Monograph:

Nicholson, H (1969) Sequential Least-Squares Estimation, Identification, Reduction and Control. Part 2. Research Report. ACSE Report CE69-3 . Department of Control Engineering, University of Sheffield, Mappin Street, Sheffield

University of Sheffield

Department of Control Engineering

Sequential least-squares estimation, identification, reduction and control

Part 2

H. Nicholson

Research report CE69 - 3

November 1969

Summary

The study of Part 1 concerned with least-squares estimation, identification and prediction is extended to include problems concerning the generalised inverse of a singular matrix, the optimal control of the linear dynamic system based on finite- and infinite-time performance criteria, and the reduction of system order. The solution of the time-varying and steady-state matrix Riccati differential equation associated with the optimal filtering and control problem is developed on the basis of a transition matrix and eigenvector representation. The sensitivity of the sequential algorithms for state estimation, prediction and optimal control is also investigated.

The paper develops a number of important computational techniques which complement the algorithms outlined in Part 1 to form a comprehensive review of certain aspects of linear system theory which are suitable for direct practical application. Current research work being undertaken in the general field of process modelling and online control which will investigate the application of many of the techniques developed in the present study is outlined.

Introduction

The study of Part 1 concerned with state estimation, parameter identification and prediction is extended in Part 2 to consider other closely related problems in order to form a comprehensive review of the computational techniques available for investigating the optimal control of the linear stochastic dynamic system. The concept of a generalised inverse of a singular matrix is associated inherently with least-squares minimisation, and practical methods of solution, including methods of polar decomposition, partitioning, and iterative and recursive algorithms similar to those developed for least-squares filtering are reviewed. The optimal control of the linear dynamic system based on finite- and infinite-time quadratic performance criteria is studied using the maximum principle and the discrete and continuous form of dynamic programming, and practical computational methods for solution of the time-varying and steady-state matrix Riccati equations are developed.

The study illustrates the close relationship of the optimal control problem based on quadratic minimisation to the problem of optimal estimation based on sequential least-squares theory. The discrete algorithms obtained for state estimation and optimal control are also shown to reduce in the limit to the form of the continuous-time Riccati covariance-type equation and to the optimal control law given by the maximum principle. The steady-state matrix Riccati equation is solved explicitly within the structure of an eigenvalue analysis, and a similar form of solution is obtained using a transition matrix representation. The techniques are also shown to be applicable for solution of the Lyapunov equation which forms an inherent component of the covariance-type Riccati equation. Practical methods for computation of the discrete time response of the linear system, including techniques based on the matrix-series expansion, an eigenvalue analysis, Sylvester's expansion theorem and a partial fraction expansion are reviewed. The constituent matrices of Sylvester's expansion related to the algebraic eigenvalue problem are seen to play an important role in many aspects of linear system theory, particularly with reference to the spectral analysis, the generalised inverse and the residual vector of least-squares theory.

The high-order process model used for dynamic optimisation and stochastic approximation will introduce problems of dimensionality, and a reduction of system order for isolating the significant dynamics represents an important area of study in linear system theory. Methods of reduction which have been developed on the basis of modal analysis and the concepts of least-squares estimation theory are outlined. The review is completed with a sensitivity analysis of least-squares estimation, of the spectral prediction algorithm based on eigenvalue and eigenvector sensitivity and of the optimal control algorithms which may be required for assessing the effects of inaccurate data and plant parameters. The paper also outlines a range of research work which is currently being undertaken in the general field of process modelling and online computer control which will investigate the application of the general techniques of least-squares estimation, identification, reduction, control and sensitivity.

6. Generalised inverse of a rectangular or singular matrix

The concept of a generalised or pseudo inverse of a rectangular or singular square matrix analogous to the regular inverse of a nonsingular matrix was introduced by Moore (1920, 1935)[92,93] and rediscovered independently by Penrose (1955)[94,95], Bjerhammar (1951)[96] and Murray and von Neumann (1936)[97]. The problem has particular relevance to the least-square minimisation solution of inconsistent sets of normal algebraic equations[98] associated with problems of filtering, identification, prediction and control. A detailed account of the development of the generalised inverse is given by Ben-Israel and Charnes[99] and Rao[100]. The various methods of solution proposed, and the close relation existing between the concept of the generalised inverse and the methods of least squares previously discussed are illustrated.

The problem is concerned with obtaining a solution of the linear equations

$$A\,x = y \tag{148}$$

where in general A is a matrix of order mxn and rank r (<n). The general solution has been defined[96]

$$x = A^{+}y + (I_n - A^{+}A)z \tag{149}$$

where A^{+} is the generalised inverse matrix of order nxm and z is an arbitrary n-column vector. The problem is associated with determining an approximate or minimum-norm least-squares solution of eqn 148 in the sense of minimising $\|Ax - y\|^2$ such that $A^{+}A$ approximates the unit matrix I_n, or AA^{+} approximates the unit matrix I_m as closely as possible. In most practical cases A will be of maximal or full-column rank n < m and the least-squares solution will be given by the form of eqn 7, Part 1. The vector of residuals in the least-squares solution of eqn 148 is[92]

$$y - Ax = (I_m - AA^{+})y \tag{150}$$

Matrix $(I_m - AA^{+})$ is symmetric and idempotent and the sum of the squared residuals is a minimum given by

$$J = y'(I_m - AA^{+})y \tag{151}$$

similar to the form of the error covariance matrix of eqn 143, Part 1. The generalised inverse also possesses the properties -[94,95,99]

$$A^{++} = A, \quad AA^{+}A = A, \quad A^{*+} = A^{+*}, \quad (A^{*}A)^{+} = A^{+}A^{*+}$$

where A^{*} denotes the conjugate transpose of the matrix A. Matrices A, $A^{*}A$, A^{+}, $A^{+}A$ all have rank equal to trace $A^{+}A$.

Various methods developed for computation of the generalised inverse are reviewed.

Method 1 Penrose[94], Ben Israel and Charnes[99] and Aoki[11] consider a polar decomposition of the matrix A with rank r which has been credited to Gibbs (1931)[99]. The matrix A^{+} is defined by the spectral form[11]

$$A^{+} = \sum_{i=1}^{r} \lambda_i^{-1} f_i g_i^{*}, \quad \lambda_i = \rho_i^{\frac{1}{2}} \ (> 0) \tag{152}$$

where ρ_i are r nonzero eigenvalues of the matrix A^*A or AA^* and g_i, f_i are n- and m-column orthonormal eigenvectors associated with the matrices A^*A and AA^* respectively. The spectral form of solution may also be defined directly in terms of the Sylvester expansion theorem applied to the nxn matrix A^*A. Thus

$$(A^*A)^+ = \Sigma \rho_i^+ G(\rho_i) \tag{153}$$

$$A^+A = \Sigma G(\rho_i) \tag{154}$$

with the projection or idempotent matrices $G(\rho_i)$ defined by the form of eqn 298 and $\rho_i^+ = \rho_i^{-1}$ $(\rho_i > 0)$ or 0 if $\rho_i = 0$. Then

$$A^+ = (A^*A)^+A^* \tag{155}$$

The spectral methods require calculation of eigenvalues and eigenvectors and may be sensitive to errors in the computation of $\rho(A^*A)$[99].

Method 2 Penrose[95] considers a method based on partitioning which permits the generalised inverse of any matrix to be expressed in terms of the regular inverses of the partitioned submatrices. Thus any matrix can be partitioned in the form

$$D = \begin{bmatrix} A & B \\ C & CA^{-1}B \end{bmatrix} \tag{156}$$

where A is any nonsingular submatrix of rank r. Then

$$D^+ = \begin{bmatrix} A^*PA^* & A^*PC^* \\ B^*PA^* & B^*PC^* \end{bmatrix}, \quad P = (AA^*+BB^*)^{-1}A(A^*A+C^*C)^{-1} \tag{157}$$

The method has the disadvantage of requiring selection of the matrix A and the forming of the submatrices in eqn 156.

Method 3 Penrose[95] also illustrates an iterative method similar to Frame's method for computing the regular inverse of a matrix. Let $B = A^*A$ and define a sequence of matrices $C_{(j)}$, $j = 1,2..$ by

$$\begin{aligned} &C_{(1)} = I, \; C_{(j+1)} = I.\frac{1}{j}\,\text{trace}\,(C_{(j)}B) - C_{(j)}B \\ &C_{(r+1)}B = 0, \quad \text{trace}\,(C_{(r)}B) \neq 0 \end{aligned} \tag{158}$$

where r is rank of B. Then

$$A^+ = \frac{rC_{(r)}}{\text{trace}\,(C_{(r)}B)}A^* \tag{159}$$

Method 4 Greville[93] considers an mxn matrix A of rank $r > 0$ expressed as the product

$$A = BC \tag{160}$$

where B is of order mxr with its columns forming a basis for the column-space of A, and C is of order rxn with rows forming a basis for the row-space of A. Then

$$A^+ = C^+B^+ = C'(CC')^{-1}(B'B)^{-1}B' \tag{161}$$

Method 5 Boot[101] considers the generalised inverse as a problem of constrained maximisation of a quadratic function and develops results similar to those of Method 4. The matrix A is partitioned with the leading rxr matrix of full rank

$$A = \begin{bmatrix} D & E \\ F & G \end{bmatrix} \tag{162}$$

Matrix G is of order (m-r)x(n-r). Since A is of rank r

$$[FG] = H[DE] \quad \text{and} \quad \begin{bmatrix} E \\ G \end{bmatrix} = \begin{bmatrix} D \\ F \end{bmatrix} K \tag{163}$$

Then

$$A = \begin{bmatrix} D & E \\ HD & HE \end{bmatrix} = \begin{bmatrix} D & DK \\ HD & HDK \end{bmatrix} = RDQ' \tag{164}$$

where $$R = \begin{bmatrix} I \\ H \end{bmatrix}, \quad Q' = [I\ K], \quad H = FD^{-1}, \quad K = D^{-1}E \tag{165}$$

H and K represent the row and column dependencies. The unique generalised inverse is then given by

$$A^{+} = Q(Q'Q)^{-1}D^{-1}(R'R)^{-1}R' \tag{166}$$

The disadvantage of the method is the requirement to make the leading sub-matrix of full rank. Eqn 166 is then reduced to a form which avoids calculation of H and K

$$A^{+} = A_2' \, (A_2A_2')^{-1} \, D(A_1'A_1)^{-1}A_1' \tag{167}$$

where $A_2 = [DE]$, $A_1 = \begin{bmatrix} D \\ F \end{bmatrix}$

An improved computational method is obtained by considering the constrained minimisation problem - minimise trace X'X subject to the nxm equations $A_1'AX = A_1'$. Thus consider

$$\text{trace } X'X - \text{trace } M\,(A_1'AX - A_1') \tag{168}$$

where M is an mxr matrix of Lagrange multipliers. Differentiating with respect to X and M' then gives an equivalent solution to eqn 166

$$X = A^{+} = A'A_1(A_1'AA'A_1)^{-1}A_1 \tag{169}$$

requiring only one rxr matrix inversion and r independent columns of A. Similarly, minimising trace YY' subject to AA'Y' = A gives

$$Y = A^{+} = A_2'(A_2A'AA_2')^{-1}A_2A' \tag{170}$$

based on a set of r independent rows of A.

Method 6 Foster[102] defines an 'optimum' inverse which reduces to the generalised inverse of Penrose and Moore in a limiting case of error-free data. For a data vector related to signal (x) and noise (η),

$$y = Ax + \eta \tag{171}$$

the 'optimum' inverse is defined by the solution

$$\hat{x} = A^{I}y + (I - A^{I}A)\bar{x} \tag{172}$$

where

$$A^I = SA^*(ASA^* + N)^\dagger \tag{173}$$

with covariance matrices

$$S = (x-\bar{x})(x-\bar{x})' , \quad N = (\eta-\bar{\eta})(\eta-\bar{\eta})' \tag{174}$$

With $N = \sigma_N^2 I , \quad S = \sigma_s^2 I$

$$A^I = A^*(AA^* + \Gamma^2 I)^{-1} = (A^*A + \Gamma^2 I)^{-1} A^* , \quad \Gamma^2 = \frac{\sigma_N^2}{\sigma_s^2} \tag{175}$$

The inverse exists since AA^* is nonnegative definite and $\Gamma^2 I$ is positive definite. The presence of noise in the data improves the conditioning of the optimum inverse and corresponds to the technique suggested for inverting ill-conditioned positive-definite matrices by the addition of small multiples of the unit matrix. The solution of eqns 173 and 175 is similar to the form of eqn 15, Part 1.

Method 7 Greville[93] develops a recursive algorithm for obtaining the pseudoinverse $A_k^\dagger$ of a matrix A_k with k columns from $A_{k-1}^\dagger$, which corresponds to the successive addition of higher-order terms in the polynomial approximation problem. Consider the partitioned matrices

$$A_k = [A_{k-1} \quad a_k] , \quad A_k^\dagger = \begin{bmatrix} B_k \\ b_k \end{bmatrix} \tag{176}$$

where a_k denotes the kth column of A_k, then with relations

$$A_{k-1}^\dagger A_k A_k^\dagger = A_{k-1}^\dagger , \quad A_{k-1}^\dagger A_{k-1} B_k = B_k \tag{177}$$

the pseudoinverse is given by

$$A_k^\dagger = \begin{bmatrix} A_{k-1}^\dagger - d_k b_k \\ b_k \end{bmatrix} , \quad d_k = A_{k-1}^\dagger a_k \tag{178}$$

With $c_k = a_k - A_{k-1} d_k \neq 0$, it is shown that $b_k = c_k^\dagger = (c_k' c_k)^{-1} c_k'$

and with $c_k = 0$, $b_k = (1 + d_k' d_k)^{-1} d_k' A_{k-1}^\dagger$ (179)

The algorithm is associated with the form of the least-squares sequential estimation algorithm incorporating new data, and the correspondence with the Kalman-Bucy filtering equations for $c_k = 0$ corresponding to the observable case with $(A^*A)^{-1}$ existing, is discussed by Kishi[103]. Kishi formulates the same problem with

$A_k = \begin{bmatrix} A_{k-1} \\ a_k^* \end{bmatrix}$, $A_k^\dagger = [B_k \; b_k]$ from which the sequential least-squares solution is given by

$$\hat{x}_k = A_k^\dagger y_k = [A_{k-1}^\dagger - b_k a_k^* A_{k-1}^\dagger , \; b_k] \begin{bmatrix} y_{k-1} \\ y_k \end{bmatrix} \tag{180}$$

$$\text{or} \quad \hat{x}_k = \hat{x}_{k-1} - b_k a_k^* \hat{x}_{k-1} + b_k y_k \tag{181}$$

Method 8 Mayne[104] also develops a sequential algorithm for the pseudoinverse matrix based on regression analysis and Wiener-Kalman filtering theory. An observation scheme of m-parameters θ is used to define the pseudoinverse by the rth scalar observation

$$x_r = y_r'\theta + v_r = A^+ y_r \quad (182)$$

where y_r is a known m vector. Successive 'observations' are then used to update the 'estimate' $\hat{\theta}_r$ of the pseudoinverse using an algorithm of the form of eqn 26, Part 1, with $\hat{\theta}_r$, P_{r-1}, y_r, σ_{vv}, x_r corresponding to $\hat{x}_{k+1}$, P_k, H'_{k+1}, V^{-1}_{k+1}, y_{k+1} respectively, where σ_{vv} is the variance of the measurement noise v_r of zero mean and $\hat{\theta}(0) = 0$, $P_o = I$, $r = 1..p$. The mxm symmetric variance matrix

$$P_r = E(\theta - \hat{\theta}_r)(\theta - \hat{\theta}_r)' \quad (183)$$

is given similarly by the form of eqn 23, Part 1. The jth row of A^+ is then given by the m-dimensional vector $\hat{\theta}_j(p)$, $j = 1..n$.

Rao[97] considers a step-by-step reduction of a matrix A to a nonsingular generalised inverse using a method of sweepout and interchange of rows. Ben-Israel and Charnes give an explicit expression for A^+ as a limit, due to den Broeder and Charnes (1962),

$$A^+ = \sum_{k=1}^{\infty} A^*(I + AA^*)^{-k} \quad (184)$$

The polar decomposition method 1 and, particularly, the iterative and sequential methods 3, 7, 8 would appear to represent the preferred methods of solution compared to the other methods requiring selection of a nonsingular submatrix. However, further work must be undertaken in order to assess the computational merits of each method. The generalised inverse matrix is associated inherently with least squares methods of estimation, and the corresponding residual vector, particularly, is seen to possess properties analogous to those associated with all other least squares minimum norm solutions.

7. Optimal control of the linear dynamic system

Optimal online control of a dynamic process subjected to random disturbances on the basis of imperfect observations and ill-defined parameters represents an important and, in general, complex problem. If the state variables cannot be measured directly in a stochastic environment, as usually encountered in practice, the control sequence must be related to an estimate of the state variables using a performance criterion related to some probability measure of the states, such as minimum expected value of a squared error criterion. For the linear system subjected to additive white Gaussian random disturbances and with linear observation containing additive white Gaussian noise, the Separation Theorem[105-107,16,21,66,152] permits independent design of an optimal linear filter and a deterministic optimal controller for overall optimum performance.

The structure of the optimal control law is then defined by the deterministic problem assuming that all the state variables are available, and the control input will be given by a functional of the past observations by the form

$$u_k = K(\hat{x}_k) \qquad (185)$$

where $\hat{x}_k$ is the optimal estimate of the present state based on the available noisy measurements, and the performance index need not be quadratic[108]. Optimal control of the noisy linear dynamic system based on quadratic performance will be related to a linear function of the estimated states and data $\{y_o..y_{N-1}\}$ by

$$u_k = K\,\hat{x}_k \qquad (186)$$

and minimises the expected performance value EJ_N. The derivation of the optimal control law for the linear system, based on the discrete and continuous forms of dynamic programming and on the maximum principle, is illustrated in Sections 13.2 and 13.3. A similar discrete solution obtained using Lagrange multipliers is given by Lee[16]. The solution for discrete linear optimal control based on unconstrained quadratic minimisation is closely related to the solution obtained for sequential least-squares estimation, as first discussed by Kalman. The analogy is illustrated by a comparison between the single-stage quadratic-form algorithm of eqn 328 for M_{r-1} and the sequential relation for the estimation error covariance matrix of eqn 58, Part 1.

8. The matrix Riccati equation[110-117]

The problems of optimal control theory, linear filtering and prediction, identification and the model-in-the-performance index problem[115], associated with the minimisation of a quadratic functional lead to the existence of a matrix Riccati-type differential or difference equation. The theory of multiwire transmission lines also produces the same type of matrix differential equation as a matrix analogue of the impedance and admittance functions[113].

Optimal control of the linear dynamic system based on a finite-time quadratic performance functional is related to the time-varying solution of a matrix Riccati differential equation as outlined in Section 13.3. The resulting optimal trajectory may be defined in terms of a transition matrix associated with the matrix-M of eqn 343 with backward time $\tau = T-t$. Thus

$$\dot{z}(\tau) = -Mz(\tau)\ , \quad x(\tau_o) = x(T)\ , \quad p(\tau_o) = p(T) = 0 \qquad (187)$$

$$z(\tau) = \emptyset(\tau)\, z(\tau_o)\ , \quad \emptyset(\tau) = \begin{bmatrix} \emptyset_{11}(\tau) & \emptyset_{12}(\tau) \\ \emptyset_{21}(\tau) & \emptyset_{22}(\tau) \end{bmatrix} \qquad (188)$$

and with boundary condition $p(\tau_o) = 0$

$$p(\tau) = \emptyset_{21}(\tau)\emptyset_{11}(\tau)^{-1}\, x(\tau) = -P(\tau)\, x(\tau) \qquad (189)$$

Differentiating eqn 189 and combining with the components of eqn 187 then gives the matrix Riccati differential equation for reverse time

$$\dot{P}(\tau) = P(\tau)A + A'P(\tau) - P(\tau)BG^{-1}B'P(\tau) + Q, \quad P(\tau_o) = 0 \qquad (190)$$

and for optimal control

$$u(t) = -G^{-1}B'P(T-t)x(t) \qquad (191)$$

The time-varying matrix P(T-t) may be calculated using the transition matrix programming algorithms of Section 13.1, or alternatively by direct numerical integration of eqn 190.

The least-squares state estimation problem formulated in Section 3.1, Part 1, also leads to a continuous-time Riccati-type equation associated with the error covariance matrix. Considering first-order approximations, with small sampling interval h

$$\emptyset = I + Ah \quad , \qquad \Gamma = B_1 h \qquad (192)$$

and with arguments k+1, k → t+h, t, and[117,118]

$$Q(t) = Q_k h \quad , \qquad R(t) = V_{k+1}^{-1} h \qquad (193)$$

then limiting conditions applied in eqns 58 and 59 give

$$\bar{P}(t+h) = (I+Ah)P(t)(I+Ah)' + (B_1 h)Q(t)/h(B_1 h)' \qquad (194)$$

and

$$P(t+h) = \bar{P}(t+h) - \bar{P}(t+h)H'[Rh^{-1} + H\bar{P}(t+h)H']^{-1}H\bar{P}(t+h) \qquad (195)$$

$$= P(t) + hAP(t) + hP(t)A' + hB_1Q(t)B_1'$$
$$- h\left\{P(t)H'[R + hH\bar{P}(t+h)H']^{-1}HP(t) + O(h)\right\} \qquad (196)$$

Then in limit h→0, [P(t+h) - P(t)]/h→$\dot{P}$, and

$$\dot{P} = AP + PA' + B_1Q(t)B_1' - PH'R(t)^{-1}HP \;, \quad P(t_o) = 0 \qquad (197)$$

representing a continuous-time nonlinear first-order matrix Riccati differential equation associated with the dual problem of optimal filtering. Solutions based on eigenvector and transition matrix components apply similarly in this case using the augmented system matrix

$$M = \begin{bmatrix} -A' & H'R^{-1}H \\ B_1QB_1' & A \end{bmatrix} \qquad (198)$$

with $P = U_{21}U_{11}^{-1}$. Solution for the optimal error covariance matrix representing the propagation of uncertainty in the continuous dynamic system under the influence of forcing by B_1QB_1' with a linear additive noise measurement system may be obtained by direct integration of eqn 197, or by computation of the sequential algorithm of eqn 58 for the discrete system. The existence of the Riccati equation illustrates the close relationship between the optimal control problem and the sequential least-squares filtering problem. Kalman and Bucy[5] develop an analytical solution of the variance equation using the form of eqn 198 and also propose a solution related to transition matrix components.

8.1 Solution of the steady-state matrix Riccati differential equation

The steady-state converging solution of the matrix Riccati differential equation is associated inherently with the infinite-time regulator problem, and with the steady-state filtering problem associated with a stationary random process. The solution of the steady-state algebraic matrix Riccati equation may be obtained by direct integration, and iterative methods of solution of the resulting simultaneous nonlinear equations have also been developed[119,66]. Simplified computational techniques have also been obtained using partitioned eigenvector and transition matrix solutions related to the set of 2n simultaneous linear differential equations in the n-state variables and the n-adjoint variables[6]. The techniques are also suitable for solution of the Lyapunov stability matrix equation which is associated inherently with the form of the algebraic matrix Riccati equation.

The existence of asymptotic stability conditions in the linear optimal control problem, with the optimisation interval extending to infinite time, enables a simple analytical control design to be obtained that is both linear and time invariant. The optimal control law can then be related to limiting values of the partitioned components of the time-varying transition matrix, and also to the eigenvector components associated with the stable modes of the augmented system of 2n differential equations defining the optimal trajectory.

Thus using eqn 189 for a steady-state solution of eqn 190 with $\dot{P}(\tau) = 0$,

$$P = -\lim_{\tau \to \infty} \phi_{21}(\tau)\, \phi_{11}(\tau)^{-1} \tag{199}$$

The transition matrix components can now be defined using matrix series expansions. Thus

$$\phi(t) = \sum_{j=o}^{\infty} M^j t^j / j! \tag{199a}$$

where $$M = \begin{bmatrix} -A & \bar{G} \\ -Q & A' \end{bmatrix}, \quad \bar{G} = -BG^{-1}B' \tag{199b}$$

and with repeated squaring

$$M^j = \begin{bmatrix} \{(M^{j-1})_{11}(-A)+(M^{j-1})_{12}(-Q)\} & \{(M^{j-1})_{11}\bar{G}+(M^{j-1})_{12}A'\} \\ \{(M^{j-1})_{21}(-A)+(M^{j-1})_{22}(-Q)\} & \{(M^{j-1})_{21}\bar{G}+(M^{j-1})_{22}A'\} \end{bmatrix} = \begin{bmatrix} (M^j)_{11} & (M^j)_{12} \\ (M^j)_{21} & (M^j)_{22} \end{bmatrix} \tag{199c}$$

Then from eqn 199a the transition matrix components will be given by

$$\phi_{11}(t) = I_n + (M^1)_{11}t + (M^2)_{11}t^2/2! + .. = \sum_{j=o}^{\infty} (M^j)_{11} t^j/j! \tag{199d}$$

where $$(M^1)_{11} = -A, \quad (M^1)_{12} = \bar{G} \tag{199e}$$

and $(M^j)_{11}$, $(M^j)_{12}$ are given by the components of eqn 199c. Then

$$\phi_{11}(t) = I_n - At + \sum_{j=1}^{\infty} \left\{[(M^j)_{11}(-A) + (M^j)_{12}(-Q)]t^{j+1}/(j+1)!\right\} \quad (199f)$$

Similarly

$$\phi_{21}(t) = \sum_{j=1}^{\infty} (M^j)_{21} t^j/j! \quad , \quad (M^1)_{21} = -Q \; , \quad (M^1)_{22} = A' \quad (199g)$$

and including the components of eqn 199c

$$\phi_{21}(t) = -Qt + \sum_{j=1}^{\infty} \left\{[(M^j)_{21}(-A) + (M^j)_{22}(-Q)]t^{j+1}/(j+1)!\right\} \quad (199h)$$

Also

$$\phi_{12}(t) = \bar{G}t + \sum_{j=1}^{\infty} \left\{[(M^j)_{11}\bar{G} + (M^j)_{12}A']t^{j+1}/(j+1)!\right\} \quad (199i)$$

$$\phi_{22}(t) = I + A't + \sum_{j=1}^{\infty} \left\{[(M^j)_{21}\bar{G} + (M^j)_{22}A']t^{j+1}/(j+1)!\right\} \quad (199j)$$

The limiting solution of eqn 199 can now be obtained by repeated squaring of $\phi(t)$ which will give the translated transition matrix components

$$\begin{aligned} \phi_{11}(nt) &= \phi_{11}^2(nt/2) + \phi_{12}(nt/2)\phi_{21}(nt/2) \\ \phi_{21}(nt) &= \phi_{21}(nt/2)\phi_{11}(nt/2) + \phi_{22}(nt/2)\phi_{21}(nt/2) \end{aligned} \quad (199k)$$

and similarly for $\phi_{12}(nt)$, $\phi_{22}(nt)$. Then the symmetrical P matrix will be given by the form of eqn 199 with $n \rightarrow \infty$ for convergence. The components ϕ_{11}, ϕ_{21} will increase in magnitude with increasing argument and cancellation is required in order to ensure the existence of convergence for the steady-state solution P.

Potter[112] derives a steady-state solution of the nth-order matrix Riccati differential equation in terms of the eigenvectors of an associated 2nx2n matrix, which leads to an explicit solution of the optimal control problem for the linear system with quadratic performance based on the maximum principle. A simple derivation of this result is now illustrated by eliminating the effects of the unstable modes for asymptotic stability of the linear system. The time response of eqn 343 of Section 13.3.1 may be defined in terms of assumed distinct eigenvalues and eigenvector components for the matrix M. Thus

$$x(t) = Ue^{\Lambda T}U^{-1} z(t_o) \; , \qquad T = t - t_o \quad (200)$$

which is associated with the eigenvector matrix equation

$$MU = U\Lambda \quad (201)$$

where U is a 2nx2n modal matrix of eigenvector columns and Λ is a diagonal matrix with elements $\lambda_1 .. \lambda_{2n}$. Matrix M possesses convergent and divergent mode pairs with $\Lambda_1 = [\lambda_i]$, i = 1..n with negative real parts, and $\Lambda_2 = [\lambda_i]$, i = n+1..2n with positive real parts. Partitioning the solution of eqn 200

$$\begin{bmatrix} x(t) \\ p(t) \end{bmatrix} = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} e^{\Lambda_1 T} & 0 \\ 0 & e^{\Lambda_2 T} \end{bmatrix} \begin{bmatrix} U'_{22} & -U'_{12} \\ -U'_{21} & U'_{11} \end{bmatrix} \begin{bmatrix} x(t_o) \\ p(t_o) \end{bmatrix} \quad (202)$$

gives the state-variable solution

$$x(t) = U_{11}e^{\wedge_1 T}\left[U'_{22}x(t_o) - U'_{12}p(t_o)\right] + U_{12}e^{\wedge_2 T}\left[U'_{11}p(t_o)-U'_{21}x(t_o)\right] \qquad (203)$$

Then for conditions of asymptotic stability

$$p(t_o) = U'^{-1}_{11} U'_{21} x(t_o) \qquad (204)$$

$$x(t) = U_{11}e^{\wedge_1 T} (U'_{22} - U'_{12}U'^{-1}_{11}U'_{21})x(t_o) = U_{11}e^{\wedge_1 T}U^{-1}_{11}x(t_o) \qquad (205)$$

Similarly

$$p(t) = U_{21}e^{\wedge_1 T}U^{-1}_{11}x(t_o) = U_{21}U^{-1}_{11} x(t) = -Px(t) \qquad (206)$$

Thus for optimal control with asymptotic stability

$$u(t) = G^{-1}B'U_{21}U^{-1}_{11}x(t) \qquad (207)$$

giving an explicit solution based on an eigenvalue analysis. Computation of the 2n eigenvalues of the matrix M is required together with the n eigenvector components associated with the n stable mode eigenvalues. The asymptotic stability condition can be shown to be associated with the algebraic matrix Riccati equation by considering the eigenvalue problem

$$\begin{bmatrix} A & BG^{-1}B' \\ Q & -A' \end{bmatrix} \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix} = \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix} [\wedge_1] \qquad (208)$$

Expanding and eliminating $\wedge_1$ with $P = -U_{21}U^{-1}_{11}$ then gives

$$PBG^{-1}B'P - PA - A'P - Q = 0 \qquad (209)$$

where P is a symmetric positive semidefinite solution for all positive semidefinite matrices Q.

The technique of constraining the equations defining optimal motion for asymptotic stability is similar in principle to the methods developed in Section 9 for reducing the order of a set of matrix differential equations by neglecting the insignificant high-order modes[121].

The solution of the steady-state Riccati equation also contains inherently the solution of the Lyapunov equation associated with the stability of the free linear system, which can be obtained in terms of eigenvector and transition matrix components. Thus the Lyapunov equation

$$PA + A'P = -Q \qquad (210)$$

where the symmetric matrix P is required to be positive definite for any symmetric positive definite matrix Q is related to the eigenvalue problem defined by

$$\begin{bmatrix} A & 0 \\ Q & -A' \end{bmatrix} \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix} = \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix} [\wedge_1] \qquad (211)$$

with $P = -U_{21}U^{-1}_{11}$. Explicit algebraic methods of solution have been developed which require the forming of $n(n+1)/2$ sets of simultaneous equations for determining the unknown elements of the matrix P. A solution can now also be obtained based on the n eigenvalues of Λ_1 associated with the matrix A and the corresponding eigenvector components $U_{11} U_{21}$ for the augmented 2nx2n system. However, such a solution presupposes knowledge of the system eigenvalues and thus stability, but may have application in performance function weighting. By specific choice of the matrix Q[146] the solution can be related to the eigenvalues and eigenvectors of the matrix A only, which will avoid the computation of the eigenvector components U_{21} for the augmented system. Thus expanding eqn 211 with

$$Q = (U_{11}U'_{11})^{-1} \tag{212}$$

gives
$$U_{21} = \tfrac{1}{2}(\Lambda_1 U'_{11})^{-1} , \quad P = -\tfrac{1}{2}(U_{11}\Lambda_1 U'_{11})^{-1} \tag{213}$$

and with a symmetric A matrix

$$P = -\tfrac{1}{2} A^{-1} , \qquad Q = I \tag{214}$$

A transition matrix solution of the Lyapunov equation may also be obtained in terms of the eigenvalue problem defined by

$$\begin{bmatrix} -A & 0 \\ -Q & A' \end{bmatrix} \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix} = \begin{bmatrix} U_{11} \\ U_{21} \end{bmatrix} [\Lambda_1] \tag{215}$$

which also reduces to the Lyapunov equation with $P = -U_{21}U_{11}^{-1}$.
Now consider the transition matrix solution

$$\Phi(t) = \exp\left\{ \begin{bmatrix} -A & 0 \\ -Q & A' \end{bmatrix} t \right\} = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} e^{-\Lambda_1 t} & 0 \\ 0 & e^{\Lambda_1 t} \end{bmatrix} \begin{bmatrix} U'_{22} & -U'_{12} \\ -U'_{21} & U'_{11} \end{bmatrix} \tag{216}$$

then with $t \to \infty$, $e^{\Lambda_1 t} \to 0$ and

$$\underset{t\to\infty}{\mathrm{Lt}}\ \Phi(t) = \underset{t\to\infty}{\mathrm{Lt}} \begin{bmatrix} \phi_{11}(t)\phi_{12}(t) \\ \phi_{21}(t)\phi_{22}(t) \end{bmatrix} = \begin{bmatrix} U_{11}e^{-\Lambda_1 t}U'_{22}, & -U_{11}e^{-\Lambda_1 t}U'_{12} \\ U_{21}e^{-\Lambda_1 t}U'_{22}, & -U_{21}e^{-\Lambda_1 t}U'_{12} \end{bmatrix} \tag{217}$$

and

$$\underset{t\to\infty}{\mathrm{Lt}} \ [\phi_{21}(t)\, \phi_{11}(t)^{-1}] = U_{21}U_{11}^{-1} = -P \tag{218}$$

thus giving a solution of the Lyapunov equation in terms of transition matrix components, provided the limit converges.

The solution of the Lyapunov equation will also exist as a limiting solution of the dynamic programming algorithm of eqn 327 with $\Delta = 0$, thus

$$P_r = \Phi' P_{r-1} \Phi + Q , \quad r \to \infty \tag{219}$$

A solution can also be obtained by expanding the transition matrix components for $\phi_{11}(t)$, $\phi_{21}(t)$ in terms of the augmented matrix of eqn 215, giving

$$\phi_{21}(t) = (-Q)t + [(-Q)(-A)-A'Q]t^2/2! + [(QA-A'Q)(-A)-A'^2Q]t^3/3! + \ldots$$
$$+ [(\text{previous term})(-A)-A'^{k-1}Q]t^k/k! + \ldots \qquad (220)$$

$$\phi_{11}(t) = e^{-At} \qquad (221)$$

Repeated squaring of $\phi(t)$ will then give the translated components required in the solution of eqn 218[6]. However, eqn 219 based on the transition matrix $\phi(A)$ would appear to offer the simplest computational algorithm for solution of the discrete equation. Barnett, et al[163], also consider a convergent infinite matrix series numerical solution of the Lyapunov equation.

The techniques developed for solution of the algebraic matrix Riccati equation will have direct application to the steady-state sequential filtering algorithms using the augmented matrix of eqn 198. The numerical problems involved in solving the Riccati equation, particularly with regard to round-off errors are discussed in ref. 66.

9. Reduction of dynamical model complexity[121-130]

The lumped-parameter process model, such as may be developed for a distillation column or a power station boiler-turboalternator unit, will generally be associated with a relatively large number of state variables and controlled inputs. The responses of such processes, however, will usually be of simple form, representing the effects of a small number of significant time constants or modes. A problem may thus exist for reducing the order of the system model to include only the essential dynamics which will produce a transient response close to that of the original system. The various methods developed for simplifying lumped-parameter process models based on mode reduction, corresponding to the filtering of high frequency components by truncating the spectrum, and the concepts of linear least-squares estimation are outlined.

The high-order modes or small time-constant effects associated with the transient response of a linear system may be eliminated by isolating the modes using the modal-matrix transformation[121]

$$x = U\,q \qquad (222)$$

which transforms the system of eqn 280, possessing distinct eigenvalues λ_i, to the normal coordinate form

$$\dot{q}(t) = \Lambda q(t) + U^{-1}B\,u(t) \qquad (223)$$

The system states and defining matrices are now partitioned

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad q = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}, \quad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$
$$\Lambda = \begin{bmatrix} \Lambda_1 & \cdot \\ \cdot & \Lambda_2 \end{bmatrix}, \quad U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}, \quad U^{-1} = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} = \begin{bmatrix} V_A \\ V_B \end{bmatrix} \qquad (224)$$

with the eigenvalues arranged in order of increasing moduli, $\lambda_1 < \lambda_2 < .. < \lambda_n$. With a reduction to order r, U_{11} is an rxr and U_{21} an (n-r)xr matrix of modal vectors associated with the first r significant modes and the required state variables. Neglecting the transient solutions $e^{\lambda_i t}$ for the higher modes, with $q_k \to 0$, k = (n-r)..n in the partitioned form of eqn 222, gives

$$x_2 = U_{21} U_{11}^{-1} x_1 \qquad (225)$$

Partitioning the free system of eqn 280 and combining with eqn 225 then gives the reduced-order system matrix

$$\bar{A} = A_{11} + A_{12} U_{21} U_{11}^{-1} \qquad (226)$$

Eqn 225 illustrates the similarity between the method of mode reduction for simplifying system response and the technique of rejecting unstable modes for asymptotic stability of the optimal controlled linear system.

The driving-matrix coefficients for the reduced system are obtained by considering the response due to forcing

$$x_f = \int_0^t U e^{\Lambda(t-\tau)} U^{-1} B\, u(\tau) d\tau \qquad (227)$$

and for the reduced system

$$x_f = \int_0^t U_{11} e^{\Lambda_1 (t-\tau)} U_{11}^{-1} \bar{B}\, u(\tau) d\tau \qquad (228)$$

Neglecting modes in the solution of eqn 227 and equating partitioned responses then gives the rxp-dimensional driving matrix

$$\bar{B} = U_{11} V_A B \qquad (229)$$

The reduced rth-order system associated with the significant eigenvalues is then represented

$$\dot{\bar{x}}_1(t) = \bar{A}\, \bar{x}_1(t) + \bar{B}\, u(t) \qquad (230)$$

with the remaining variables x_2 given by the algebraic relation of eqn 225. It is important to select dependent variables for the reduced system which will avoid ill-conditioned matrices, such as might be obtained with variables of similar character, such as steam and metal temperatures in a boiler model. With such conditions it may then be appropriate to consider the application of the generalised inverse matrix.

The modal method of reduction is valid for real and complex eigenvalues of the matrix A, and also for repeated eigenvalues with non-degenerate eigenvectors[125]. It is also valid for repeated eigenvalues with degenerate eigenvectors if the eigenvectors are related to the Jordan canonical form of A. The method, however, introduces error in the steady-state values of the reduced model compared to the original system. These may be compensated by combining the dynamics of the reduced system with the new state variable[125]

$$\hat{x}_1 = \bar{x}_1 + \left\{ \bar{A}^{-1}\bar{B} - [\hat{A}_{11} \hat{A}_{12}] B \right\} u \qquad (231)$$

with the inverse of the original system matrix partitioned in terms of the retained variables

$$A^{-1} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \quad (232)$$

Witn a singular system A-matrix the steady-state values may be obtained by suitable partitioning.

A similar method of reduction has been considered by assuming derivatives $\dot{q}_k \rightarrow 0$[122]. This produces a reduced system which retains the original steady-state values but may introduce errors into the form of the transient response. Combining the partitioned forms of eqns 222 and 223 gives

$$x_2 = U_{21}U_{11}^{-1} x_1 - V_{22}^{-1}\Lambda_2^{-1} (V_{21}B_1 + V_{22}B_2)u \quad (233)$$

and substituting into the original partitioned system equations gives the reduced-order form

$$\dot{x}_1 = (A_{11} + A_{12}U_{21}U_{11}^{-1})\bar{x}_1 + [B_1 - A_{12}V_{22}^{-1}\Lambda_2^{-1}(V_{21}B_1 + V_{22}B_2)]u \quad (234)$$

A method of reduction based on the geometric projection of a known state vector onto a linear subspace associated with a reduced number of states of a discrete linear system represented by transition equations, has also been considered[126]. The method is based on a non-sequential least-squares fitting of r known response components of the original vector to the discrete response of the reduced system

$$x_i = \hat{\phi}(T)x_{i-1} + \hat{\Delta}(T)\, u_{i-1} \quad (235)$$

where $\hat{\phi}$ and $\hat{\Delta}$ are the transition and driving matrices of reduced orders rxr and rxp respectively. Including r states and p control components at each interval i eqn 235 is then reformed to

$$[x_i] = [x_{i-1} \mid u_{j-1}][\hat{\phi}\hat{\Delta}]' \;, \quad i = 1..k+1 \;(\geq r+p+2) \quad (236)$$

$$\text{or} \quad X_{k+1} = M_j\, C \quad (237)$$

of orders $(k+1)xr$, $(k+1)x(r+p)$, $(r+p)xr$

Then for a least-squares solution of the transition and driving matrices

$$C = (M_j' M_j)^{-1} M_j' X_{k+1} \quad (238)$$

The method represents a particular application of the results for least-squares parameter estimation of Section 4.3 with offline fitting based on the use of block data obtained over an extended response period. In online application the reduced-order model identification algorithm could be developed for sequential solution using the methods of Section 4.3. The fitting of data to the reduced-order model could then be extended beyond the range of the available data X_{k+1} to give improved identification using the additional information available with changing control and disturbance conditions.

Mitra[128,129] considers an optimal method of model reduction related to the best reduced-order subspaces spanned by the set of eigenvectors corresponding to the r largest eigenvalues of the outer product matrix of vector impulse responses of the linear system defined over an extended response period. The best subspace is thus specified in terms of the eigenstructure of the matrix of impulse functions

$$W = \int_0^T \left\{ r(t)\, r(t)' \right\} dt$$

with the minimum value of the projection error given by the sum of the remaining n-r eigenvalues

$$\sum_{i=n-r+1}^{n} \lambda_i(W)$$

For the linear system with input

$$u(t) = \sqrt{D}\delta(t) \qquad (239)$$

then

$$r(t) = e^{At} B\sqrt{D} \qquad (240)$$

and

$$W(T) = \int_0^T e^{At}\, BD\, B' e^{At'} dt \qquad (241)$$

It is shown that matrix W satisfies the matrix Riccati equation

$$\dot{W} = AW + WA' + BDB' \qquad (242)$$

With asymptotic stability the resulting algebraic matrix equation can be solved using the results of Section 8.1. The methods involve properties of the covariance matrix of impulse responses which are similar in concept to those associated with the spectral analysis of the covariance data matrix of Section 5. They introduce similar properties with the value of the performance function related to summated eigenvalues. The methods imply a rejection of the smallest eigenvalues of a covariance matrix by projection onto a linear subspace and are analogous to the methods of modal reduction which define the best subspace for the reduced model by neglecting the largest eigenvalues or short time-constant modes associated with the system A-matrix.

Brown[127] considers a procedure based on a minimisation of the difference between the time derivatives of the states of the reduced-order model and the original system model, specified by the conditional ensemble expectation

$$E\left\{(\dot{\bar{x}}_1 - \dot{x}_1)'(\dot{\bar{x}}_1 - \dot{x}_1)\right\} \qquad (243)$$

where

$$\dot{\bar{x}}_1 = \bar{A}_{11}(t)\, x_1 + \bar{B}_1(t)u \qquad (244)$$

and u is an assumed nonstationary random control input.
The variables $s = (x_1 u)'$ and x_2 are assumed jointly normal with given covariance matrix

$$E(sx_2)'(sx_2) = \begin{bmatrix} V_{11}(t) & V_{21}'(t) \\ V_{21}(t) & V_{22}(t) \end{bmatrix} \qquad (245)$$

The low-order model coefficients obtained by minimisation of eqn 243 are then given by

$$[\bar{A}_{11}(t),\ \bar{B}_1(t)] = [A_{11}, B_1] + A_{12}V_{21}(t)V_{11}^{-1}(t),\ E[x_2|\ s] = V_{21}(t)V_{11}^{-1}(t)s \tag{246}$$

The method produces a reduced model similar in form to that obtained by mode reduction but requires the computation of time-varying coefficients. Similar methods based on integrating the fit of derivatives of a function have been used successfully in numerical analysis[156].

Reduced models obtained by the above techniques may not, in general, be considered appropriate for use in an overall system design study, but they have usually been found to be sufficiently accurate for control design based on dynamic optimisation.

10. Linear least-squares sensitivity

The effects of the accuracy of measured data and of changes in the elements of the matrices H, $\emptyset$, $\triangle$, V, Q and G associated with the least-squares estimation, identification and control problems are of considerable theoretical and practical importance. The least-squares problem will often be affected by ill-conditioning, and round-off errors may be introduced by numerical integration and matrix inversion. Model parameters will be affected by uncertainty, and linearisation of non-linear dynamic equations will also introduce inaccuracies. In practice knowledge of the noise statistics will be limited and may not be known with certainty and will, inherently, contain inaccuracies. It is thus important to consider the effects of perturbations on the transient response, the controlled inputs and performance criterion and on the filtering algorithms for estimated states and error covariance matrices. The sensitivity of the properties of linear systems produced by parameter change thus represents a fundamental and important problem and provides a more detailed understanding of the behaviour of linear multivariable systems which may be used to form the basis for overall system design.

Boot[101] has considered the sensitivity of the convex quadratic programming problem and Neal[131] studied the sensitivity of the Kalman estimator to small changes in the transition matrix elements. Aoki[11] considers the effects of variation of gain and of changes in the transition matrix and also the effects of imprecise noise statistics, and also considers methods of simplification in a sensitivity and error analysis of the Kalman filter. Griffin and Sage[141] discuss both large and small scale sensitivity of the optimum estimation algorithm and consider the effect of model errors and of errors in the plant and measurement noise covariance matrices. Price[150] and Heffes[151] also consider the effect of modelling errors and of errors in the noise models on the filter estimates. Nishimura[153] studies the effect of errors in the a priori information on the variance of the estimates and determines an upper bound for the variance and its effect on the final estimation. The effects of differential changes in the elements of a matrix have been considered with respect to changes produced in the system eigenvalues and eigenvectors[132-140]. Such variations will also affect the components of the state-transition equation and will produce corresponding changes in the discrete state variables.

10.1 Sensitivity of least-squares estimation

The effect of differential changes in the elements of ϕ, Δ, Γ, H, V on the sequential state estimation algorithm for the dynamic system with control can be determined by perturbing eqns 56-59, or the solution given by the form of eqn 44. Including control related to the estimated states, eqns55, 57 and 59 will combine to give

$$\hat{x}_{k+1} = (\phi + P_{k+1}\bar{P}_{k+1}^{-1}\Delta K)\hat{x}_k + P_{k+1}H'_{k+1}V_{k+1}(y_{k+1} - H_{k+1}\phi\hat{x}_k) \quad (247)$$

then with perturbation the differential change in the estimated state will be given by

$$d\hat{x}_{k+1} = P_{k+1}\bar{P}_{k+1}^{-1}(\phi+\Delta K)d\hat{x}_k + [P_{k+1}\bar{P}_{k+1}^{-1}d\phi + d(P_{k+1}\bar{P}_{k+1}^{-1}\Delta K)]\hat{x}_k$$
$$+ P_{k+1}H'_{k+1}V_{k+1}[dy_{k+1} - (dH_{k+1})\phi\hat{x}_k] + d(P_{k+1}H'_{k+1}V_{k+1})(y_{k+1}-H_{k+1}\phi\hat{x}_k) \quad (248)$$

where

$$d(P_{k+1}\bar{P}_{k+1}^{-1}\Delta K) = (dP_{k+1})\bar{P}_{k+1}^{-1}\Delta K - P_{k+1}\bar{P}_{k+1}^{-1}(d\bar{P}_{k+1})\bar{P}_{k+1}^{-1}\Delta K + P_{k+1}\bar{P}_{k+1}^{-1}(d\Delta)K + P_{k+1}\bar{P}_{k+1}^{-1}\Delta(dK) \quad (249)$$

$$d(P_{k+1}H'_{k+1}V_{k+1}) = (dP_{k+1})H'_{k+1}V_{k+1} + P_{k+1}(dH'_{k+1})V_{k+1} + P_{k+1}H'_{k+1}(dV_{k+1}) \quad (250)$$

$$dP_{k+1} = P_{k+1}[\bar{P}_{k+1}^{-1}(d\bar{P}_{k+1})\bar{P}_{k+1}^{-1} - d(H'_{k+1}V_{k+1}H_{k+1})]P_{k+1} \quad (251)$$

$$d\bar{P}_{k+1} = d(\phi P_k\phi' + \Gamma Q_k\Gamma') \quad (252)$$

It will be noted that the perturbed equation for $d\hat{x}_{k+1}$ is of sequential form with the coefficient of the previous-stage change $d\hat{x}_k$ unaffected by perturbation. Eqn 251 gives the sensitivity of the actual covariance matrix related to error in the model process and in the weighting matrices. The equations will simplify with H_{k+1} and Γ existing as constant incidence matrices and with constant weighting matrix V_{k+1}, and the process would appear to be stable with terms of the form $P_{k+1}\bar{P}_{k+1}^{-1}$ producing effects of cancellation with small changes. Persistent component variations must not be permitted to affect the positive-definite character of certain matrices, and changes in off-diagonal matrix elements must be offset by appropriate changes in the diagonal elements[101]. Changes in the matrices ϕ, Δ and K will be associated with the differential changes introduced into the defining system A-matrix, and these will correspondingly affect both the state estimation problem and the time solution of the system model[139]. Similar techniques may also be applied to obtain the sensitivity of the identification algorithm of eqns 94-98 Part 1, and the effect of small variations on the minimum value of the performance index will also follow directly.

10.2 Sensitivity of the spectral prediction algorithm

The sensitivity of the prediction algorithm of Section 5.2, Part 1, based on spectral expansion will require knowledge of the sensitivity of the eigenvalue analysis of the defining covariance data matrix. Various techniques have been developed for relating eigenvalue change to differential matrix element changes, and the basic role of Sylvester's expansion theorem associated with the spectral decomposition of a matrix and of the constituent matrices in this development have been illustrated[139]. The inverse eigenvalue sensitivity problem concerned with the requirement to synthetise a differential change in the elements of a matrix to produce a desired eigenvalue change, with applications in control design, has also been considered by reformulating the eigenvalue sensitivity problem to obtain a direct relationship between the matrix elements and the corresponding eigenvalue changes.

A differential change of dA in the matrix A possessing distinct eigenvalues has been shown to produce the small first-order eigenvalue variation

$$d\lambda_r = v_r dA u_r \qquad (253)$$

where the eigenvector and eigenrow u_r and v_r respectively associated with the eigenvalue λ_r are defined by eqns 290-292. The eigenvalue sensitivity has also been represented as the sum of inner products formed by the rows and columns of two nxn matrices[133]

$$d\lambda_r = u_r v_r * dA = \text{trace}\,(u_r v_r dA) \qquad (254)$$

The n-square idempotent matrices $u_r v_r$ represent the constituent matrices of Sylvester's expansion theorem which expresses the polynomial function of a matrix, with n distinct eigenvalues, in the form of eqn 298. Then

$$G(\lambda_i) = u_i v_i = \frac{Q(\lambda_i)}{g'(\lambda_i)} \qquad (255)$$

where $Q(\lambda_i)$ is the adjoint of the characteristic matrix $(\lambda I-A)$ and $g'(\lambda)$ represents the derivative of the characteristic determinant $|\lambda I-A|$ with respect to λ. Laughton[134] and Crossley and Porter[140] define $G'(\lambda_i)$ as the matrix of eigenvalue sensitivity coefficients or condition numbers with respect to the elements of A, for all eigenvalues, by the form

$$G'(\lambda_i) = [\partial\lambda_i/\partial a_{jk}] \qquad (256)$$

Now $$g'(\lambda) = \text{trace}\; Q(\lambda) \qquad (257)$$

thus $$d\lambda_r = \frac{\text{trace}[Q(\lambda_r)dA]}{\text{trace}\; Q(\lambda_r)} \qquad (258)$$

The eigenvector sensitivity problem has also been studied.[134,137,140] Reddy[137] considers a method based on the properties of the derivative of a determinant associated with the adjoint matrix $Q(\lambda)$ of the characteristic matrix of A. The method can be explained simply by considering the characteristic matrix

$$M_i = \lambda_i I - A \qquad (259)$$

The corresponding adjoint matrix or transposed matrix of cofactors, with columns proportional to the eigenvector u_r and rows proportional to the eigenrow v_r [139], is then defined by

$$Q(\lambda_i) = \begin{bmatrix} C_{11}(M_i) & \cdot\ \cdot & C_{j1}(M_i) & \cdot\ \cdot & C_{n1}(M_i) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ C_{1n}(M_i) & \cdot\ \cdot & C_{jn}(M_i) & \cdot\ \cdot & C_{nn}(M_i) \end{bmatrix} \tag{260}$$

where elements $C_{jk}(M_i)$ represent the cofactors of the matrix M_i. An eigenvector associated with the eigenvalue λ_i can now be defined in terms of any column of the matrix $Q(\lambda_i)$, and the sensitivity of the eigenvector to a parameter change will then be given by

$$du_i = d[C_{j1}(M_i) \ \cdot\ \cdot\ C_{jn}(M_i)]' \tag{261}$$

Now the derivative of a determinant with respect to a parameter is given by the sum of n determinants obtained by replacing individual rows (or columns) by their derivatives with respect to the parameter. Thus for the cofactor element $C_{j1}(M_i)$ given by an n-1-order determinant

$$dC_{j1}(M_i) = \begin{vmatrix} d(C_{j1})_{11} & \cdot\ \cdot & d(C_{j1})_{1,n-1} \\ (C_{j1})_{21} & \cdot\ \cdot & (C_{j1})_{2,n-1} \\ \cdot & \cdot & \cdot \\ (C_{j1})_{n-1,1} & \cdot & (C_{j1})_{n-1,n-1} \end{vmatrix} + \begin{vmatrix} (C_{j1})_{11} & \cdot\ \cdot & (C_{j1})_{1,n-1} \\ d(C_{j1})_{21} & \cdot\ \cdot & d(C_{j1})_{2,n-1} \\ \cdot & \cdot & \cdot \\ (C_{j1})_{n-1,1} & \cdot & (C_{j1})_{n-1,n-1} \end{vmatrix}$$

$$+ \cdot\ \cdot\ \cdot + \begin{vmatrix} (C_{j1})_{11} & \cdot\ \cdot & (C_{j1})_{1,n-1} \\ (C_{j1})_{21} & \cdot\ \cdot & (C_{j1})_{2,n-1} \\ \cdot & \cdot & \cdot \\ d(C_{j1})_{n-1,1} & \cdot\ \cdot & d(C_{j1})_{n-1,n-1} \end{vmatrix}$$

$$= \sum_{r=1}^{n-1} K_r \tag{262}$$

and expanding each determinant K_r in terms of cofactor elements $C(K_r)_{ij}$

$$dC_{j1}(M_i) = \sum_{k=1}^{n-1} d(C_{j1})_{1,k}\, C(K_1)_{1,k} + \cdot\ \cdot\ \cdot + \sum_{k=1}^{n-1} d(C_{j1})_{n-1,k}\, C(K_{n-1})_{n-1,k} \tag{263}$$

$$= \text{trace} \begin{bmatrix} C(K_1)_{11} & \cdot\ \cdot & C(K_1)_{1,n-1} \\ \cdot & \cdot & \cdot \\ C(K_{n-1})_{n-1,1} & \cdot\ \cdot & C(K_{n-1})_{n-1,n-1} \end{bmatrix} \begin{bmatrix} d(C_{j1})_{11} & \cdot\ \cdot & d(C_{j1})_{n-1,1} \\ \cdot & \cdot & \cdot \\ d(C_{j1})_{1,n-1} & \cdot\ \cdot & d(C_{j1})_{n-1,n-1} \end{bmatrix} \tag{264}$$

Other derivative eigenvector components $dC_{j2}(M_i)$. . in eqn 261 will be given similarly in terms of the cofactor components C_{j2}. . The result is similar to the form of eqn 18 in reference 139 which states that for any square matrix $M(\lambda)$ with elements m_{ij}, the derivative of the determinant is

$$\frac{d}{d\lambda}\left| M(\lambda) \right| = \sum_{i,j} [C(m_{ij})] \frac{dm_{ij}}{d\lambda} \qquad (265)$$

and for $\quad M(\lambda) = \lambda I - A$

$$\frac{d}{d\lambda}\left| \lambda I - A \right| = \text{trace } [Q(\lambda) \frac{d}{d\lambda}(\lambda I - A)] = g'(\lambda) \qquad (266)$$

Reddy[158] uses the eigenvalue and eigenvector sensitivity functions to obtain a closed-form solution for the sensitivity of the response of the linear system subjected to parameter variations.

The above results will now form the basis for the sensitivity of the prediction algorithm of Section 5 based on a spectral expansion. Thus small variations in the elements of the covariance matrix R will perturb the eigenvalues by $d\Lambda$ and the eigenvectors by $d\phi$ according to eqns 258 and 264. Then from eqn 114, Part 1 the coefficients C will change according to

$$dC = (dX)\phi + X(d\phi) \qquad (267)$$

The differential changes $d\phi$ will determine particularly the sensitivity of the sequential prediction algorithm of eqns 139 and 140, Part 1, to changes in the data X defining the characteristic functions, and in the data x_{M+1} occurring during the prediction interval. Perturbing eqn 137 then gives the differential change in the prediction coefficients

$$d[c_{M+1,i}]_j = d[c_{M+1,i}]_{j-1} P_{j-1}^{-1} P_j + \left\{ dx_{M+1}(j) - [c_{M+1,i}]_{j-1}(d\phi_i'(j))\right\}\phi_i(j)P_j$$

$$+ \left\{ x_{M+1}(j) - [c_{M+1,i}]_{j-1}\phi_i'(j)\right\}\left\{(d\phi_i(j))P_j + \phi_i(j)(dP_j)\right\} \qquad (268)$$

and from eqn 135,

$$dP_j = P_j[P_{j-1}^{-1}(dP_{j-1})P_{j-1}^{-1} - d(\phi_i'(j)\ \phi_i(j))]P_j \qquad (269)$$

Changes in the eigenvectors $\phi_i(j)$ will be based on changes in the original data using the form of eqns 261 and 264. Corresponding variations in the coefficients $[c_{M+1,i}]_j$ will then affect the predicted values $[\hat{x}_{M+1}(j+1)..\hat{x}_{M+1}(N)]$ according to the perturbed form of eqn 140. The sensitivity analysis may be used as a basis for improving the accuracy of prediction, and may have application particularly for compensating the errors encountered during the peak periods of electrical load prediction based on a spectral analysis[71].

10.3 Sensitivity of the linear optimal control problem

The sensitivity of the control law, performance index and state vector to parameter perturbation in the linear optimal control problem has been considered by Pagurek (1965)[142], Barnett (1966)[143] and Barnett and Storey (1966)[144]. Barnett and Storey develop conditions for an insensitive optimal control law and show that optimal control is more sensitive to small changes in the system B matrix than to small changes in the corresponding A matrix.

The sensitivity of the optimal control law of eqn 353 related to differential changes in the elements of the matrices G, B and P is given by

$$dK = G^{-1}(dG)G^{-1}B'P - G^{-1}(dB')P - G^{-1}B'dP \qquad (270)$$

With asymptotic stability, the differential change dP will be associated with the perturbed steady-state matrix Riccati equation

$$(dP)A + P(dA) + (dA')P + A'(dP) - (dP)BG^{-1}B'P - Pd(BG^{-1}B')P - PBG^{-1}B'dP + dQ = 0 \qquad (271)$$

Then for changes in matrices A and B with P_o representing the original P matrix

$$(dP)\bar{A} + \bar{A}'(dP) + (Q_1 + Q_2 + dQ) = 0 \qquad (272)$$

where

$$\bar{A} = A - BG^{-1}B'P_o = A + BK$$

$$Q_1 = P_o(dA) + (dA')P_o \qquad (273)$$

$$Q_2 = -P_o[(dB)G^{-1}B' + BG^{-1}(dB')]P_o$$

Eqn 272 defines the perturbed equation for the nxn symmetric sensitivity matrix dP which is governed by a modified Lyapunov-type equation. Matrices Q_1 and Q_2 are symmetrical forms related to the changes dA and dB respectively and must be constrained in order to retain the positive-definite character of the overall Q-matrix. A solution for dP can now be obtained using the techniques of Section 8.1 based on transition matrix and eigenvector components. The sensitivity of the matrix P of eqn 206 related to eigenvector components will be given by

$$dP = (PdU_{11} - dU_{21})U_{11}^{-1} \qquad (274)$$

with eigenvector changes dU_{11}, dU_{21} related to differential changes in the defining system matrix.[11] Similarly, the sensitivity dP may be related to perturbed transition matrix components in the limiting relation of eqn 199, which may then be related to A-matrix element changes[139]. The sensitivity of the reduced-order model based on modal reduction will also be related to the sensitivity of the corresponding eigenvalue problem and particularly to the differential changes produced in the components U_{11}, U_{21}.

The discrete optimal control algorithm of eqns 325, 327 and 328 will also perturb in a similar manner to the sequential algorithm for state estimation. Thus from eqn 325

$$dK_r = \hat{P}_{r-1}^{-1}[(d\hat{P}_{r-1})\hat{P}_{r-1}^{-1}\Delta'P_{r-1}\phi - (d\Delta')P_{r-1}\phi - \Delta'(dP_{r-1})\phi - \Delta'P_{r-1}(d\phi)] \qquad (275)$$

$$\hat{P}_{r-1} = \Delta'P_{r-1}\Delta + G \qquad (276)$$

and from eqns 327 and 328 the perturbed matrix dP_{r-1} reduces to

$$\begin{aligned} dP_{r-1} = {} & (d\phi')R_{r-2}P_{r-2}\phi + \phi'P_{r-2}R'_{r-2}(d\phi) + \phi'R_{r-2}(dP_{r-2})R'_{r-2}\phi \\ & -\phi'[P_{r-2}\Delta\hat{P}_{r-2}^{-1}(d\Delta')P_{r-2}R'_{r-2} + R_{r-2}P_{r-2}(d\Delta)\hat{P}_{r-2}^{-1}\Delta'P_{r-2} \\ & - P_{r-2}\Delta\hat{P}_{r-2}^{-1}(dG)\hat{P}_{r-2}^{-1}\Delta'P_{r-2}]\phi + dQ \end{aligned} \qquad (277)$$

$$\hat{P}_{r-2} = \Delta' P_{r-2} \Delta + G \qquad (278)$$

$$R_{r-2} = I - P_{r-2} \Delta \hat{P}_{r-2}^{-1} \Delta' \qquad (279)$$

The above relations will permit a study of the effects of all parameter changes and can be developed to consider, in particular, the sensitivity of the controlled response to variations in the control-law weighting parameters and in the system model equations, and also the conditions required to minimise the performance sensitivity. Chen and Shen[145] investigate the effect of variations in the controlled system A-matrix and the corresponding differential changes dP produced by the elements of dQ. The eigenvalue sensitivity equation is then used in an iterative algorithm for determining the weighting matrix Q and the optimal feedback gain matrix based on desired closed-loop eigenvalues.

Sensitivity of the estimation, identification and control algorithms associated with the linear dynamic system is particularly important for assessing the effects of inaccurate plant and noise model parameters and also the effects of the weighting coefficients on performance. In particular, the covariance functions associated with the state estimates and the noise sequences will not be known exactly in practice, and computational techniques for adaptive sequential estimation within the framework of a sensitivity analysis are required for estimating and updating these functions from available data.

11. Research applications[162]

A wide range of research activities relevant to the general field of linear system theory has been discussed and referenced. Research work is now being undertaken in the Department of Control Engineering, University of Sheffield which will also include online application of the techniques of state estimation, parameter identification, prediction and control of processes which require detailed investigation under near-actual operating conditions. The process computer is an extremely versatile machine and it is believed to possess enormous potential for extending the present areas of application to include more advanced adaptive optimal control techniques using detailed and changing plant models related to the steady-state and dynamic characteristics of the plant. The research work will thus be concerned essentially with extending the application of the process computer for online control using techniques based on modern control and linear system theory for obtaining integrated control of large-scale systems in the power, steel and glass industries.

Optimal online scheduling of a multimachine power system is being investigated on a multilevel basis. A process computer will operate as a grid process controller for the real-time automatic scheduling and control of a power system simulated in a large central scientific computer. The computers will be interconnected through a data link, and the study will incorporate techniques of linear programming for optimal scheduling and methods of network solution based on partitioning and updating using sequential algorithms, together with load prediction based on spectral analysis. The application of pattern recognition techniques for line security assessment will also be investigated. Process modelling of a glass-tube manufacturing process, marine boiler, multi-stand cold rolling mill, electric arc furnace, steel bar mill and billet furnace is also being undertaken. These studies, and other work concerned with factory production control, road tunnel ventilation and traffic control, online identification of a boiler model, and the analysis of medical data will also consider the application of many of the sequential computational algorithms developed in the papers.

12. Conclusions

An attempt has been made to review the development of the closely related sequential algorithms for state estimation, parameter identification, prediction and control of the noisy linear dynamic system using classical least-squares theory. The techniques have particular application in the fields of automatic control and stochastic approximation, and also link together many other areas of study involving online data processing for the fitting of model parameters and decision making in stochastic environments. The linear least squares theory which has developed naturally from the original work of Gauss, and Aitken and Plackett among others, and Wiener and Kolmogorov[154] also provides a basis for the study of the more difficult problems of nonlinear stochastic estimation and control. Computational algorithms have been developed and illustrated in a readily accessible form suitable for direct application, which it is hoped will promote the understanding and motivate the use of these powerful techniques of data processing in many other fields.

There is an increasing and active interest in the use of sequential algorithms in problems concerned with nonstationary processes, and particularly in problems of pattern recognition and data classification. In the power system control problem there is an application for obtaining online sequential solutions for load flow and optimal load scheduling. There is undoubtedly many fields remaining to be explored, particularly those concerned with process system identification and control and with the modelling of biological systems. The techniques of pattern recognition and machine classification incorporating sequential processing using the methods of potential functions and spectral analysis will also find application in these fields. Collaboration between different scientific disciplines in the universities and in industry is now essential in order to obtain increased effort for investigating applications of the developments which have been achieved in the theory of state estimation, system identification, prediction, reduction and control, and for reducing the gulf now existing between theory and practice. Present-day knowledge of the underlying theoretical basis for optimal control and stochastic approximation is at an advanced stage, and wider experience in practical application is now required in order to demonstrate the effectiveness of the techniques which can be applied for the solution of an increasing range of scientific problems.

13. Appendix

13.1 The discrete time solution of linear state equations

The complete time solution of the linear continuous-time dynamic system

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{280}$$

is defined by

$$x(t) = e^{At}x_o + \int_o^t e^{A(t-\tau)}B\, u(\tau)d\tau \tag{281}$$

representing solutions of the homogeneous equation and convolution integral, similar to the form for scalar variables. Tabulated integral solutions are available[159] for impulse, step and ramp-type forcing functions, which can be used to approximate continuous inputs. A discrete time solution, with the forcing functions changing only at the sampling time iT as step inputs, is represented by the recursive vector-difference eqn 1, Part 1, where $\phi(i+1,i)$ or $\phi(T)$ represents the nxn transition matrix defined for a sampling period T by the relation

$$\phi(T) = e^{AT} = \sum_{j=0}^{\infty} A^j T^j / j! \qquad (282)$$

and $\Delta(T)$ is the nxp driving matrix defining the effect of an input u_i at time $t = iT$ on the state of the solution at time $t = (i+1)T$, given by

$$\Delta(T) = \int_0^T e^{A(T-\tau)} B d\tau = \left\{ \sum_{j=0}^{\infty} A^j T^{j+1}/(j+1)! \right\} B \qquad (283)$$

$$= A^{-1}[\phi(T) - I]B \qquad (284)$$

with nonsingular matrix A. With disturbance and control inputs u_{1i}, u_{2i} the matrix Δ will include partitioned components Δ_1, Δ_2. Various computational algorithms have been developed for obtaining the matrices ϕ and Δ. A closed-form solution can be obtained using an eigenvalue analysis, the properties of the constituent matrices or a partial-fraction expansion of the equivalent transfer matrix. An arbitrary sampling period T can be defined but the methods usually require a complicated programming routine. A simpler computational algorithm with arbitrary sampling-period can be obtained using the squaring and translation properties associated with the matrix series expansions[160].

13.1.1 Solutions based on matrix-series expansions

A simple programming loop can be used to derive the expansions

$$\phi(t) = I + (At) + (At)At/2 + (A^2t^2/2!)At/3 + \ldots \qquad (285)$$

$$\Delta(t) = It + (It)At/2 + (At^2/2!)At/3 + (A^2t^3/3!)At/4 + \ldots \qquad (286)$$

where the brackets contain the previous terms in the expansion and each is multiplied by the expression At/j. Rapid convergence with limited iteration will be restricted to relatively small values of step length t for avoiding computer overflow. An extended-period solution can then be obtained using the repeated squaring and translation properties of the matrices ϕ and Δ. Thus for a required sampling period $T = nt$, where t is a reduced step length producing rapid convergence of $\phi(t)$ and $\Delta(t)$, the squaring properties of ϕ can be used to give

$$\begin{aligned} \phi(2t) &= \phi^2(t) \\ &\cdots\cdots \\ \phi(nt) &= \phi^2(nt/2) \end{aligned} \qquad (287)$$

Similarly the following relations may be developed for obtaining the driving matrix $\Delta(T)$

$$\Delta(2t) = \int_0^{2t} e^{A(2t-\tau)} B d\tau = A^{-1}[e^{A2t} - I]B \qquad (288)$$

reducing to

$$\Delta(2t) = [I + \phi(t)]\Delta(t)$$

$$\cdots\cdots\cdots\cdots \qquad (289)$$

$$\Delta(nt) = [I + \phi(nt/2)]\Delta(nt/2)$$

This algorithm may be programmed directly in conjunction with eqn 287, giving both $\phi(T)$ and $\Delta(T)$ based on the converging results for $\phi(t)$ and $\Delta(t)$ obtained from eqns 285 and 286.

13.1.2 Solutions based on an eigenvalue analysis

The response of the linear system may also be investigated in terms of the transient modes associated with the solutions of the characteristic determinant. These are related to the eigenvalue problem, which for a matrix A is defined by the relations

$$(\lambda_r I - A)u_r = 0 \qquad (290)$$

$$v_r(\lambda_r I - A) = 0 \qquad (291)$$

$$v_r u_r = 1, \quad v_r u_k = 0, \quad k \neq r, \quad r = 1..n \qquad (292)$$

where u_r and v_r represent the eigenvector and eigenrow respectively, associated with the eigenvalue λ_r. The eigenrow v_r is also defined as an adjoint eigenvector of the transposed matrix A'. In terms of a square modal matrix of eigenvector columns u_r the resulting eigenvector matrix equation is

$$AU = U\Lambda \qquad (293)$$

where Λ is a diagonal matrix with elements $\lambda_1, \lambda_2, ..\lambda_n$. The transition matrix based on calculated eigenvalues and eigenvectors can then be defined as

$$\phi(T) = e^{AT} = Ue^{\Lambda T}U^{-1} \qquad (294)$$

where $e^{\Lambda T}$ is a diagonal matrix with elements $e^{\lambda_1 T}, e^{\lambda_2 T}, ..e^{\lambda_n T}$. Similarly

$$\Delta(T) = \int_0^T Ue^{A(T-\tau)}U^{-1}Bd\tau = U\bar{\Lambda}U^{-1}B \qquad (295)$$

where

$$\bar{\Lambda} = \Lambda^{-1}[e^{\Lambda T} - I] \qquad (296)$$

is a diagonal matrix with elements $(e^{\lambda_1 T}-1)/\lambda_1, ..(e^{\lambda_n T}-1)/\lambda_n$.

13.1.3 Solutions based on the constituent matrices of Sylvester's expansion theorem

A discrete time solution can be defined in terms of the spectral decomposition of the matrix exponential obtained using Sylvester's theorem which expresses the polynomial function of a matrix, with n distinct eigenvalues, in the form

$$F(A) = \sum_{i=1}^{n} F(\lambda_i)G(\lambda_i) \qquad (297)$$

where $G(\lambda_i)$ is the n-square constituent matrix defined by

$$G(\lambda_i) = \frac{Q(\lambda_i)}{g'(\lambda_i)} = \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{(A - \lambda_j I)}{(\lambda_i - \lambda_j)} \qquad (298)$$

$Q(\lambda)$ is the adjoint of the characteristic matrix $(\lambda I-A)$ and $g'(\lambda)$ represents the derivative of the characteristic determinant $|\lambda I-A|$ with respect to λ. Thus the matrix exponential can be defined by the spectral decomposition

$$\phi(T) = e^{AT} = \sum_{i=1}^{n} e^{\lambda_i T} G(\lambda_i) \qquad (299)$$

Similarly for the driving matrix

$$\Delta(T) = \int_o^T e^{A(T-\tau)} B d\tau = \left\{ \sum_{i=1}^{n} \frac{(e^{\lambda_i T} - 1)}{\lambda_i} G(\lambda_i) \right\} B \qquad (300)$$

The constituent matrix may also be defined by the form of a normalised dyad expression, in terms of the outer product of an eigenvector and an eigenrow

$$G(\lambda_i) = u_i v_i \qquad (301)$$

Then $$e^{AT} = \sum_{i=1}^{n} e^{\lambda_i T} (u_i v_i) \qquad (302)$$

The constituent matrix sequence $G(\lambda_i)$ can thus be used in a computational algorithm for determining ϕ and Δ in terms of the eigenvalues and eigenvectors of the matrix A and its transpose, or in terms of the eigenvalues and the adjoint matrix of eqn 298. The matrices $G(\lambda_i)$ represent the set of constituent idempotent matrices of A with the properties

$$\sum_{i=1}^{n} G(\lambda_i) = UV = I, \; [G(\lambda_i)]^k = G(\lambda_i), \; G(\lambda_i)G(\lambda_j) = 0, \quad i \neq j \qquad (303)$$

13.1.4 <u>Solutions based on the partial-fraction expansion</u>

Laplace transforming the continuous-time state eqn 280 gives the transformed state solution

$$X(s) = (sI - A)^{-1} x(o) + (sI - A)^{-1} BU(s) \qquad (304)$$

where $(sI-A)^{-1}B$ represents the system transfer matrix. The transition matrix will then be represented by the closed-form solution

$$\phi(T) = \mathcal{L}^{-1} [(sI - A)^{-1}] \qquad (305)$$

Now

$$\lim_{\lambda \to \lambda_r} (\lambda - \lambda_r)(\lambda I - A)^{-1} = u_r v_r = \frac{Q(\lambda_r)}{g'(\lambda_r)} \qquad (306)$$

which represents the matrix of residues corresponding to the distinct poles in the partial-fraction solution associated with

$$(sI-A)^{-1} = \sum_{r=1}^{n} C_r (s-s_r)^{-1} \tag{307}$$

where $$C_r = \lim_{s \to s_r} (s-s_r)(sI-A)^{-1} \tag{308}$$

Thus $$(\lambda I-A)^{-1} = \sum_{r=1}^{n} (u_r v_r)(\lambda-\lambda_r)^{-1} \tag{309}$$

and inverse transforming with Laplace variable λ gives the transition matrix

$$\phi(T) = \mathcal{L}^{-1} \left[\sum_{r=1}^{n} (u_r v_r)(\lambda-\lambda_r)^{-1} \right] \tag{310}$$

corresponding to eqn 302. Columns of the matrix of residues C_r given by the outer product $u_r v_r$, and thus also of the adjoint matrix $Q(\lambda_r)$, are proportional to the eigenvector u_r, and the rows are proportional to the eigenrow v_r. Eqns 307 and 309 illustrate the relationship of the constituent matrices with the matrix of residues associated with the equivalent partial-fraction expansion of the transfer matrix. The Leverrier algorithm can also be used as a computational method for determining $(sI-A)^{-1}$ 161.

The constituent matrices play an important basic role in the theory of linear systems, and are also of fundamental importance in the eigenvalue sensitivity problem associated with linear systems, as discussed in Section 10. They are related to the algebraic eigenvalue problem and provide a unifying relationship between the various methods of solution developed for the time solution of linear vector differential equations. The use of the algorithm of eqn 310 will avoid the problems of convergence using the matrix series expansions based on arbitrary step lengths. It will, however, require the calculation of eigenvalues and residues, and for the usual, relatively large asymmetrical process model, difficulties will be introduced with the existence of complex eigenvalues. In this case, the preferred method of numerical solution for determining the matrices ϕ and Δ will be given by the series expansions of eqns 285 and 286 combined with the squaring and translation properties of eqns 287 and 289.

13.2 Optimal control of the discrete linear dynamic system

For optimal control we require to determine the control sequence $[u_o, u_1 .. u_{N-1}]$ which minimises the N-stage process performance index

$$J_N = \sum_{i=1}^{N} [x_i' Q x_i + u_{i-1}' G u_{i-1}] \tag{311}$$

subject to the constraints of eqn 1, Part 1, with $w_i = 0$. Q is an nxn positive semi-definite symmetric matrix and G is a pxp positive definite symmetric matrix. Including a measurement matrix H then $Q = H'QH$.

From the 'principle of optimality' in dynamic programming with a backward tracing of the trajectory, the minimum cost over any 'future' interval is a function only of the present state and the future control. Optimal performance at the Nth period associated with a single-stage process is then given by

$$J_1^o = \min_{u_{N-1}} (x_N' Q\, x_N + u_{N-1}' \, G\, u_{N-1}) \qquad (312)$$

Now relating the state x_N to conditions at the instant N-1 for proceeding backwards, using eqn 1, Part 1, gives

$$J_1 = x_{N-1}' \phi' Q \phi\, x_{N-1} + u_{N-1}' \Delta' Q \phi x_{N-1} + x_{N-1}' \phi' Q \Delta u_{N-1}$$

$$+ u_{N-1}' (\Delta' Q \Delta + G) u_{N-1} \qquad (313)$$

Setting $\partial J_1 / \partial u_{N-1} = 0$ for a single extreme (minimum) value then gives

$$u_{N-1} = -(\Delta' P_o \Delta + G)^{-1} \Delta' P_o \phi x_{N-1} = K_1 x_{N-1}, \; P_o = Q \qquad (314)$$

Matrix K_1 represents a linear feedback control law based on a single-step decision. For suboptimal control on the forward trajectory it relates control inputs to system states at each instant.

For control at instant N-2, summated performance

$$J_2^o = \min_{u_k} (J_1^o + x_{N-1}' Q x_{N-1} + u_{N-2}' G\, u_{N-2}), \; k = N-2, N-1 \qquad (315)$$

Then including eqn 312, related to conditions at the instant N-1, and eqn 314, gives

$$J_2 = x_{N-1}' P_1 x_{N-1} + u_{N-2}' G\, u_{N-2} \qquad (316)$$

where

$$P_1 = (\phi + \Delta K_1)' Q (\phi + \Delta K_1) + K_1' G K_1 + Q \qquad (317)$$

Now translating the state x_{N-1} to conditions at the instant N-2 using eqn 1 and setting $\partial J_2 / \partial u_{N-2} = 0$ gives

$$u_{N-2} = -(\Delta' P_1 \Delta + G)^{-1} \Delta' P_1 \phi x_{N-2} = K_2 x_{N-2} \qquad (318)$$

Similarly $J_3^o = \min_{u_k} (J_2^o + x_{N-2}' Q x_{N-2} + u_{N-3}' G u_{N-3}), \; k = N-3, N-2, N-1$

Then including eqn 316 related to N-2 and eqn 318

$$J_3^o = x_{N-2}' P_2\, x_{N-2} + u_{N-3}' \, G\, u_{N-3} \qquad (320)$$

and minimising with respect to u_{N-3} gives

$$u_{N-3} = -(\Delta' P_2 \Delta + G)^{-1} \Delta' P_2 \phi\, x_{N-3} = K_3\, x_{N-3} \qquad (321)$$

where

$$P_2 = (\phi + \Delta K_2)' P_1 (\phi + \Delta K_2) + K_2' G K_2 + Q \qquad (322)$$

Continuing this procedure to the initial state gives the optimal inputs as a solution of the functional equation

$$J_N^o = \min_{u_k} [x_1'Qx_1 + u_o'Gu_o + J_{N-1}^o(x_2,u_2,J_{N-2}^o)], \quad k = 0,1..N-1 \quad (323)$$

The problem thus reduces to minimisation with respect to a single input u_o $(=K_N x_o)$ as a single-point boundary value problem. In the limit as $N^o \to \infty$ with an infinite time interval, the control law and quadratic form sequence converge to limiting steady-state values. The dynamic programming algorithm for optimal control of the discrete linear system with quadratic index may be summarised -

$$u_{N-r} = K_r x_{N-r}, \quad r = 1,2..N \quad (324)$$

where K_r is a pxn feedback gain matrix determined by the recursive relations

$$K_r = -(\Delta'P_{r-1}\Delta + G)^{-1}\Delta'P_{r-1}\phi \quad (325)$$

$$P_{r-1} = (\phi + \Delta K_{r-1})'P_{r-2}(\phi + \Delta K_{r-1}) + K_{r-1}'GK_{r-1} + Q \quad (326)$$

$$P_o = Q, \quad r = 1$$

Substituting the relations for K_r into those for P_r the general algorithm may also be stated by eqns 324 and 325 with

$$P_{r-1} = \phi' M_{r-1}\phi + Q \quad (327)$$

$$M_{r-1} = P_{r-2} - P_{r-2}\Delta(\Delta'P_{r-2}\Delta + G)^{-1}\Delta'P_{r-2} \quad (328)$$

$$P_o = Q = H'\bar{Q}H \quad (329)$$

With the performance criterion including a term relating to the final state $(x_N'Fx_N)$ then

$$P_o = H'\bar{Q}H + F \quad (330)$$

With given matrices ϕ, Δ, Q, G and F the feedback control law matrices P_{r-1}, M_{r-1} and K_r will be computed offline, and with convergence in a desired interval the linear invariant feedback control law is given by

$$u_k = K x_k \quad (331)$$

In the limit as $N \to \infty$, the control algorithm can be related to a non-linear algebraic matrix equation by combining eqns 327 and 328 to give

$$P = \phi'P\phi - \phi'P\Delta(\Delta'P\Delta + G)^{-1}\Delta'P\phi + Q \quad (332)$$

13.3 Optimal control of the linear continuous dynamic system

Consider the optimal control of the system

$$\dot{x}_i(t) = f_i(x,u,t), \quad i = 1..n \quad (333)$$

for transferring the initial state $x(t_o)$ to a terminal state $x(T)$ with control $u(t)$ which minimises the performance functional

$$J = \min_{u(t)} \int_{t_o}^{T} f_o(x,u,t)dt \quad (334)$$

13.3.1 Solution based on the maximum principle

Define additional state variables

$$\dot{x}_o(t) = f_o(x,u,t), \quad x_o(t_o) = 0, \; x_o(T) = J \tag{335}$$

and adjoint variables

$$\dot{p}_i = -\sum_{j=o}^{n} \partial f_j/\partial x_i \cdot p_j, \quad i = 0,1..n \tag{336}$$

The Hamiltonian is then defined

$$H(p,x,u) = p'f(x,u,t) \tag{337}$$

thus $$\dot{x}_i = \partial H/\partial p_i, \quad \dot{p}_i = -\partial H/\partial x_i, \quad i = 0,1..n \tag{338}$$

For the linear system with quadratic performance

$$f_o = (x'Qx + u'Gu)/2 \tag{339}$$

$$H = p_o(x'Qx + u'Gu)/2 + p'(Ax + Bu) \tag{340}$$

$$\dot{p} = -p_oQx - A'p, \quad \dot{p}_o = 0 \tag{341}$$

with boundary conditions

$$p_o(t) = -1, \quad 0 \leqslant t \leqslant T, \quad p(T) = 0$$

For maximum H along an optimal trajectory, differentiating eqn 340 with respect to control u gives

$$u(t) = G^{-1} B' p(t) \tag{342}$$

Combining eqns 280, 342 and 341 then gives a 2n-dimensional vector differential equation defining the optimal trajectory

$$\begin{bmatrix} \dot{x}(t) \\ \dot{p}(t) \end{bmatrix} = \begin{bmatrix} A & BG^{-1}B' \\ Q & -A' \end{bmatrix} \begin{bmatrix} x(t) \\ p(t) \end{bmatrix} \quad \text{or} \quad \dot{z}(t) = M\, z(t) \tag{343}$$

with boundary conditions $x(t_o) = 0$, $p(T) = 0$.

Including the assumed relation $p(t) = -P\,x(t)$ in eqn 343 then gives the nonlinear matrix Riccati differential equation

$$\dot{P} = PBG^{-1}B'P - PA - A'P - Q \tag{344}$$

For the linear system with a quadratic performance criterion defined over a finite-time interval, the optimal control problem reduces to two sets of single-point boundary value problems for which explicit solutions can be obtained as in Section 8.

13.3.2 Solution based on the continuous form of dynamic programming

The continuous form of dynamic programming can be developed as a partial differential equation[118]

$$\underset{u(t)}{\text{Min}} \left\{ f_o(x,u,t) + \sum_{i=1}^{n} f_i\,(x,u,t)\frac{\partial J}{\partial x_i} \right\} + \frac{\partial J}{\partial t} = 0 \qquad (345)$$

For the linear system with

$$J = x'Px\,, \quad \partial J/\partial t = x'\dot{P}x, \quad \partial J/\partial x = 2Px \qquad (346)$$

eqn 345 gives

$$\underset{u}{\text{Min}} \left\{ (x'Qx + u'Gu) + (Ax + Bu)'Px + x'P(Ax + Bu) \right\} + x'\dot{P}x = 0 \qquad (347)$$

and differentiating with respect to u gives the optimal control law

$$u(t) = -G^{-1}B'P\,x(t) \qquad (348)$$

Combining eqns 347 and 348 then gives the nonlinear matrix Riccati differential eqn 344.

For integration with reverse time, $\tau = T-t$,

$$\dot{P}(\tau) = PA + A'P + Q - PBG^{-1}B'P\,, \quad P(\tau_o) = P(T) = 0 \qquad (349)$$

The optimal control law is thus time varying and related to the nxn symmetrical matrix P given by the solution of the matrix Riccati differential equation. Integration in reverse time requires storage of $P(\tau)$ for implementation of $u(t)$ and may cause realisation to be uneconomic.

The discrete reverse-time control algorithm of eqns 327 and 328 of Section 13.2 developed by dynamic programming can also be shown to reduce, in the limit, to the reverse-time form of eqn 349 for the continuous system. Thus combining eqns 327 and 328 and using the first-order approximations of eqns 192 with small sampling interval h, and arguments $r, r-1 \rightarrow t+h, t$, and following the derivation in Section 8 gives

$$M(t+h) = P(t) - P(t)(Bh)[(Bh)'P(t)(Bh) + hG]^{-1}(Bh)'P(t) \qquad (350)$$

Then

$$\begin{aligned} P(t+h) &= (I+Ah)'M(t+h)(I+Ah) + hQ \\ &= P(t) + hA'P(t) + hP(t)A - hP(t)B[hB'P(t)B+G]^{-1}B'P(t) \\ &\quad + O(h^2) + hQ \end{aligned} \qquad (351)$$

reducing in the limit $h \rightarrow o$ to the form of eqn 349. The control law algorithm of eqn 325 will reduce similarly to the continuous-time optimal control solution. Thus

$$K(t) = -[(Bh)'P(t-h)(Bh) + hG]^{-1}(Bh)'P(t-h)(I+Ah) \qquad (352)$$

and with $h \rightarrow 0$

$$K = -G^{-1}B'P \qquad (353)$$

14. References

1. LANCZOS, C.: 'Applied analysis', Pitman, 1957
2. PLACKETT, R.L.: 'A historical note on the method of least squares', Biometrikia, 1949, 36, pp.458-460
3. ROSENBROCK, H.H.: 'On the connection between discrete linear filters and some formulae of Gauss', Control Group report, Univ. of Cambridge, 1965.
4. KALMAN, R.E.: 'A new approach to linear filtering and prediction problems', Trans. Amer. Soc. Mech. Engrs., 1960, 82, [D], pp.35-45
5. KALMAN, R.E., and BUCY, R.S.: 'New results in linear filtering and prediction theory', ibid, 1961, 83, pp.95-108
6. MARSHALL, S.A., and NICHOLSON, H.: 'Optimal control of linear multivariable systems with quadratic performance criteria', Research report CE69-1, Dept. of Control Eng., Univ. of Sheffield
7. AITKEN, A.C.: 'On least squares and linear combination of observations', Proc. Roy. Soc. Edinburgh, 1934-35, LV, pp.42-48
8. DWYER, P.S.: 'A matrix representation of least squares and correlation theory with matrix justification of improved methods of solution', Annals Math. Stats., 1944, 15, pp.82-89
9. SMITH, G.L.: 'The scientific inferential relationships between statistical estimation, decision theory, and modern filter theory', JACC, 1965, New York, pp.350-359
10. SAGE, A.P., and MASTERS, G.W.: 'Least-squares curve fitting and discrete optimum fitting', IEEE Trans. Educ., 1967, E-10, 1, pp.29-36
11. AOKI, M.: 'Optimization of stochastic systems', Academic Press, 1967
12. LUENBERGER, D.G.: 'Optimization by vector space methods', Wiley, 1969
13. DeRUSSO, P.M., ROY, R.J., and CLOSE, C.M.: 'State variables for engineers', Wiley, 1965
14. ROSENBERG, A.E., and SHEN, D.W.C.: 'Regression analysis and its application to the system identification problem', JACC, 1963, Univ. of Minnesota, pp.446-451
15. HO, Y.C.: 'On the stochastic approximation method and optimal filtering theory', Jl. Math. Anal. Applicns., 1962, 6, pp.152-154
16. LEE, R.C.K.: 'Optimal estimation, identification, and control', MIT Press, 1964
17. GURA, I.A.: 'Extension of linear estimation techniques to nonlinear problems', Jl. Astro. Sci., 1968, XV, 4, pp.194-206
18. PLACKETT, R.L.: 'Some theorems in least squares', Biometrika, 1950, 37, pp.149-157
19. HOUSEHOLDER, A.S.: 'Principles of numerical analysis', McGraw-Hill, NY, 1953, p.79
20. GRABBE, E.M., RAMO, S., and WOOLRIDGE, D.E.: 'Handbook of automation, computation, and control', Vol.1, 1958, Wiley
21. HO, Y.C., and LEE, R.C.K.: 'A Bayesian approach to problems in stochastic estimation and control', JACC, 1964, Stanford Univ., pp.382-387
22. AOKI, M., and HUDDLE, J.R.: 'Estimation of the state vector of a linear stochastic system with a constrained estimator', IEEE Trans. Auto. Control, 1967, 12, 4, pp.432-433
23. COX, H.: 'Estimation of state variables via dynamic programming', JACC, 1964, Stanford Univ., pp.376-381
24. COX, H.: 'On the estimation of state variables and parameters for noisy dynamic systems', Trans. Auto. Control, 1964, 9, 1, pp.5-12
25. MOWERY, V.O.: 'Least squares recursive differential-correction estimation in nonlinear problems', IEEE Trans. Auto. Control, 1965, 10, 4, pp.399-407
26. OHAP, R.F., and STUBBERUD, A.R.: 'A technique for estimating the state of a nonlinear system', ibid., 1965, 10, 2, pp.150-155

27. GIESE, C., and McGHEE, R.B.: 'Estimation of nonlinear system states and parameters by regression methods', JACC, 1965, New York, pp.46-55
28. NOTON, A.R.M., and CHOQUETTE, P.: 'The application of modern control theory to the operation of an industrial process', Research report, Univ. of Waterloo, 1967
29. LUBBOCK, J.K.: 'Nonlinear filters', Control Group report, Univ. of Cambridge, 1963
30. BUCY, R.S.: 'Recent results in linear and non-linear filtering', paper in 'Stochastic problems in control', JACC, 1968, Univ. of Mich., (ASME), pp.87-105
31. ATHANS, M., and TSE, E.: 'A direct derivation of the optimal linear filter using the maximum principle', IEEE Trans. Auto. Control, 1967, 12, 6, pp.690-698
32. KRON, G.: 'Multidimensional curve-fitting with self-organizing automata', Jl. Math. Anal. Applicns., 1962, 5, pp.46-69
33. HARRISON, B.K.: 'A discussion of some mathematical techniques used in Kron's method of tearing', Jl. Soc. Indust. Appl. Math., 1963, 11, 2, pp.258-280
34. FU, K.S.: 'Sequential methods in pattern recognition and machine learning', Academic Press, 1968
35. NICHOLSON, H.: 'Construction of dynamical system models', IEE Summer school on control theory, survey lecture, Manchester, July 1967
36. NICHOLSON, H.: 'Control of sequential subsystem processes interconnected with transport delays', Proc. IEE, 1967, 114, 10, pp.1562-1568
37. KUMAR, K.S.P.: 'Identification of nonlinear, nonstationary processes', Proc. IFAC Tokyo Sympm. on Systems eng. for control system design, 1965, 6, pp.237-243
38. HO, Y.C., and LEE, R.C.K.: 'Identification of linear dynamic systems', Third symposium on discrete adaptive processes, Chicago, IEEE, 1964, pp.86-101
39. CAMERON, A.V.: 'A closed form recursive learning filter for a class of processes with unknown parameters', IFAC Symposium, Sydney, 1968, pp.161-163
40. JENKINS, K.W., and ROY, R.J.: 'A design procedure for discrete adaptive control systems', JACC, 1966, Univ. of Washington, pp.624-633
41. KOPP, R.E., and ORFORD, R.J.: 'Linear regression applied to system identification for adaptive control systems', AIAA Jl.,1963,1,10,pp.2300-2306
42. BRIGGS, P.A.N., CLARKE, D.W., and HAMMOND, P.H.: 'Introduction to statistical identification methods in control systems', Control, March 1968, pp.233-238
43. REES, N.W.: 'The adaptive control of an absorption tower', IFAC Symposium, Sydney, 1968, pp.129-132
44. BERMAN, M., SHAHN, E., and WEISS, M.F.: 'The routine fitting of kinetic data to models', Biophysical Jl., 1962, 2, pp.275-287
45. ROBINSON, E.A.: 'An introduction to infinitely many variates', Griffin,1959
46. BARTLETT, M.S.: 'An introduction to stochastic processes', CUP, 1966
47. BOX, G.E.P., and JENKINS, G.M.: 'Some statistical aspects of adaptive optimisation and control', Jl. Royal Stat. Soc., 1962, 24, pp.297-343
48. WOLD, H.O.A.: 'On prediction in stationary time series', Annals Math. Stats., 1948, 19, pp.558-567
49. FRAZER, R.A., JONES, W.P., and SKAN, S.W.: 'Approximations to functions and to the solutions of differential equations', R and M No. 1799, March 1937, pp.517-549
50. DUNCAN, W.J.: 'Galerkin's method in mechanics and differential equations', R and M No. 1798, Aug. 1937, pp.484-516
51. BOX, G.E.P., and JENKINS, G.M.: 'Mathematical models for adaptive control and optimisation', Procs. Sympm. on Applns. of Math. Models in Chem. Eng. Research, Design, and Prodn., A.I.Chem.E.-I.Chem.E., London, 1965, 4, pp.61-68

52. JENKINS, G.M.: 'An example of the estimation of a linear open loop transfer function', Technometrics, 1963, 5, 2, pp.227-245
53. EYKHOFF, P., van der GRINTEN, P.M.E.M., KWAKERNAAK, H., and VELTMAN, B.P.T.: 'Systems modelling and identification', Survey paper, Procs. third congress IFAC, London, 1966
54. PEARSON, J.D.: 'Deterministic aspects of the identification of linear discrete dynamic systems', IEEE Trans. Auto. Control, 1967, 12, 6, pp.764-766
55. FARISON, J.B.: 'Parameter identification for a class of linear discrete systems', ibid., 1967, 12, 1, p.109
56. SWANICK, B.H., and GRAUPE, D.: 'Analysis of on-line identification and control of a nuclear reactor', Jl. British Nucl. Energy Soc., 1968, 7, 1, pp.100-113
57. GRAUPE, D., and CASSIR, G.R.: 'A predictive identification and mode-recognition technique for adaptive control and tracking systems', Proc. IFAC Sympm. on Identification, 1967, Prague, Part 2, paper 4.1, pp.1-7
58. BECK, M.S.: 'Learning method for identification of process dynamics', Electronics Letters, 1968, 4, 4, pp.64-66
59. FUKAO, T.: 'System identification by Bayesian learning process', Proc. IFAC Tokyo Sympm. on Systems eng. for control system design, 1965, 4, pp.137-146
60. PEARSON, A.E.: 'A regression analysis in function space for the identification of multivariable linear systems', Procs. third congress IFAC, London, 1966
61. ROY, R., and SCHLEY, C.: 'Process identification using a mode learning machine', JACC, 1966, Univ. of Washington, pp.639-648
62. SKLANSKY, J.: 'Learning systems for automatic control', IEEE Intern. Conv., 1965, Part 6, pp.117-122
63. HSIEH, H.C.: 'The least squares estimation of linear and nonlinear system weighting function matrices', Inf. and Control, 1964, 7, pp.84-115
64. NICHOLSON, H.: 'Dual-mode control of a time-varying boiler model with parameter and state estimation', Proc. IEE, 1965, 112, 2, pp.383-395
65. TOU, J.T.: 'Optimum design of digital control systems', Academic Press,1963
66. BUCY, R.S., and JOSEPH, P.D.: 'Filtering for stochastic processes with applications to guidance', Interscience, 1968
67. ALBERT, G.E.: 'Statistical methods in prediction, filtering, and detection problems', J. Soc. Indust. Appl. Math., 1960, 8, 4, pp.640-653
68. HAWKINS, J.K.: 'Self-organising systems - a review and commentary', Proc. IRE, 1961, 49, 1, pp.31-48
69. FARMER, E.D.: 'A method of prediction for non-stationary processes and its application to the problem of load estimation', Procs. second congress IFAC, Basle, 1963, Theory, pp.47-54
70. FARMER, E.D., and POTTON, M.J.: 'The prediction of load on a power system', Procs. third congress IFAC, London, 1966, paper 21F, pp.1-7
71. MATTHEWMAN, P.D., and NICHOLSON, H.: 'Techniques for load prediction in the electricity-supply industry', Proc. IEE, 1968, 115, 10, pp.1451-1457
72. TOYAMA, M., and YAMADA, K.: 'Daily load prediction on the basis of Farmer's method', Elect. Eng. Japan, 1967, 87, 7, pp.50-58
73. TSUI, H.T.: 'The computer control of a distribution network', Ph.D. thesis, Univ. of Birmingham, 1969
74. DAVENPORT, W.B., and ROOT, W.L.: 'An introduction to the theory of random signals and noise', McGraw-Hill, 1958
75. LANING, J.H., and BATTIN, R.H.: 'Random processes in automatic control', McGraw-Hill, 1956

76. SELIN, I.: 'The sequential estimation and detection of signals in normal noise, I', Inf. and Control, 1964, 7, pp.512-534
77. MAYNE, D.Q.: 'Optimal non-stationary estimation of the parameters of a linear system with Gaussian inputs', Jl. Electronics and Control, 1963, XIV, 1, pp.101-112
78. BROWN, J.L.: 'Mean square truncation error in series expansions of random functions', Jl. Soc. Indust. Appl. Math., 1960, 8, 1, pp.28-32
79. OESTREICHER, H.L., and MOORE, D.R.: 'Cybernetic problems in bionics', Bionics Symp., 1966, Gordon and Breach, p.700
80. KONDO, B., and SUZUKI, T.: 'Optimalizing and learning control by a statistical model', Proc. IFAC Tokyo Sympm. on Systems eng. for control system design, 1965, 2, pp.69-76
81. TOU, J.T.: 'Feature extraction in pattern recognition', Pattern Recognition, vol. 1, Pergamon Press, 1968, pp.3-11
82. THOMAS, J.B.: 'The representation of stochastic processes by orthogonal expansions', JACC, 1963, Univ. of Minnesota, pp.287-291
83. WHITTLE, P.: 'Prediction and regulation by linear least-square methods', EUP, 1963
84. GRENANDER, U.: 'Stochastic processes and statistical inference', Arkiv Mat 1, 1950, 195
85. KELLY, E.J., REED, I.S., and ROOT, W.L.: 'The detection of radar echoes in noise, Parts I and II', Jl. Soc. Indust. Appl. Math., 1960, 8, Nos 2 and 3
86. SLEPIAN, D.: 'Estimation of signal parameters in noise', IRE Prof. Group Inf. Theory, 1954, 3, pp.68-89
87. LENGYEL, B.A.: 'Estimation of the accuracy of approximate wave functions', Jl. Math. Anal. Applcns., 1962, 5, pp.451-460
88. AIZERMAN, M.A., BRAVERMAN, E.M., and ROZONOER, L.I.: 'Theoretical foundations of the potential function method in pattern recognition learning', Automation and Remote Control, 1964, pp.821-837
89. TOU, J.T.: Editor, 'Computer and Information Sciences - II', Acad. Press, 1967, Article by TSYPKIN, Ya.Z., 'Optimisation, adaptation, and learning in automatic systems', p.15
90. TOU, J.T.: ibid., Article by TOU, J.T., and HEYDORN, R.P., 'Some approaches to optimum feature extraction', p.57
91. TOU, J.T.: ibid., Article by SIMMONS, G.J., 'Iterative storage of multidimensional functions in discrete distributed memories', p.261
92. GREVILLE, T.N.E.: 'The pseudoinverse of a rectangular or singular matrix and its application to the solution of systems of linear equations', SIAM Review, 1959, 1, 1, pp.38-43
93. GREVILLE, T.N.E.: 'Some applications of the pseudoinverse of a matrix', SIAM Review, 1960, 2, 1, pp.15-22
94. PENROSE, R.: 'A generalised inverse for matrices', Proc. Camb. Phil. Soc., 1955, 51, 3, pp.406-413
95. PENROSE, R.: 'On best approximate solutions of linear matrix equations', ibid., 1956, 52, 1, pp.17-19
96. BJERHAMMAR, A.: 'Application of calculus of matrices to method of least squares; with special reference to geometric calculations', Trans. Roy. Inst. Tech. Stockholm, 1951, 49, pp.1-86 and pp.188-220
97. RAO, C.R.: 'A note on a generalised inverse of a matrix with applications to problems in mathematical statistics', Jl. Royal Stat. Soc., 1962, 24, pp.152-158
98. GREVILLE, T.N.E.: 'Note on fitting of functions of several independent variables', Jl. Soc. Indust. Appl. Math., 1961, 9, 1, pp.109-115
99. BEN-ISRAEL, A., and CHARNES, A.: 'Contributions to the theory of generalised inverses', ibid., 1963, 11, 3, pp.667-699
100. RAO, C.R.: 'Calculus of generalized inverses of matrices, Part 1 - general theory', The Indian Jl. of Statistics, 1967, A, 29, 3, pp.317-342

101. BOOT, J.C.G.: 'Quadratic programming', North-Holland, 1964
102. FOSTER, M.: 'An application of the Wiener-Kolmogorov smoothing theory to matrix inversion', Jl. Soc. Indust. Appl. Math., 1961, 9, 3, pp.387-392
103. LEONDES, C.T.: Editor, 'Advances in control systems', Vol 1, Acad. Press, 1964, Article by KISHI, F.H., 'On line computer control techniques and their application to re-entry aerospace vehicle control', pp.245-357
104. MAYNE, D.: 'An algorithm for the calculation of the pseudo-inverse of a singular matrix', Computer Jl., 1966-67, 9, pp.312-317
105. JOSEPH, P.D., and TOU, J.T.: 'On linear control theory', Trans. AIEE, Appl. and Ind., Pt. II, 1961, 80, pp.193-196
106. GUNCKEL, T.L., and FRANKLIN, G.F.: 'A general solution for linear sampled-data control', Trans. ASME, 1963, D, 85, pp.197-201
107. KALMAN, R.E.: 'On the general theory of control systems', Procs. first congress IFAC, Moscow, 1960, pp.481-492
108. WONHAM, W.M.: 'Optimal stochastic control', paper in 'Stochastic problems in control', JACC, 1968, Univ. of Mich., (ASME), pp.107-120
109. COSAERT, R.: 'On the estimation of the plant state, plant parameters and external disturbances from observations of the output of a control system', JACC, 1964, Stanford Univ., pp.340-348
110. BELLMAN, R.: 'Introduction to the mathematical theory of control processes, Vol 1, Linear equations and quadratic criteria', Acad. Press, 1967
111. FREESTED, W.C., WEBBER, R.F., and BASS, R.W.: 'The 'GASP' computer program - an integrated tool for optimal control and filter design', JACC, 1968, Univ. of Mich., (ASME), pp.198-202
112. POTTER, J.E.: 'Matrix quadratic solutions', Jl. Soc. Indust. Appl. Math., 1966, 14, 3, pp.496-501
113. STERNBERG, R.L., and KAUFMAN, H.: 'Applications of the theory of systems of differential equations to multiple non-uniform transmission lines', Jl. Maths. and Physics, 1952, 31, pp.244-252
114. MOORE, J.B.: 'Application of Riccati equations in system engineering', IFAC Symposium, Sydney, 1968, pp.179-184
115. TYLER, J.S.: 'The characteristics of model-following systems as synthesized by optimal control', IEEE Trans. Auto. Control, 1964, 9, 4, pp.485-498
116. MEDITCH, J.S.: 'A note on optimal stochastic linear control', Procs. JACC, 1968, Univ. of Mich., pp.626-633
117. MEDITCH, J.S.: 'Stochastic optimal linear estimation and control', McGraw-Hill, 1969
118. NOTON, A.R.M.: 'Introduction to variational methods in control engineering', Pergamon Press, 1965
119. BLACKBURN, T.R.: 'Solution of the algebraic matrix Riccati equation via Newton-Raphsen iteration', JACC, 1968, Univ. of Mich., pp.940-945
120. MAYNE, D.Q.: 'Parameter estimation', Automatica, 1966, 3, 3/4, pp.245-255
121. NICHOLSON, H.: 'Dynamic optimisation of a boiler', Proc.IEE, 1964, 111, 8, pp.1479-1499
122. MARSHALL, S.A.: 'An approximate method for reducing the order of a linear system', Control, 1966, Dec., pp.642-643
123. DAVISON, E.J.: 'A method for simplifying linear dynamic systems', IEEE Trans. Auto. Control, 1966, 11, 1, pp.93-101
124. DAVISON, E.J.: 'On simplifying linear dynamic systems', ibid., 1967, 12, 1, pp.119-121
125. DAVISON, E.J.: 'The simplification of large linear systems', Control, 1968, May, pp.418-419
126. ANDERSON, J.H.: 'Geometrical approach to reduction of dynamical systems', Proc. IEE, 1967, 114, 7, pp.1014-1018

127. BROWN, R.F.: 'Reduction of order of linear state equations', Electronics Letters, 1968, 4, 19, pp.416-417
128. MITRA, D.: 'On the reduction of complexity of linear dynamical models', AEE, Winfrith, Reports 520, 535
129. MITRA, D.: 'W matrix and the geometry of model equivalence and reduction', Proc. IEE, 1969, 116, 6, pp.1101-1106
130. HAMMOND, P.H., and WILLIAMS, B.J.: 'On determining a dynamic model of a binary distillation column using a computer program', Second UKAC Control Convention, 'Advances in computer control', 1967, C9, pp.1-22
131. NEAL, S.R.: 'Linear estimation in the presence of errors in assumed plant dynamics', IEEE Trans. Auto. Control, 1967, 12, 5, pp.592-594
132. ROSENBROCK, H.H.: 'Sensitivity of an eigenvalue to changes in the matrix', Electronics Letters, 1965, 1, pp.278-279
133. MORGAN, B.S.: 'Computational procedure for the sensitivity of an eigenvalue', ibid., 1966, 2, pp.197-198
134. LAUGHTON, M.A.: 'Sensitivity in dynamical system analysis', Jl. Electronics and Control, 1964, 17, pp.577-591
135. VAN NESS, J.E., BOYLE, J.M., and IMAD, F.P.: 'Sensitivities of large, multiple-loop control systems', IEEE Trans. Auto. Control, 1965, 10, 3, pp.308-315
136. REDDY, D.C.: 'Sensitivity of an eigenvalue of a multivariable control system', Electronics Letters, 1966, 2, 12, p.446
137. REDDY, D.C.: 'Sensitivity of an eigenvector of a multivariable control system', ibid., 1967, 3, 5, pp.205-207
138. REDDY, D.C.: 'Eigenfunction sensitivity and the parameter variation problem', Int. Jl. Control, 1969, 9, 5, pp.561-568
139. NICHOLSON, H.: 'Eigenvalue and state-transition sensitivity of linear systems', Proc. IEE, 1967, 114, 12, pp.1991-1995
140. CROSSLEY, T.R., and PORTER, B.: 'Eigenvalue and eigenvector sensitivities in linear systems theory', Int. Jl. Control, 1969, 10, 2, pp.163-170
141. GRIFFIN, R.E., and SAGE, A.P.: 'Large and small scale sensitivity analysis of optimum estimation algorithms', JACC, 1968, Univ. of Mich., pp.977-988
142. PAGUREK, B.: 'Sensitivity of the performance of optimal linear control systems to parameter variations', Int. Jl. Control, 1965, 1, pp.33-45
143. BARNETT, S.: 'Sensitivity of optimal linear systems to small variations in parameters', ibid., 1966, 4, 1, pp.41-48
144. BARNETT, S., and STOREY, C.: 'Insensitivity of optimal linear control systems to persistent changes in parameters', ibid., 1966, 4, 2, pp.179-184
145. CHEN, R.T.N., and SHEN, D.W.C.: 'Sensitivity analysis and design of multivariable regulators using a quadratic performance criterion', JACC, 1968, Univ. of Mich., pp.229-238
146. REES, N.W.: 'The adaptive control of an absorption tower', IFAC Sympm., Sydney, 1968, pp.129-132
147. KOPP, R.E., and ORFORD, R.J.: 'Linear regression applied to system identification for adaptive control systems', AIAA Jl., 1963, 1, 10, pp.2300-2306
148. GRAUPE, D., SWANICK, B.H., and CASSIR, G.R.: 'Application of regression analysis to reduction of multivariable control problems and to process identification', Proc. of Natl. Elects. Conf., 1967, XXIII, pp.20-25
149. NEAL, S.R.: 'Nonlinear estimation techniques', IEEE Trans. Auto. Control, 1968, 13, 6, pp.705-708

150. PRICE, C.F.: 'An analysis of the divergence problem in the Kalman filter', ibid., 1968, 13, 6, pp.699-702
151. HEFFES, H.: 'The effect of erroneous models on the Kalman filter response', ibid., 1966, 11, 3, pp.541-543
152. NOTON, A.R.M.: 'The extension of certain results in the treatment of inaccessible state variables', ibid., 1966, 11, 4, pp.669-675
153. NISHIMURA, T.: 'On the a priori information in sequential estimation problems', ibid., 1966, 11, 2, pp.197-204
154. KAILATH, T.: 'An innovations approach to least-squares estimation, Part 1, Linear filtering in additive white noise', ibid., 1968, 13, 6, pp.646-655
155. SWERLING, P.: 'First order error propagation in a stagewise smoothing procedure for satellite observations', Jl. Astro. Sci., 1959, 6, 3, pp.46-52
156. GLUSS, B.: 'An alternative method for continuous line segment curve-fitting', Inf. and Control, 1964, 7, 2, pp.200-206
157. BLUM, M.: 'Best linear unbiased estimation by recursive methods', J. Siam Appl. Math., 1966, 14, 1, pp.167-180
158. REDDY, D.C.: 'Eigenfunction sensitivity and the parameter variation problem', Int. Jl. Control, 1969, 9, 5, pp.561-568
159. ENNS, M., GREENWOOD, J.R., MATHESON, J.E., and THOMPSON, F.T.: 'Practical aspects of state-space methods, Part 2', JACC, 1964, Stanford Univ., pp.501-513
160. NICHOLSON, H.: 'Dynamic optimisation of a boiler-turboalternator model', Proc. IEE, 1966, 113, 2, pp.385-399
161. MORGAN, B.S.: 'Sensitivity analysis and synthesis of multivariable systems', IEEE Trans. Auto. Control, 1966, 11, 3, pp.506-512
162. NICHOLSON, H.: 'On-line digital control of industrial processes - a research programme', Instrument Practice, Sept. 1969, pp.640-644
163. BARNETT, S., FORD, B., and STOREY, C.: 'A comparison of several methods for calculating the general functional matrix for linear systems', Third UKAC Control Convention, 'Advances in control of systems', 1968, paper A4, pp.1-11.