varying polynomial system operator description for the signal and noise sources. The notion of a continuously changing signal spectrum (*Priestley* [3]) arises and an *innovations signal* description is employed (*Anderson and Moore* [4]).

One of the first predictors for non-stationary processes was derived by *Whittle* [5] but the solution was only explicit in certain special cases. *Priestley and co-workers* generalised and extended this approach but multi-channel estimation problems and models including coloured measurement noise were not considered. These generalisations are obtained in the following. The results should be valuable in feedback control problems and in signal processing applications, or time-series prediction problems.

2. SIGNAL GENERATING PROCESS

Since both control and signal processing problems are of interest the discrete system model shown in Fig. 1 can represent either an industrial plant or a message generating process. The system is assumed to be linear and possibly time-varying. The noise sources can be non-stationary where their second-order properties are known. The noise signals $\xi(t) \in R^{qd}$, $\omega(t) \in R^{qn}$ and $v(t) \in R^r$ are mutually independent and trend free, with zero means and covariances defined as: $cov[\xi(t),\xi(\tau)] = Q_d(t)\delta_{t\tau}$, $cov[\omega(t),\omega(\tau)] = Q_n(t)\delta_{t\tau}$, $cov[v(t),v(\tau)] = R(t)\delta_{t\tau}$ respectively. Here $\delta_{t\tau}$ denotes the Kronecker delta function and the assumption is made that $R(t) = R^T(t) \quad 0$. Let $Q(t) = block\ diagonal\ \{Q_d(t), Q_n(t)\}$.

2.1 *Operator form of signal model*

The system is assumed to be in operation from time $t_0 \to -\infty$ and for simplicity any known inputs (such as control signals) are neglected, since these contribute to the filter output in a straightforward manner. The results obtained are to be related to the Kalman filtering problem and thus the system is assumed to have an underlying state-space description which is completely observable, and controllable from the noise inputs. This system model $y = W\xi$ may be represented by the one-sided moving average processes:

$$y(t) = (W_d\xi)(t) \triangleq \sum_{\tau=-\infty}^{t} W_d(t,\tau)\xi(\tau) = \sum_{j=0}^{\infty} W_{dj}(t)\xi(t-j)$$

and (1)

$$n(t) = (W_n\omega)(t) \triangleq \sum_{\tau=-\infty}^{t} W_n(t,\tau)\omega(\tau) = \sum_{j=0}^{\infty} W_{nj}(t)\omega(t-j)$$

where (2)

$W_{dj}(t) \triangleq W_d(t, t-j)$ and $W_{nj}(t) \triangleq W_n(t, t-j)$ for $j = 0, 1, 2, \ldots$ where $\{v(t)\}$ denotes white measurement noise and $\{n(t)\}$ represents a coloured noise or output disturbance. The time-varying signal $\{y(t)\}$ might for example represent a *fading* and *wandering* radio signal. All signals are considered to be members of the spaces of doubly infinite vector sequences $\{f(t): t = \ldots -1, 0, 1, \ldots\}$ with a unity sampling interval.

2.2 *ARMA signal descriptions*

To identify the form of the signal W_d and measurement noise W_n models, consider the time-varying difference equations which are assumed to represent these operators:

$$A(t;z^{-1})y(t) = C_d(t;z^{-1})\xi(t) \quad (4)$$

$$A(t;z^{-1})n(t) = C_n(t;z^{-1})\omega(t) \quad (5)$$

where without loss of generality (A, C_d) are of the form:

$$A(t;z^{-1}) = I_r + A_1(t)z^{-1} + ... + A_n(t)z^{-n} \quad (6)$$

$$C_d(t;z^{-1}) = C_{do}(t) + C_{d1}(t)z^{-1} + ... + C_{dn}(t)z^{-n} \quad (7)$$

$$C_n(t;z^{-1}) = C_{no}(t) + C_{n1}(t)z^{-1} + ... + C_{nn}(t)z^{-n} \quad (8)$$

The measurement noise model or output disturbance model (A_n, C_n) can be taken to be asymptotically stable. The time-varying operators W_d and W_n can be expressed, as for time-invariant systems, in the form:

$$W_d(t;z^{-1}) = A(t;z^{-1})^{-1}C_d(t;z^{-1}) \quad (9)$$

$$W_n(t;z^{-1}) = A(t;z^{-1})^{-1}C_n(t;z^{-1}) \quad (10)$$

3. OUTPUT FILTERING PROBLEM

The first filtering problem is concerned with output estimation, that is the problem of finding the best estimate of the signal $y(t)$ in the presence of the noise terms $v(t)$ and $n(t)$.

3.1 *Minimum variance criterion*

Let the output estimation error be defined as:

$$\tilde{y}(t) \triangleq y(t) - \hat{y}(t) \quad (11)$$

where $\hat{y} \triangleq \hat{y}(t|t)$ is the estimate of $y(t)$, given observations $\{z(\tau)\}$ over the semi-infinite interval, $\tau\varepsilon(-\infty, t]$. The average variance is to be minimised and is given, in terms of the trace function, as:

$$J = \lim_{T\to\infty} \frac{1}{2T} \sum_{t=-T}^{T} E\{\tilde{y}^T \tilde{y}(t)\}$$

$$= \lim_{T\to\infty} \frac{1}{2T} \sum_{t=-T}^{T} Tr\{E\{\tilde{y}(t)\tilde{y}^T(t)\}\} \quad (12)$$

The time-averaging operation is necessary since for time-invariant plants and stationary noise the steady-state value of $E\{\tilde{y})^T(t)\tilde{y}(t)\}$ is a constant, and the sum of an infinite number of terms would be infinity. However, using the above expression in this case J reduces to the usual steady-state criterion $J = E\{\tilde{y}^T(t)\tilde{y}(t)\}$. Moreover, by minimizing J the function $E\{\tilde{y}^T(t)\tilde{y}(t)\}$ is minimized, for each t, even in the time-varying case.

3.2 *Estimation error equation*

The filtered estimate $\hat{y}(t)$ is assumed to be generated from a linear causal time-varying estimator of the form:

$$\hat{y}(t) = H_f(t;z^{-1})z(t) \quad (13)$$

where H_f denotes a minimal realisation of the optimal estimator. Since an infinite time ($t_o \to -\infty$) problem is of interest no initial condition term is required.

To obtain an expression for the estimation error $\tilde{y}(t)$ note from (3) to (5):

$$z = v + A^{-1}C_n\omega + A^{-1}C_d\xi \quad (14)$$

but a realisation of $\{z(t)\}$ can be obtained using the innovations signal model:

$$z = A^{-1}D_f\varepsilon \quad (15)$$

where the zero mean white noise signal $\{\varepsilon(t)\}$ has identity covariance matrix. From (11) and (13) obtain:

$$\tilde{y} = y - \hat{y} = A^{-1}C_d\xi - H_f z \quad (16)$$

where statistically the signal $\{z(t)\}$ can be represented in either of the equivalent forms (14) or (15).

3.3 *Solution of the output estimation problem*

The signal and noise models were defined in Section 2 and the variance to be minimised was given as (12). Proceeding with the solution, the average covariance of the estimation error can be expressed, using (16) as:

$$J = \lim_{T\to\infty} \frac{1}{2T} E\{\tilde{y}, \tilde{y}\} \quad (17)$$

$$= \lim_{T\to\infty} \frac{1}{2T} E\{< H_f A^{-1} D_f \varepsilon, H_f A^{-1} D_f \varepsilon >_{H_r}$$

$$- < H_f A^{-1} C_d \xi, A^{-1} C_d \xi >_{H_r}$$

$$- < A^{-1} C_d \xi, H_f A^{-1} C_d \xi >_{H_r} + < A^{-1} C_d \xi, A^{-1} C_d \xi >_{H_r} \quad (18)$$

(where the independence of ξ and ω and v has been used to simplify the expression). Note for later use that (17) may also be written, using (14), in the form:

$$J = \lim_{T\to\infty} \frac{1}{2T} E\{< (I - H_f) A^{-1} C_d \xi, (I - H_f) A^{-1} C_d \xi >_{H_r}$$

$$+ < H_f v, H_f v >_{H_r} + < H_f A^{-1} C_n \omega, H_f A^{-1} C_n \omega >_{H_r}\} \quad (19)$$

Using (34) equation (18) may be written as:

$$J = \lim_{T\to\infty} \frac{1}{2T} \sum_{t=-T}^{T} Tr\{H_f A^{-1} D_f D_f^* A^{*-1} H_f^*$$

$$+ A^{-1} C_d Q_d C_d^* A^{*-1}$$

$$- H_f A^{-1} C_d Q_d C_d^* A^{*-1} - A^{-1} C_d Q_d C_d^* A^{*-1} H_f^* - T(z;z^{-1})\} \quad (20)$$

where $T(z;z^{-1})$ is defined so that the kernel of the above summation includes no terms involving powers of z or z^{-1}. Thus, in expanding the operators within the kernel the contribution to the cost J is from terms in z^0. Completing squares gives:

$$J = \lim_{T\to\infty} \frac{1}{2T} \sum_{t=-T}^{T} Tr\{(H_f A^{-1} D_f - A^{-1} C_d Q_d C_d^* D_f^{*-1})$$

$$x(H_f A^{-1}D_f - A^{-1}C_dQ_dC_d^*D_f^{*-1})^* + (A^{-1}C_dQ_dC_d^*A^{*-1}$$

$$-A^{-1}C_dQ_dC_d^*D_f^{*-1}D_f^{-1}C_dQ_dC_d^*A^{*-1} - T(z;z^{-1})))\} \quad (21)$$

Diophantine equations

To minimise this cost-function two diophantine equations must be introduced. The existence and uniqueness of the solution to these equations is discussed in Section 3.5. Assuming for the present the existence of a solution let $(G_o(t;z^{-1}),\ S_o(t;z^{-1}),$ and $F_o(t;z^{-1})$, denote the minimal degree solution, with respect to $F_o(t;z^{-1})$, of the equations:

$$AF_o + G_oD_f^*z^{-g} = C_dQ_dC_d^*z^{-g} \quad (22)$$

$$-AF_o + S_oD_f^*z^{-g} = (ARA^* + C_nQ_nC_n^*)z^{-g} \quad (23)$$

where $g \triangleq deg\ D_f$. From (22) obtain:

$$F_oz^gD_f^{*-1} + A^{-1}G_o = A^{-1}C_dQ_dC_d^*D_f^{*-1}$$

since deg $F_o = g$ the first term can be expanded as a convergent sequence in terms of positive powers of z. The *squared* term in the cost-function (21) can thus be written as:

$$(H_fA^{-1}D_f - A^{-1}C_dQ_dC_d^*D_f^{*-1}) = [H_fA^{-1}D_f - A^{-1}G_o]$$
$$-F_oz^gD_f^{*-1} \quad (24)$$

To show that the first term [.] on the right of (24) can be expanded as a convergent series in terms of negative powers of z, an *implied* diophantine equation must be introduced. Appropriately multiplying (22) and (23):

$$F_oD_f^{*-1} + A^{-1}G_oz^{-g} = A^{-1}C_dQ_dC_d^*D_f^{*-1}z^{-g} \quad (25)$$

$$-F_oD_f^{*-1} + A^{-1}S_oz^{-g} = (R + A^{-1}C_nQ_nC_n^*A^{*-1})A^*D_f^{*-1}z^{-g}$$

Adding (25) and (26) gives: (26)

$$G_o + S_o = D_f \quad (27)$$

The first term in (24) may now be expressed, using (27), as:

$$[H_fA^{-1}D_f - A^{-1}G_o] = [(H_f - I)A^{-1}G_o + H_fA^{-1}S_o] \quad (28)$$

For the cost-function (20) to be finite $(I-H_f)A^{-1}$ and $H_fA^{-1}C_n$ must represent asymptotically stable causal system matrices (using (23) this can also be shown to apply to $H_fA^{-1}S_o$). Thus, (24) separates into an asymptotically stable causal term [.] and a strictly non-causal term $F_oz^gD_f^{*-1}$

Minimization step

The cost-function (21) may now be expanded as:

$$J = \lim_{T\to\infty}\frac{1}{2T}\sum_{t=-T}^{T} Tr\{H_fA^{-1}D_f - A^{-1}G_o]$$

$$x[H_fA^{-1}D_f - A^{-1}G_o]^*$$

$$+[F_oD_f^{*-1}D_f^{-1}F_o^* + A^{-1}C_dQ_dC_d^*A^{*-1}$$

$$-A^{-1}C_dQ_dC_d^*D_f^{*-1}D_f^{-1}C_dQ_dC_d^*A^{*-1}]$$

$$-([H_fA^{-1}D_f - A^{-1}G_o]D_f^{-1}F_o^*z^g + F_oz^gD_f^{*-1}$$

$$x(H_fA^{-1}D_f - A^{-1}G_o]^* + T(z;z^{-1})))\} \quad (29)$$

Inspection of the cross terms in the kernel of the above summation reveals that the term $[H_fA^{-1}D_f - A^{-1}G_o]D_f^{*-1}F_oz^{-g}$ involves only negative powers of z *(deg F < g)* and

$F_oz^gD_f^{*-1}[H_f - A^{-1}D_fA^{-1}G_o]^*$ involves only positive powers of z. Thus, the contribution to the kernel, at time *t*, of the final terms in (29) is zero, since the coefficient of z^0 is null. The second group of terms in (29) (within the square brackets) are independent of the filter operator H_f and do not enter the minimisation procedure. The minimum variance is therefore achieved by setting the coefficients of z^0 in the first term of (29) to zero, or equivalently by setting the term $[H_fA^{-1}D_f - A^{-1}G_o]$ to zero.

The optimal filter is obtained in its minimal form using:

$$H_f = A^{-1}G_oD_f^{-1}A \quad (30)$$

or from equation (27) noting $A^{-1}G_o = A^{-1}(D_f - S_o)$:

$$H_f = I_r - A^{-1}S_oD_f^{-1}A \quad (31)$$

Collecting the above results the following theorem is obtained:

Theorem 3.1 : ***Optimal linear filter***

The optimal time-varying causal linear filter for the signal model described in Section 2, to minimise the average covariance of the estimation error (17), is given as:

$$H_f = A^{-1}G_oD_f^{-1}A \quad (32)$$

or alternatively as:

$$H_f = I_r - A^{-1}S_oD_f^{-1}A \quad (33)$$

The polynomial operators G_o or S_o are obtained from the minimal degree solutions (G_o, F_o) or (S_o, F_o), with respect to F_o, of the diophantine equations (22) or (23), respectively. The operator D_f^{-1} is asymptotically stable and is defined from the time-dependent spectrum:

$$D_fD_f = C_dQ_dC_\alpha^* + ARA^* + C_nQ_nC_n^* \quad (34)$$

3.4 *White measurement noise : Kalman filtering problem*

The special case of the time-varying Kalman filtering problem where the measurement noise is white $(C_n = 0)$, will now be considered. These results should be compared with those of for the time-invariant situation.

Lemma 3.1 : *Kalman filtering problem*

The Kalman filter is obtained as:

$$H_f = I_r - \tilde{S}_o D_f^{-1} A \tag{35}$$

where $S_o = A\tilde{S}_o$ follows from (22) and (23), and D_f from (34).

Proof: If $C_n = 0$ then from (23) S_o must be of the form $A\tilde{S}_o$, and (35) follows from (33). Note that in this case (23) may be replaced by the reduced equation:

$$-F_o + \tilde{S}_o D_f^{-1} z^{-g} = RA^* z^{-g} \tag{36}$$

Example 1 : ***Time-Varying System Filtering Problem***

Consider the system model defined in state equation form:

$$x(t+1) = a(t)x(t) + \xi(t) \tag{37}$$
$$y(t) = x(t)$$
$$z(t) = y(t) + v(t)$$

The independent zero mean noise sources have variances $E\{\xi^2(t)\} = Q_d = 1$ and $E\{v^2(t)\} = R = 1$. The output disturbance is absent in this problem ($Q_n = 0$).

Solution

This example is used to relate the present results to the familiar Kalman filtering problem. The system model may be represented by the polynomial operator:

$$W_d(t;z^{-1}) = A(t;z^{-1})^{-1} C_d(t;z^{-1})$$

$$= \frac{1}{(1-a(t-1)z^{-1})} z^{-1}\xi(t)$$

Computation of D_f: The spectral factor satisfies equation (34):

$$D_f D_f^* = C_d Q_d C_d^* + ARA^* \tag{38}$$

where $D_f(t;z^{-1}) = d_0(t) + d_1(t-1)z^{-1}$. A method of computing $d_0(t)$ and $d_1(t)$ is described below but note that in self-tuning filtering problems D_f will be identified directly and this calculation can be avoided.

General diophantine equations: In a general filtering problem either G_o can be found from (22), or S_o from (23). For the sake of illustration both of these calculations are considered below. Since $\deg(F_o) \le g - 1$, then $F_o = f_o(t)$, $G_o = g_o(t)$ and $S_o = s_o(t) + s_1(t)z^{-1}$; thus

$$(1-a(t-1)z^{-1})f_o(t) + g_o(t)(d_o(t)+d_1(t)z)z^{-1} = z^{-1} \tag{39}$$

and

$$-(1-a(t-1)z^{-1})f_o(t) + (s_o(t)+s_1(t)z^{-1})$$
$$\times(d_o(t)+d_1(t)z)z^{-1}$$
$$= (1-a(t-1)z^{-1})(1-za(t-1))z^{-1} \tag{40}$$

Multiplying out these equations and equating coefficients

$$f_o(t) + g_o(t)d_1(t) = 0$$
$$-a(t-1)f_o(t-1) + g_o(t)d_o(t) = 1$$

and

$$-f_o(t) + s_o(t)d_1(t) = -a(t)$$
$$a(t-1)f_o(t-1) + s_o(t)d_o(t) + s_1(t)d_1(t-1)$$
$$= 1 + a^2(t-1)$$
$$s_1(t)d_o(t-1) = -a(t-1)$$

These equations can be solved recursively for $(f_o(t), g_o(t))$ or $(f_o(t), s_o(t), s_1(t))$ given $f_o(t-1)$.

Kalman filtering diophantine equations

For this particular problem (with $C_n = 0$) the second diophantine equation (23) can be simplifed as described in Lemma 3.1. From (36) noting $\deg(F_o) \le g - 1$ gives:

$$-f_o(t) + s_o(t)(d_o(t)+d_1(t)z)z^{-1} = (1-a(t)z)z^{-1} \tag{41}$$

Multiplying out and equating coefficients obtain:

$$-f_o(t) + s_o(t)d_1(t) = -a(t) \tag{42}$$
$$s_o(t)d_o(t) = 1 \tag{43}$$

giving $s_o(t) = d_o(t)^{-1}$ and $f_o(t) = d_o(t)^{-1}d_1(t) + a(t)$. The filter now follows from (35).

$$H_f(t;z^{-1}) = 1 - d_o(t)D_f(t;z^{-1})^{-1}(1-a(t-1)z^{-1}) \tag{44}$$

5. Conclusions

The general approach at representing time-varying non-stationary systems in polynomial operator form and the subsequent optimal solution procedure is original, and can be applied to a range of linear estimation and control problems. The signal model is also in a very appropriate form for signal processing applications.

When the signal model is unknown and must be identified on-line, a self-tuning approach may be taken based upon the proposed estimator. The time-varying spectral factor $D_f(t;z^{-1})$ can then be identified from the signal measurements and a single diophantine equation can be solved to obtain the filter transfer operator. By this means simple adaptive estimators may be constructed.

REFERENCES

1. V. Kucera, *Discrete Linear Control*, John Wiley & Sons, Chichester, 1979.
2. R.E. Kalman, *A new approach to linear filtering and prediction problems*, Jour. of Basic Eng, 1960, 35-45.
3. M.B. Priestley, *Spectral Analysis and Time Series*, Vol. 1, Univariate Series, Vol. 2, Multivariate Series Prediction and Control, London: Academic Press, 1981.
4. B.D.O. Anderson, and J.B. Moore, *Optimal Filtering*, New Jersey; Prentice Hall, 1979.
5. P. Whittle, *Recursive relations for predictors of non-stationary processes*, J. Roy Stat. Soc. Ser. B. Vol 27, pp 523-532, 1965.
6. P.T.K. Fung, X.P. Chen, and M.J. Grimble, *Adaptive tracking of slowly varying processes with coloured noise disturbances*, Trans. Inst. Meas, and Control, 1985.