COVARIANCE AND GRAMIAN MATRICES IN CONTROL AND SYSTEMS THEORY

KURUKULASURIYA VINCENZA FERNANDO

A thesis submitted for the degree of DOCTOR OF PHILOSOPHY in the FACULTY OF ENGINEERING, the UNIVERSITY OF SHEFFIELD

September 1982

DEPARTMENT OF CONTROL ENGINEERING

I have fought a good fight, I have finished my course

2 Timothy 4.7

ACKNOWLEDGEMENTS

The author is grateful to his mentor, Professor H. Nicholson, for his guidance. The author also wishes to acknowledge the University of Sheffield for financial assistance in the form of the Edgar Allen Scholarship and the Linley Scholarship for doctoral research. The manuscript was skilfully prepared by Mrs. Josephine Stubbs.

Covariance and Gramian matrices in control and systems theory

K. V. FERNANDO

SUMMARY

Covariance and Gramian matrices in control and systems theory and pattern recognition are studied in the context of reduction of dimensionality and hence complexity of large-scale systems. This is achieved by the removal of redundant or 'almost' redundant information contained in the covariance and Gramian matrices. The Karhunen-Loeve expansion (principal component analysis) and its extensions and the singular value decomposition of matrices provide the framework for the work presented in the thesis. The results given for linear dynamical systems are based on controllability and observability Gramians and some new developments in singular perturbational analysis are also presented.

CONTENTS

PART 1

Exordium

CHAPTER 1

Exordium

1. The prelude

The classical historian Edward Gibbon (1737-1794) noted in his autobiography that independence is the first of the earthly blessings. In an earthly subject such as mathematics, especially in linear analysis, 'independence' is often the first concept which has to be comprehended.

The independence in linear analysis (see any standard text on linear algebra or matrix theory, Mirsky[6], Gantmacher[7]) can be defined as follows. If x^i, $i = 1,m$ denotes a set of real n-vectors, then this set $\{x^i\}$ is said to be linearly dependent if there are scalar values a_i, $i = 1,m$ (not all zeros) such that

$$\sum_{i=1}^{m} a_i x^i = 0 \qquad (1)$$

In the contrary case, that is equation (1) implies that $a_i = 0$ for all i, then the set $\{x^i\}$ is said to be independent.

One of the important concepts associated with linear independence is rank. Let $\{x^{(i)}\}$ denote all possible permutations of indexing of the vectors x^i, $i = 1,m$. The rank is then defined as the largest integer r such that

$$\sum_{i=1}^{r} a_i x^{(i)} \neq 0$$

for non-zero a_i, $i = 1,r$. If r is equal to m, then the set $\{x^i\}$ is said to be full rank. If the null vector does not belong to this set, then

$$1 \leq r \leq n$$

According to the definition, verification of linear independence requires an uncountable infinite number of tests for all possible values of a_i, i = 1,m which is obviously impractical. Fortunately, there is a simple test for verification of independence which is due to J. P. Gram (1850-1916). Gram's monumental contribution[1] was submitted in the year 1881 and was published in 1883, and thus we are in the midst of important centennial anniversaries.

2. The Gramian matrix

If the vector set $\{x^i\}$ is considered as column vectors, then a matrix of n,m dimensions can be formed as follows.

$$X = [x^1 \mid x^2 \mid \ldots \mid x^m]$$

The matrix product defined by $G^2 = X^T X$ (X^T is the transpose of X) is known as the Gram matrix or the Gramian, and its determinant as the Gram determinant. The Gramian G^2 can also be written in the format,

$$G^2 = \sum_{i=1}^{m} (x^i)^T x^i$$

The following may be considered as the fundamental result in linear algebra.

Theorem: The Gram determinant is zero if and only if the set $\{x^i\}$ is linearly dependent.

Interesting variations of this result are available in most texts on linear algebra and matrix theory. Since the determinant of a matrix is given by the product of its eigenvalues, then the linear independence is guaranteed if all the eigenvalues of the Gramian are non-zero. Equivalently, if the Gramian G^2 is positive definite,

then the set $\{x^i\}$ is linearly independent. The notation G^2 is used to show the non-negative definiteness of the Gramian. This notation is used in this thesis whenever it is required to emphasise this property.

If the spectral expansion of the Gramian G^2 is written in the form,

$$G^2 = UD^2U^T \quad , \quad U^TU = I \quad , \quad D^2 = \text{diag}\,(d_1^2,\ldots,d_m^2)$$

where U is an orthonormal matrix, I is the identity matrix and D^2 is a diagonal matrix, then the Gram determinant is given by

$$\det G^2 = \prod_{i=1}^{m} d_i^2$$

If the determinant is 'small' due to small eigenvalues, then this determinant is 'almost' zero showing that the set $\{x^i\}$, although theoretically independent, is almost near to degeneration to a set of rank less than m. This theme is present throughout the thesis where almost dependent subspaces are removed to simplify theoretical or computational problems.

Similarly, as in the problem of determining the linear independence of column vectors, if $\{y^i\}$ is a set of row vectors, then a matrix Y can be formed as

$$Y = \begin{pmatrix} y^1 \\ \hline y^2 \\ \vdots \\ \hline y^m \end{pmatrix}$$

The Gram matrix in this case can be defined as

$$\bar{G}^2 = YY^T = \sum_{i=1}^{m} y^i (y^i)^T$$

This Gramian $\bar{G}^2$ can be used in exposing the linear independence of the row vector set $\{y^i\}$. Gram matrices of the form, X^TX and XX^T appear in this thesis, especially in connection with the singular value decomposition of the matrix X (see section 4).

3. Covariance matrices

If M^2 is a positive definite matrix of n,n dimensions and α is an n-vector, then a function p(x) can be defined as

$$p(x) = \frac{1}{(2\pi)^{n/2}\det M} e^{-\frac{1}{2}(x-\alpha)^T M^{-2}(x-\alpha)} \qquad (2)$$

which has the property

$$\int_{-\infty}^{\infty} p(x)dx = 1 \qquad (3)$$

where dx is the infinitesimal element

$$dx = dx_1 \ldots dx_n, \quad x = (x_1, \ldots x_n)^T$$

If x is a random vector with the probability density function p(x) then $\{x\}$ is called a Gaussian (or normal) process[8].

The following first- and second-order moments can be easily verified,

$$E\{x\} = \alpha$$

$$E\{(x-\alpha)(x-\alpha)^T\} = M^2$$

where $E\{\cdot\}$ denotes the expectation operator. The vector α is called the 'average value' and the matrix M^2 the covariance matrix of the process.

If the spectral decomposition of the matrix M^2 is of the form

$$M^2 = UD^2U^T \ , \ U^TU = I \ , \ D^2 = \mathrm{diag}(d_1^2, \ldots, d_n^2)$$

then equation (2) can be written in the form

$$p(y) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{\frac{1}{2}} d_i^{\,2}} e^{-\frac{1}{2}(y_i-\beta_i)^2/d_i^{\,2}}$$

where $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = U^T x$, $\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = U^T \alpha$

If $d_i^{\,2}$ is almost zero then

$$\frac{1}{(2\pi)^{\frac{1}{2}} d_i^{\,2}} e^{-(y_i-\beta_i)^2/d_i^{\,2}} \approx 0 \qquad \text{except near } y_i = \beta_i$$

Thus, as in the case of Gramian matrices, we may ignore y_i due to 'statistical' dependence. Such, removal of dependent data is a recurrent theme in this thesis. We also observe that the validity of equation (3) can be verified using this canonical transformation $y = U^T x$.

If there are a large number of observations of the process, then the first- and second-order moments can be estimated in the following manner

$$\alpha = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} x_i$$

$$M^2 = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} (x^i-\alpha)(x^i-\alpha)^T$$

We observe that the covariance matrix M^2 is a Gramian matrix formed by the infinite set $\{x^i-\alpha\}$.

4. The singular value decomposition

We have already indicated the form $x^T(M^{-2})x$ into its canonical form. A general quadratic form x^TPx where P is a symmetric matrix can be simplified using the spectral decomposition

$$P = UDU^T \quad , \quad U^TU = I \quad , \quad D = \text{diag}(d_1,\ldots,d_n)$$

Thus, $$x^TPx = (U^Tx)^TD(U^Tx) = \sum_{i=1}^{n} d_i y_i^2$$

where $y = U^Tx$.

If the rank of the matrix P is r, then n-r number of the diagonal values of D will be zero.

A natural extension of the above simplification is concerned with the bilinear form x^TQw where x and w are n and m dimensional vectors respectively and Q is an n,m dimensional matrix.

For the special case m = n, the problem was solved independently by three celebreties in the theory of matrices namely, Beltrami[3] (1873), Jordan[4] (1874), and Sylvester[2] (1889).

If the spectral decomposition of the Gramian matrices (which share the same eigenvalue set D^2) are written in the format

$$QQ^T = UD^2U^T \quad , \quad Q^TQ = VD^2V^T$$

$$U^TU = V^TV = I \quad , \quad D^2 = \text{diag}(d_1^2,\ldots,d_n^2)$$

then the singular value decomposition of Q is given by

$$Q = UDV^T$$

where the diagonal matrix D can be chosen to have diagonal positive values (called the singular values).

The bilinear form x^TQw in its canonical form is given by

$$x^TQw = (U^Tx)^TD(V^Tw) = \sum_{i=1}^{n} d_i y_i z_i$$

where $y = U^Tx$, $z = V^Tz$

However, the discovery of the singular value decomposition should be attributed to Jacobi[5], one of the founders of matrix theory. In the year 1832, that is exactly 150 years ago, Jacobi derived this decomposition for the special case n = 3, which was used in the simplification of a double integral via the canonical form described earlier. We observe that the generalization of the singular value decomposition from the 3-dimensional to the n-dimensional case is obvious to us although at the time of Jacobi higher dimensional spaces above 3 were not generally considered as physically meaningful.

The generalization of the decomposition for rectangular matrices is due to Eckart and Young[9] (1939). For the rectangular matrix θ of rank r, the decomposition is given by[11]

$$\theta = [U|\bar{U}] \left[\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array}\right] [V|\bar{V}]^T = UDV^T$$

where

$$[U|\bar{U}]^T[U|\bar{U}] = I_n \ , \ [V|\bar{V}]^T[V|\bar{V}] = I_m$$

$$U^TU = V^TV = I_r$$

$$\theta\theta^T = [U|\bar{U}] \left[\begin{array}{c|c} D^2 & 0 \\ \hline 0 & 0 \end{array}\right] |U|\bar{U}|^T = UD^2U^T$$

$$\theta^T\theta = [V|\bar{V}] \left[\begin{array}{c|c} D^2 & 0 \\ \hline 0 & 0 \end{array}\right] [V|\bar{V}]^T = VD^2V^T$$

where D is a diagonal r,r dimensional matrix.

5. Computation of Gramians and the singular value decomposition

As shown in the previous section, the information present in the Gramian matrices $\theta^T\theta$ and $\theta\theta^T$ are explicitly present in the decomposition $\theta = UDV^T$. Furthermore, the natural method of computation of the singular value decomposition is to form the Gramian matrices $\theta^T\theta$ and $\theta\theta^T$ and then to compute the spectral decompositions of those matrices. However, this approach has certain pitfalls if finite word length arithmetic is used in the computation as the following example indicates.

Let θ be the full rank matrix

$$\theta = \begin{pmatrix} 1.0 & 0 \\ 1.0 & \varepsilon \end{pmatrix}$$

where ε is a small real number such that the floating arithmetic unit can distinguish between 1.0 and ε. That is

$$1.0 \neq fl(1.0 + \varepsilon)$$

where $fl(\cdot)$ denotes floating point operations.

However, we assume that the floating point arithmetic is 'blind' to values of ε^2.

$$1.0 = fl(1.0 + \varepsilon^2)$$

The Gramian matrix $\theta\theta^T$ is of the form

$$\theta\theta^T = \begin{pmatrix} 1.0 & 1.0 \\ 1.0 & 1.0+\varepsilon^2 \end{pmatrix}$$

Thus, the Gramian $\theta\theta^T$ is internally represented in the computer as the rank one matrix

$$\begin{pmatrix} 1.0 & 1.0 \\ 1.0 & 1.0 \end{pmatrix}$$

if the above described floating point arithmetic unit is used, implying that the matrix θ is not full rank.

The above example clearly indicates that formation of Gramian matrices should be avoided in rank determination and in the computation of the singular value decomposition.

Fortunately, there is a computational scheme for the decomposition which does not require the formation of the Gramian matrices which is due to Golub et al[10] (1970) and which is based on an extended QR algorithm. Thus, the linear independence/dependence and the associated rank can be determined using the singular value decomposition without formation of the Gramian matrices.

6. The objectives and the organization of the thesis

The main aim of this thesis is to study the appearance of Gramian and canonical matrices in control and systems theory including pattern recognition and signal processing. More specifically, linearly dependent or 'almost' linearly dependent subspaces are removed so that more attention can be given to the more 'robust' linearly independent subspaces. Such removal of dependent data reduces the dimensionality and hence the complexity of such systems pertaining to that data. Thus data reduction is paramount in the analysis of large-scale systems.

The thesis is in seven parts and this introductory chapter forms Part 1.

Part 2 which consists of Chapters 2, 3 and 4 is about the Karhunen-Loeve expansion/transform, which is fundamental in the analysis of random processes. In Chapter 2, we study the

relationships between this expansion, the singular value decomposition and the technique of separation of variables. In Chapter 3, 2-dimensional data reduction is investigated in the context of the singular value decomposition. Finally, in Chapter 4, extrapolation or prediction of 'future' data using a random expansion which is structurally similar to the Karhunen-Loeve expansion, is described.

Part 3 is concerned with model-order reduction of linear dynamical systems which is increasingly becoming more important in the context of large-scale systems theory. The traditional methods of model-order reduction are based on modal methods where the slow-time behaviour and the fast-time behaviour as characterized by the 'small' and 'large' eigenvalues respectively, of the system matrix are the criteria for order reduction. The more modern approach is to delete the least controllable and the least observable parts of the system. This is achieved by means of 'balancing' transformations which transform the controllability and observability Gramians into their canonical diagonal forms. The theme in this part is to harmonize the singular perturbational approach with that of the balanced reduction. In Chapters 5 and 6, respectively, continuous-time and discrete-time systems are studied. In Chapter 7, the inter-relationships between continuous-time and discrete-time model-order reduction are investigated through the Cayley transformation. Finally in Chapter 8, the combined singular perturbational balanced method is exposed in the context of reciprocal transformations.

Part 4 of the thesis which encompasses Chapters 9, 10 and 11 describes a new Gramian matrix called the cross-Gramian which has properties of a cross-covariance matrix. This matrix denoted by W_{co} contains information pertaining to both controllability and observability of linear single-input single-output systems. In Chapter 9, this matrix is studied in relation to balanced and other principal representations of linear systems. The minimality or joint controllability/observability is the subject of Chapter 10. In Chapter 11, we demonstrate that the matrix W_{co} contains information about the Cauchy index of the system.

Part 5 is concerned with the quantification of input and output behaviour of linear systems based on properties of the Gramian matrix. In Chapter 12, the use of a Mahalanobis distance measure is discussed in relation to the degree of controllability. Due to the well known duality between controllability and observability, this naturally extends to observability as well. In Chapter 13, measures for describing inter-relationships between inputs and outputs are proposed. In Chapter 14, which is the final Chapter in this part, a method based on the Mahalanobis distance is described which can be used to discriminate system inputs or outputs.

The concluding Chapter 15 forms Part 6 of the thesis and Part 7 contains the appendices.

7. Originality of the research and style of presentation

The research work presented in this thesis is based on completely original work. All the Chapters and the four appendices are best described as 'essays' since they can be read and understood almost independently of other essays. This style is considered as the best way to present the broad range of research work undertaken in this thesis which ranges from control and systems theory to image processing, pattern recognition, time series analysis and signal processing.

The material in Chapters 2 to 6 and 9 to 13 and the Appendices 1 to 3 have been published or accepted for publication in the journals of the Institution of Electrical Engineering (IEE) London or the Institute of Electrical and Electronic Engineers (IEEE) New York. Chapters 7, 8, and 14 and Appendix 4 have been submitted for possible publication.

References

1. Gram, J.P.: 'Ueberdie Entwickelung reeller Functionen in Reihen mittelst der Methode der Kleinsten Quadrate', Jurnal reine angew. Mathematik, 1883, 94, pp.41-73.

2. Sylvester, J.J.: 'Sue la réduction biorthogonale d'une forme linéo-linéaire à sa forme canonique', Comptes Rendus Acad. Sci. Paris, 1889, 108, pp.651-653.

3. Beltrami, E.: 'Sulle funzioni bilineari', Giornale de Mat. Battaglini, 1873, 11, pp.98-106.

4. Jordan, C.: 'Mémoire sur les formes bilinéaires', Journal de Mathematiques Pures et Appliquees, 1874, series 2, 19, pp.35-54.

5. Jacobi, C.G.J.: 'de transformatione integralis duplicis indefiniti

$$\int \frac{\partial\phi\partial\psi}{A + B\cos\phi + C\sin\phi+(A'+B'\cos\phi+C'\sin\phi)\cos\psi+(A''+B''\cos\phi+C''\sin\phi)\sin\psi}$$

in formam simpliciorem $\int \frac{\partial\eta\partial\theta}{G-G'\cos\eta\cos\theta - G''\sin\eta\sin\theta}$,

Crelle's Journal (Jurnal reine angew Mathematik), 1832, 8, pp.253-279, and 321-357.

6. Mirsky, L.: 'An Introduction to Linear Algebra', Clarendon Press, Oxford, 1955.

7. Gantmacher, F.R.: 'The Theory of Matrices', Vol.1, Chelsea, New York, 1959.

8. Miller, K.S.: 'An Introduction to Vector Stochastic Processes', Krieger, New York, 1980.

9. Eckart, C., and Young, G.: 'A principal axis transformation for non-Hermitian matrices', Bull. Amer. Math. Soc., 1939, 45, 118-121.

10. Golub, G.H., and Reinsch, C.: 'Singular value decomposition and least-squares solutions', Numer. Math., 1970, 14, pp.403-420.

11. Klema, V.C., and Laub, A.J.: 'The singular value decomposition: Its computation and some applications', IEEE Trans. Automatic Control, 1980, AC-25, (2), pp.164-176.

PART 2

The Karhunen-Loève Expansion and Extensions

CHAPTER 2

The Karhunen-Loève expansion with reference to singular value decomposition and separation of variables

Abstract The Karhunen-Loève expansion for random processes, the method of principal component analysis, the singular value decomposition of rectangular matrices and the method of separation of variables used in mathematical physics and functional analysis, are shown to possess the same basic structure based on orthonormal basis functions or vectors and associated eigenproblems.

1. Introduction

The Karhunen-Loève expansion is one of the fundamental expansions used for describing random processes, and has been used widely in control, estimation and information theory and also in image processing and pattern recognition. The continuous expansion is based on orthogonal functions derived from eigenfunction solutions of covariance functions.

The continuous form of the expansion is well known and the associated optimal properties can be found in texts on probability and communication theory and pattern recognition[1-9]. However, it is difficult to obtain numerical solutions since it involves eigenfunction problems defined by Fredholm integrals.

The discrete form of the expansion leads to matrix eigenvalue problems which are well suited for digital computation[8-11,23]. The extension from the continuous to the discrete case has been

motivated by these numerical considerations and, unfortunately, the full algebraic properties of the expansion have not been fully investigated or utilized in the published literature. The discrete case is also equivalent to the method of principal component analysis used in mathematical statistics, which has had extensive applications in the social sciences[12-14,23].

The aim of this Chapter is to show that the discrete Karhunen-Loève expansion is algebraically equivalent to the singular value decomposition of a rectangular matrix[15,16]. For the continuous case, the expansion is equivalent to the classical technique of separation of variables using orthonormal basis functions (Bernoulli's separation method) which is a well recognized method of solution of partial differential equations in physics[17,24]. A more formal approach can be based on approximation theory, spectral theory and generalized functions in a Hilbert space setting.

2. The singular value decomposition of a rectangular matrix[15,16]

The singular value decomposition of an m,n dimensional matrix X is given by

$$X = UC^{\frac{1}{2}}V^T$$

or

$$x_{ij} = \sum_{k=1}^{r} \sqrt{c_k}\, u_{ik} v_{kj} \qquad (1)$$

where U and V are orthonormal matrices and C is a diagonal matrix. The rank of the matrix X is taken as r, and $X \in M_{m,n}$, $C = \mathrm{diag}(c_1 \ldots c_r)$, $c_k > 0$, $k = 1..r$, $U \in M_{m,r}$, $V \in M_{n,r}$, $r \leq \min(m,n)$. Also

$$U^T U = I \quad \text{or} \quad \sum_{k=1}^{m} u_{ki} u_{kj} = \delta_{ij} \tag{2}$$

$$V^T V = I \quad \text{or} \quad \sum_{k=1}^{n} v_{ki} v_{kj} = \delta_{ij} \tag{3}$$

where δ_{ij} denotes the Kronecker delta function.

The matrix U can be obtained as the eigenvector solution of the problem defined by

$$SU = UC \quad \text{or} \quad \sum_{k=1}^{m} s_{ik} u_{ki_1} = u_{ii_1} c_{i_1} \quad , \quad i_1 = 1..r \tag{4}$$

$$\text{where} \quad S = XX^T \quad \text{or} \quad s_{ik} = \sum_{\ell=1}^{m} x_{i\ell} x_{k\ell} \tag{5}$$

The matrix V is similarly given by

$$RV = VC \quad \text{or} \quad \sum_{k=1}^{n} r_{jk} v_{kj_1} = v_{jj_1} c_{j_1} \quad , \quad j_1 = 1..r \tag{6}$$

$$\text{where} \quad R = X^T X \quad \text{or} \quad r_{jk} = \sum_{\ell=1}^{n} x_{\ell j} x_{\ell j_1} \tag{7}$$

The nonnegative matrices S and R can also be written in the dyadic format

$$S = UCU^T \quad \text{or} \quad s_{ii_1} = \sum_{k=1}^{r} c_k u_{ki} u_{ki_1} \tag{8}$$

$$R = VCV^T \quad \text{or} \quad r_{jj_1} = \sum_{k=1}^{r} c_k v_{kj} v_{kj_1} \tag{9}$$

3. The separation of variables of a function of two variables[17,18,24,25]

The equivalent decomposition for a continuous function of two variables is given by the method of separation of variables. A function $x(w,t)$ of two independent variables w and t, can be represented by

$$x(w,t) = \sum_{k=1}^{r} c_k^{\frac{1}{2}} u_k(w) v_k(t) \quad , \quad w \in W \quad , \quad t \in T \qquad (1^*)$$

$r = \infty$, $c_k \geq 0$, $\sum_{k=1}^{r} c_k < \infty$, where the sets $\{u_k\}$ and $\{v_k\}$ are orthonormal functions with

$$\int_W u_k(w) u_\ell(w) dw = \delta_{k\ell} \qquad (2^*)$$

$$\int_T v_k(t) v_\ell(t) dt = \delta_{k\ell} \qquad (3^*)$$

The analogies between eqns (1), (2) and (1*), (2*), etc are obvious, where the summation has been replaced by integration.

The orthonormal functions $u_k(w)$ can be obtained from the eigenproblem defined by the Fredholm integral equation[18,25]

$$\int_W s(w,w_1) u_k(w_1) dw_1 = c_k u_k(w) \qquad (4^*)$$

where the kernal function $s(w,w_1)$ is given by

$$s(w,w_1) = \int_T x(w,t) x(w_1,t) dt \qquad (5^*)$$

Similarly, for $v_k(t)$

$$\int_T r(t,t_1) v_k(t_1) dt_1 = c_k v_k(t) \qquad (6^*)$$

where $$r(t,t_1) = \int_W x(w,t) x(w,t_1) dw \qquad (7^*)$$

By Mercer's theorem, the kernal functions $s(w,w_1)$ and $r(t,t_1)$ can be written in the form

$$s(w,w_1) = \sum_{k=1}^{r} c_k u_k(w) u_k(w_1) \qquad (8^*)$$

$$r(t,t_1) = \sum_{k=1}^{r} c_k v_k(t) v_k(t_1) \qquad (9^*)$$

4. Karhunen-Loève expansions The continuous form of the Karhunen-Loève expansion is given by eqn (1*) if it is assumed that w is the probability space variable of a second order random process. Integration with respect to the variable w in eqns (2*) and (7*) can then be replaced by the expectation operator $E[\cdot]$. Thus

$$E[u_k u_\ell] = \delta_{k\ell} \qquad (10^*)$$

$$r(t,t_1) = E[x(t)x(t_1)] \qquad (11^*)$$

and the kernal r can be identified as a covariance function.
(Note - It is usual to suppress the variable w from $x(w,t), u_k(w)$, etc).

Similarly, for the discrete case, the Karhunen-Loève expansion is given by eqn 1 with i denoting the discrete probability variable of a random process or of the 'experiments'. Equations 2 and 7 can then be modified as

$$E[u_k u_\ell] = \delta_{k\ell} \quad \text{or} \quad E[U^T U] = I \qquad (10)$$

$$r_{jj_1} = nE[x_j x_{j_1}] \quad \text{or} \quad R = E[X^T X] \qquad (11)$$

A dual set of results can be obtained by assuming t to be the probability space variable instead of w.

5. Conclusions The role of the singular value decomposition of rectangular matrices in random signal analysis and, particularly, the relation to the discrete Karhunen-Loève expansion has been highlighted. For the continuous case, the expansion is also equivalent to the method of separation of variables used in classical physics and functional analysis.

In most engineering and other problems involving numerical values and large samples, the expectation operator can be suppressed, and the Karhunen-Loève expansion and the singular value decomposition then become numerically identical.

In applications concerned, for example with the forecasting of load data[11], the signal matrix X is formed from load data at day (row) i and at hour (column) j. It is usual to consider that the rows of the matrix are due to different 'experiments', and to obtain the eigenvector matrix V which gives the modes of the system data in the row direction. This matrix will contain information pertaining to the variation of the load in a day due to industrial and domestic peaks, etc at each hour of the day. However, if the data in each column is taken as being due to different experiments, then the matrix U will show the weekly pattern, indicating the reduced demands at weekends. Hence, it is not always necessary to know which subscript refers to the probability space or the 'experiments' in practice.

The problem can also be modified to include basis functions $u_k(w)$ and $v_k(t)$ which are orthonormal with respect to weighting functions $p_{k\ell}(w)$ and $q_{k\ell}(t)$, respectively. Then

$$\int_W p_{k\ell}(w)u_k(w)u_\ell(w)dw = \delta_{k\ell}$$

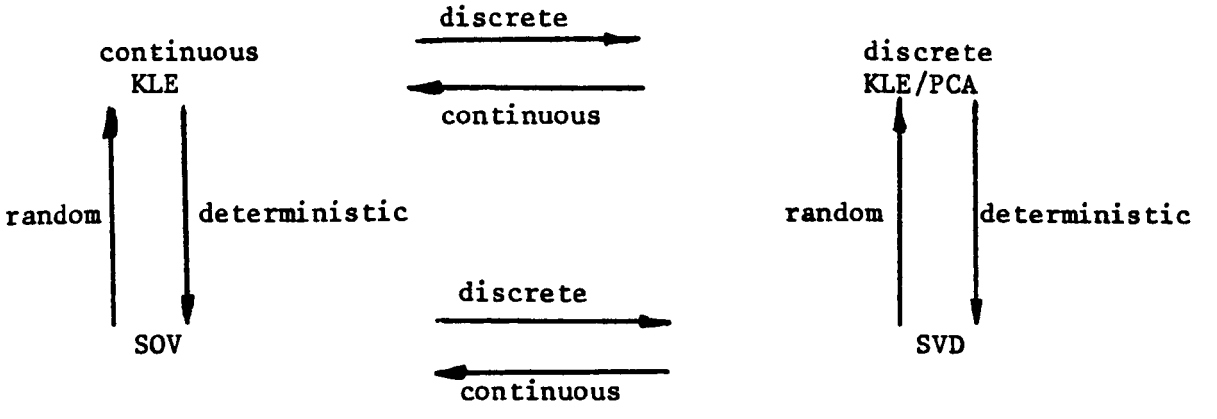
and

$$\int_T q_{k\ell}(t)v_k(t)v_\ell(t)dt = \delta_{k\ell}$$

If w is the probability space variable, then $p_{k\ell}(w)$ will correspond to the probability density. The results due to this extension are straightforward and hence omitted. For the discrete case, weighting matrices P and Q can be introduced accordingly.

The singular value decomposition has been used in model-order reduction of systems within the framework of principal component analysis[19,22] (see Part 3 and 4 for details).

The relationships between the various techniques are illustrated below. It is hoped that this exposition will further strengthen the links between different disciplines which can be studied within the general framework of systems theory.



KLE - Karhunen-Loève expansion
SOV - Separation of variables
SVD - Singular value decomposition
PCA - Principal component analysis

6. References

1. Miller, K.S.: 'Complex stochastic processes', Addison-Wesley, Reading, MA, 1974.

2. Wong, E.: 'Stochastic processes in information and dynamical systems', McGraw-Hill, New York, 1971.

3. Kailath, T.: 'A view of three decades of linear filtering theory', IEEE Trans. Information Theory, 1974, IT-20, 2, pp.146-181. Also in 'Linear least-squares estimation', (Ed. Kailath, T.) Benchmark papers in elct. eng. & comp. sci., 1977, vol.17, Dowden, Hutchinson & Ross, Strudsburg, Pennsylvania.

4. Kelly, E.J., and Root, W.L.: 'A representation of vector-valued random processes', J. of Mathematics & Physics, 1860, 39, pp.211-216.

5. Garcia, A.M.: 'On the smoothness of functions satisfying certain integral inequalities', in 'Functional analysis', (ed. Wilde, C.O.), Academic Press, 1970.

6. Davenport, W.B., and Root, W.L.: 'An introduction to the theory of random signals and noise', McGraw-Hill, New York, 1958.

7. Baggeroer, A.B.: 'State variables and communication theory', M.I.T. Press, Cambridge, MA, 1970.

8. Fukunga, K.: 'Introduction to statistical pattern recognition', Academic Press, New York, 1972.

9. Fu, K.S.: 'Sequential methods in pattern recognition and machine learning', Academic Press, New York, 1968.

10. Ahmed, N., and Rao, K.R.: 'Orthogonal transforms for signal processing', Springer-Verlag, Berlin, 1975.

11. Nicholson, H.: 'Structure of interconnected systems', Peter Peregrinus (for IEE), London, 1978.

12. Glaser, E.M., and Ruchkin, D.S.: 'Principles of neurobiological signal analysis', Academic Press, New York, 1976.

13. Kshirsagar, A.M.: 'Multivariate analysis', Marcel Dekker, New York, 1972.

14. Hotelling, H.: 'Analysis of a complex of statistical variables into principal components', J. Educational Psychology, 1933, 24, pp.417-441 and pp.498-520.

15. Forsyth, G., and Moler, C.B.: 'Computer solution of linear algebraic systems', Prentice-Hall, Englewood Cliffs, NJ, 1967.

16. Albert, A.: 'Regression and the Moore-Penrose pseudoinverse', Academic Press, New York, 1972.

17. Gould, S.H.: 'Variational methods for eigenvalue problems', Univ. of Toronto Press, Toronto, 1957.

18. Dieudonné, J.: 'Foundations of modern analysis', Academic Press, 1960.

19. Moore, B.C.: 'Principal component analysis in linear systems: controllability, observability, and model reduction', Univ. of Toronto, Dept. of Elect. Eng., Systems Control Reports 7801,7802, January 1980.

20. Andrews, H.C., and Hunt, B.R.: 'Digital image restoration', Prentice-Hall, Englewood Cliffs, NJ, 1977.

21. Brier, G.W., and Meltesen, G.T.: 'Eigenvector analysis for prediction of time series', J. Applied Meteorology, 1976, 15, 12, pp.1307-1312.

22. Moore, B.C.: 'Singular value analysis of linear systems', Proc. Conf. on Decision & Control, January 1979, San Diego, CA, USA, pp.66-73.

23. Ozeki, K.: 'A coordinate-free theory of eigenvalue analysis related to the method of principal components and the Karhunen-Loève expansion', Information and Control, 1979, 42, pp.38-59.

24. Sagan, H.: 'Boundary and eigenvalue problems in mathematical physics', Wiley, New York, 1961.

25. Hille, E.: 'Methods in classical and functional analysis', Addison-Wesley, Reading, MA, 1972.

CHAPTER 3

The discrete double-sided Karhunen-Loève expansion

Abstract. A new expansion for the representation of random data - the double-sided Karhunen-Loève expansion - is hypothesized, with application for data analysis, contraction and prediction in two-dimensional processes.

---

1. Introduction. The Karhunen-Loève (K-L) expansion is one of the basic forms used for describing random signals and has had wide application in pattern recognition, feature selection, image processing, data compression and prediction[1-7]. The expansion is formed using a set of orthonormal basis functions which can be obtained as a set of eigenvectors of a data covariance matrix, and optimal properties are associated with the expansion which is closely related to the least-squares estimation problem[8,23]. The truncated series minimizes the summated mean-square error and also the entropy function defined over the variance of the random coefficients of the expansion from the information theoretic point of view[2].

The pattern recognition or feature selection problem can be concerned with identifying the modes or the energy spectrum of the process. These properties can then be modified to emphasize or restrict certain aspects which may be required for data compression. Thus, an image can be enhanced by altering the corresponding energy values, and non-dominant terms of the expansion attributed to noise can be filtered or suppressed[9,10]. Data compression techniques also

have application in the analysis of biomedical data and in the coding and transmission of picture signals. Such data contraction can be considered as smoothing if the non-dominant neglected modes of the expansion are due to high-frequency components. From a statistical point of view and in geographical applications, this is also known as trend surface analysis.

Prediction or forecasting of non-stationary random processes which cannot be modelled exactly is essentially a problem of extrapolation of past data using known patterns. The K-L expansion has been used successfully for this problem, and particularly for the forecasting of electrical power and water system demands, and of air pollution and traffic flows[11-13,26].

A double-sided form of the K-L expansion is now developed for application within two-dimensional[15,25] space/time coordinate systems, and is related to the double-sided least-squares problem[14]. The double-sided form of the discrete K-L expansion is based on the singular value decomposition of matrices[16,21]. In numerical problems with large samples, the expectation operator can be suppressed and then the K-L expansion and the singular value decomposition technique are identical[27]. The decomposition has been used in the study of glass properties[28], meteorology[22], and image processing[29-31]. The expansion can be used for the analysis of spatially-correlated patterns, for example in geographically-located data, and for time-mapped data, which is becoming increasingly important in many fields of study including engineering, econometrics, ecology, meteorology, geology, planning and regional science.

2. The Karhunen-Loève expansion. The one-dimensional K-L expansion is concerned with the representation of mn-data points obtained from m experiments each with n observations. The data are ordered as an mxn-dimensional random signal matrix of the form

$$X = \begin{bmatrix} x_1(1) & \cdots & x_1(n) \\ \cdots & \cdots & \cdots \\ x_m(1) & \cdots & x_m(n) \end{bmatrix}$$

In the electrical load prediction problem, for example, the element $x_i(j)$ would represent the demand at time j hours (1..n) on day i (1..m).

The K-L expansion is now defined by the row-problem representation[8]

$$X = AV^T \quad , \quad X \in M_{m,n} \quad , \quad A \in M_{m,r} \quad , \quad V \in M_{n,r}$$

where A is a random coefficient matrix with expected value $E[A] = 0$.

Note - The notation $M_{m,n}$ denotes a real matrix of dimension mxn, etc.

The matrix V represents a set of basis functions and contains the orthonormalized n eigenvectors of the positive semi-definite covariance matrix R_1, defined as

$$R_1 = E[X^T P_1 X] \quad , \quad R_1 \in M_{n,n} \quad , \quad P_1 \in M_{m,m}$$

where P_1 is the apriori probability matrix associated with the m experiments, with elements p_{ij}, $1 \leq p_{ii} < 0$, $p_{ij} = 0$, $i \neq j$. In the load prediction problem, the probabilities would be assigned to each row depending on whether it is representative of demand for that particular day, which may include, for example, the effects of a freak weather condition.

The system modes are then identified with the eigenvalue problem defined by

$$R_1 V = V\Lambda_1 \quad , \quad \Lambda_1 \in M_{n,n}$$

where Λ_1 is the diagonal eigenvalue matrix which correlates the coefficient matrix A, with

$$R_1 = E[VA^T P_1 AV^T] \qquad \text{and} \quad A^T P_1 A = \Lambda_1$$

If the series is truncated, to include only the first k eigenvectors (with the eigenvalues ordered in decreasing order of magnitude), then

$$\tilde{X} = \bar{A}\bar{V}^T \quad , \quad \bar{A} \in M_{m,k} \quad , \quad \bar{V} \in M_{n,k}$$

where $\bar{A}$ denotes the truncated A-matrix with k columns, and $\tilde{X}$ $M_{m,n}$ is the reconstructed data X-matrix obtained using the truncated series. The expansion contains the first k modes, with

$$\tilde{R}_1 = E\left[\bar{V}\bar{\Lambda}_1\bar{V}^T\right] \quad , \quad \bar{\Lambda}_1 \in M_{k,k}$$

The error function is then given by

$$J = \text{trace } E\left[(X-\tilde{X})^T P_1 (X-\tilde{X})\right]$$
$$= \text{trace } \Lambda_1 - \text{trace } \bar{\Lambda}_1$$

representing the sum of the omitted eigenvalues.

A similar problem could also be considered with m observations resulting from n experiments, with priori probabilities assigned to each column. In the load prediction problem, the probabilities could emphasize the probable occurrence of demand at a particular hour of each day, say for example, during periods of peak TV viewing. In this case, the K-L expansion could be represented by the column-problem format

$$X = UB \quad , \quad U \in M_{m,r} \quad , \quad B \in M_{r,n}$$

where U contains the orthonormalized set of m vectors, or eigenvectors of the covariance matrix R_2, defined as

$$R_2 = E\left[XP_2X^T\right] \quad , \quad R_2 \in M_{m,m} \quad , \quad P_2 \in M_{n,n}$$

where P_2 is the apriori probability matrix associated with the n experiments. Then

$$R_2U = U\Lambda_2 \quad \text{and} \quad BP_2B^T = \Lambda_2$$

where $\Lambda_2 \in M_{m,m}$ is the diagonal eigenvalue matrix.

3. The double-sided K-L expansion. The double-sided K-L expansion can now be formulated to introduce the possibility of correlation between both the row and column data associated with either m or n experiments containing either n or m observations respectively. The expansion will define this dual behaviour inherently, based on the properties of the two covariance matrices R_1 and R_2 and the two spectra corresponding to row and column correlations respectively.

The expansion can be developed within a space(time)/space(time)- or space/time-coordinate framework. For example, in contrast to the time/time-coordinate load prediction problem, the measurement of river-water quality (such as biochemical oxygen demand (BOD), dissolved oxygen (DO), etc) could form a data matrix with spatial-coordinate rows and time-coordinate columns. For each independent physical cause, both spatial and temporal variations will be present and these will be characterized by an eigenvalue and corresponding row and column eigenvectors.

3.1 Probability and weighting matrices and 'energy'. For the case with a priori probability matrices $P_1 = I_m$, $P_2 = I_n$, representing absolute certainty of the experiments, the covariance matrices will be given by

$$R_1 = E[X^T X] = E[V\Lambda_1 V^T] \quad , \quad \Lambda_1 \in M_{n,n}$$
$$R_2 = E[XX^T] = E[U\Lambda_2 U^T] \quad , \quad \Lambda_2 \in M_{m,m}$$

The eigenvalue matrices Λ_1 and Λ_2 differ only by the number of zero eigenvalues. If r is the rank of the matrix X ($r \leq \min(m,n)$), then the ranks of R_1 and R_2 also will be equal to r, and

$$\Lambda_1 = \text{diag}\ (\lambda_1,\ldots,\lambda_r,0,\ldots 0)$$

$$\Lambda_2 = \text{diag}\ (\lambda_1,\ldots,\lambda_r,0,\ldots\ldots 0)$$

If the non-zero eigenvalues are contained in the diagonal matrix Λ defined as $\Lambda = \text{diag}\ (\lambda_1,\ldots,\lambda_r)$

then $\Lambda_1 = \Lambda \oplus 0_{n-r,n-r}$, $\Lambda_2 = \Lambda \oplus 0_{m-r,m-r}$

where $\oplus$ denotes the direct sum and $0_{n-r,n-r}$ is the null matrix of order n-r.

The traces of the covariance matrices are given by

$$\text{trace } R_1 = \text{trace } R_2 = E\left(\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}^2\right) = \text{trace } E[\Lambda]$$

This is equal to the sum of the eigenvalues and is a measure of the total 'energy' content of the system.

For the probability matrices P_1 and P_2 considered as general positive definite weighting matrices, we have

$$\text{trace } R_1 = E\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{m} p_{ik}(1)x_{ij}x_{kj}\right) = E[\text{trace } \Lambda_1] = E[\text{trace}(X^T P_1 X)]$$

$$\text{trace } R_2 = E\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{\ell=1}^{n} p_{j\ell}(2)x_{ij}x_{i\ell}\right) = E[\text{trace } \Lambda_2] = E[\text{trace}(XP_2 X^T)]$$

In general, the traces are not equal and can be considered as directional energies. The selection of the probability or weighting matrices P_1 and P_2 will be dictated by experience and the requirements of the problem. For example, in TV image processing, relative weighting could be used to increase the information content in the centre of the image compared to the edge regions. Also, row or horizontal scanning will preserve continuity in that direction and

reduce the correlation in the vertical direction due to the discrete nature of scanning and time delays, and the resulting effects could be de-emphasized by assigning appropriate weighting values to the matrices P_1 and P_2.

Note - If P_1 and P_2 are positive definite matrices, then rank R_1 = rank R_2 = rank X = r, and the number of non-zero eigenvalues will be equal to r, in each direction.

Since the above energies are directional, it may be possible to define a combined non-directional energy term. A possible scalar candidate function which is balanced in each direction, is given by

$$J_1 = E\left[P_1*(XP_2X^T)\right] = E\left[P_2*(X^TP_1X)\right]$$

$$= E\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{m}\sum_{\ell=1}^{n} p_{ik}(1)p_{j\ell}(2)x_{ij}x_{k\ell}\right)$$

where * denotes the matrix inner product or sum of inner products of corresponding rows or columns. From the two-dimensional point of view, any 'energy' maximization should then be with respect to J_1.

In the least-squares formulation, maximization of 'energy' with respect to J_1 corresponds to the minimization of 'energy' due to the error terms. Minimization could then be attempted using the non-directional function defined as

$$J_2 = E\left[P_1*(EP_2E^T)\right] = E\left[P_2*(E^TP_1E)\right]$$

where error matrix $E = X-\tilde{X}$.

3.2 The unweighted double-sided K-L expansion. The double-sided K-L expansion is now defined by

$$X = UCV^T = AV^T = UB \qquad (1)$$

or by the spectral decomposition

$$X = \sum_{i=1}^{r} c_{ii} u_i v_i^T$$

with dimensions $X \in M_{m,n}$, $U \in M_{m,r}$, $C \in M_{r,r}$, $V \in M_{n,r}$, where X is a data matrix of rank r and U, C and V are full rank (=r) matrices. Note - If the expansion $X = UCV^T$ is used for a time/space system, then the eigenvector matrix U will contain time-series information and the matrix V will relate to a space-series. Similarly, for purely spatial or temporal systems, the matrices U and V will be of the same kind. If the time variable is t and the space variable s, then the signal matrix can be written in the forms $U(t)CV(s)^T$, $U(t_1)CV(t_2)^T$ or $U(s_1)CV(s_2)^T$ for space/time, time/time or space/space systems respectively.

If U and V are the orthonormal eigenvector matrices, formed from the non-zero eigenvalues of the matrices XX^T, X^TX respectively, then the decomposition of eqn 1 becomes the singular value decomposition of the rectangular matrix X[16,21].

Now, since

$$R_1 = E[X^TX] = E[V\Lambda V^T] \quad , \quad \Lambda \in M_{r,r}$$

and $$R_2 = E[XX^T] = E[U\Lambda U^T] \quad ,$$

eqn 1 can be considered as the double-sided K-L expansion. It can also be shown that the matrix C is diagonal and equal to the square-root of the eigenvalue matrix. Thus

$$C^2 = \Lambda \quad , \quad C = \Lambda^{\frac{1}{2}} = C^T$$

The matrix C is also given by

$$C = U^TXV$$

Further, it can be shown that $A = UC$ and $B = CV^T$ are also K-L expansions.

If now the expansion is truncated to include only k ($\leq r$) modes, the matrix C can be solved in terms of the least-squares problem defined by[14] (see Appendix 1)

$$X = \bar{U}\bar{C}\bar{V}^T + E$$

where E is a residual error matrix. With an error function

$$J = \min E\left[\text{trace}(E^T E)\right]$$

the truncated solution for $\hat{\bar{C}}$ is then given by

$$\hat{\bar{C}} = \bar{U}^T X \bar{V} \quad , \quad \bar{U} \in M_{m,k} \quad , \quad \bar{C} \in M_{k,k} \quad , \quad \bar{V} \in M_{n,k} \quad , \quad k \leq r$$

assuming the orthogonality conditions $\bar{U}^T\bar{U} = I$, $\bar{V}^T\bar{V} = I$, and the reconstructed $\tilde{X}$ is given by

$$\tilde{X} = \bar{U}\hat{\bar{C}}\bar{V}^T = \bar{U}(\bar{U}^T X \bar{V})\bar{V}^T$$

Since C is a diagonal matrix, its least-squares solution is equivalent to truncation, ie, $\hat{\bar{C}} = \bar{C}$ and

$$\bar{A} = \bar{U}\bar{C} \quad , \quad \bar{B} = \bar{C}\bar{V}^T$$

The reconstructed covariance matrices are given by

$$\tilde{R}_1 = E\left[\tilde{X}^T\tilde{X}\right] = E\left[\bar{V}\bar{C}^T\bar{C}\bar{V}^T\right] = E\left[\bar{V}\bar{\Lambda}\bar{V}^T\right]$$
$$\tilde{R}_2 = E\left[\tilde{X}\tilde{X}^T\right] = E\left[\bar{U}\bar{C}\bar{C}^T\bar{U}^T\right] = E\left[\bar{U}\bar{\Lambda}\bar{U}^T\right]$$

which contain only the first k modes. The minimized error function is then given by

$$\text{trace } J = \text{trace } E\left[\Lambda - \bar{\Lambda}\right] \qquad (2)$$

which is equal to the sum of the omitted eigenvalues. Similarly, the maximized total energy is given by

$$\text{trace } \tilde{R}_1 = \text{trace } \tilde{R}_2 = \text{trace } E\left[\bar{\Lambda}\right].$$

3.3 The general double-sided K-L expansion. The least-squares problem with weighting matrices can be obtained by considering the truncated expansion

$$X = \bar{\psi}\bar{C}\bar{\phi}^T + E \quad , \quad \bar{\psi} \in M_{m,r} \quad , \quad \bar{\phi} \in M_{n,r} \quad , \quad \bar{C} \in M_{r,r}$$

where $\bar{\psi}$, $\bar{\phi}$ and $\bar{C}$ are full rank (=r) matrices.

With an error function

$$J = E\left[P_1*(EP_2E^T)\right] = E\left[P_2*(E^TP_1E)\right]$$

the least-squares estimate $\hat{\bar{C}}$ is then given by[14]

$$\hat{\bar{C}} = (\bar{\psi}^TP_1\bar{\psi})^{-1}\bar{\psi}^TP_1XP_2\bar{\phi}(\bar{\phi}^TP_2\bar{\phi})^{-1}$$

Without loss of generality we can assume that

$$\bar{H} = P_1^{\frac{1}{2}}\bar{\psi} \quad , \quad \bar{G} = P_2^{\frac{1}{2}}\bar{\phi} \quad , \quad \bar{H} \in M_{m,r} \quad , \quad \bar{G} \in M_{n,r}$$

and the matrices $\bar{H}$ and $\bar{G}$ are orthogonal, with

$$\bar{H}^T\bar{H} = \bar{G}^T\bar{G} = I_r$$

Then $$\hat{\bar{C}} = \bar{H}^TP_1^{\frac{1}{2}}XP_2^{\frac{1}{2}}\bar{G} = \bar{H}^TY\bar{G} \quad , \quad Y = P_1^{\frac{1}{2}}XP_2^{\frac{1}{2}}$$

and the reconstructed value of $\tilde{Y}$ is

$$\tilde{Y} = \bar{H}\hat{\bar{C}}\bar{G}^T$$

Also
$$\tilde{R}_1(\tilde{Y}) = E\left[\tilde{Y}^T\tilde{Y}\right] = E\left[\bar{G}\hat{\bar{C}}^T\hat{\bar{C}}\bar{G}^T\right]$$
$$\tilde{R}_2(\tilde{Y}) = E\left[\tilde{Y}\tilde{Y}^T\right] = E\left[\bar{H}\hat{\bar{C}}\hat{\bar{C}}^T\bar{H}^T\right]$$

Comparing these results with the un-weighted solution, the matrices $\bar{G}$ and $\bar{H}$ can now be identified as the eigenvector matrices of the covariance matrices $\tilde{R}_1(\tilde{Y})$ and $\tilde{R}_2(\tilde{Y})$ respectively. Then

$$\bar{C} = \bar{\Lambda}^{\frac{1}{2}}$$

and the reconstructed data matrix is given by

$$\tilde{X} = P_1^{-\frac{1}{2}} \bar{H} \bar{C} \bar{G}^T P_2^{-\frac{1}{2}}$$

The minimized value of the error function is given similarly by eqn 2, which is equal to the sum of the neglected eigenvalues associated with Y, and the maximized energy value is given by

$$\max E[P_1 * (\tilde{X} P_2 \tilde{X}^T)] = E[\text{trace } \bar{\Lambda}]$$

4. Computational procedure. The procedure for determining the reconstructed data matrix from the original matrix X is developed with the following steps.

(a) $Y = P_1^{\frac{1}{2}} X P_2^{\frac{1}{2}} \rightarrow R_1, R_2 \rightarrow H, G \rightarrow H^T Y G \ (=C)$

(b) choose k to truncate $H, G, C \rightarrow \bar{H}, \bar{G}, \bar{C}$

(c) $\tilde{Y} = \bar{H} \bar{C} \bar{G}^T$, $\tilde{X} = P_1^{-\frac{1}{2}} \tilde{Y} P_2^{-\frac{1}{2}}$

(d) trace $(E^T P_1 E)$, trace $(E P_2 E^T)$

(e) $J = P_1 * (E P_2 E^T) = P_2 * (E^T P_1 E) = \text{trace}(\Lambda - \bar{\Lambda})$

5. Example. A data matrix (Table 1) representing the number of passengers (x1000) carried on scheduled international airlines for monthly periods from 1949 to 1960[19,24] is used to illustrate the application of the double-sided K-L expansion technique for data contraction. Fig. 1 illustrates the hyper-surface generated by the two-dimensional process. The data is cyclic with summer peaks and has a rising pattern or trend as the popularity of air travel increased. In classical time series analysis, such behaviour has been explained using complicated and empirical composite models.

Table 1

X^d =

	JAN →											→ DEC
1949	112	118	132	129	121	135	148	148	136	119	104	118
↓	115	126	141	135	125	149	170	170	158	133	114	140
	145	150	178	163	172	178	199	199	184	162	146	166
	171	180	193	181	183	218	230	242	209	191	172	194
	196	196	236	235	229	243	264	272	237	211	180	201
	204	188	235	227	234	264	302	293	259	229	203	229
	242	233	267	269	270	315	364	347	312	274	237	278
	284	277	317	313	318	374	413	405	355	306	271	306
	315	301	356	348	355	422	465	467	404	347	305	336
↓	340	318	362	348	363	435	491	505	404	359	310	337
	360	342	406	396	420	472	548	559	463	407	362	405
1960	417	391	419	461	472	535	622	606	508	461	390	432

'Steady states' were removed using the method described in the appendix. The steady part X^s of the matrix X^d is found to be very dominant and it is given by

$$X^s = c_s \bar{a}\bar{b}^T$$

where

$\bar{a} = 10^{-1}(1.21\ 1.34\ 1.63\ 1.89\ 2.15\ 2.29\ 2.72\ 3.14\ 3.53\ 3.55\ 4.10\ 4.56)^T$

$\bar{b} = 10^{-1}(2.47\ 2.40\ 2.76\ 2.72\ 2.77\ 3.18\ 3.58\ 3.58\ 3.08\ 2.72\ 2.38\ 2.67)^T$

$c_s = 3.652\ 10^3$

The average trend could be identified from the vector $\bar{a}$ and the cyclic pattern from the vector $\bar{b}$.

The signal matrix X after the removal of steady states is shown in table 2.

The priori probability weighting matrices are selected as

P_1 = diag (0.45 0.50 0.55 0.60 0.65 0.70 0.75 0.80 0.85 0.90 0.95 1.00)

P_2 = diag (0.60 0.70 0.80 0.90 1.00 1.00 1.00 1.00 0.90 0.80 0.70 0.60)

with the elements of P_1 emphasizing the later data and the elements of P_2 providing increased weighting for the summer months.

The coefficient matrix is then given by

C = diag (68.47 36.65 -30.69 23.65 19.55 -11.19 -9.21 6.10 -4.49 -1.82
-0.30 0.00)

and the reconstructed data matrix based on k = 5 modes is illustrated in Table 3.

The optimized values of the error functions with the retention of different numbers of modes associated with the un-truncated signal and the corresponding maximized directional 'energy' functions are given in Table 4.

It can be seen from the reconstructed matrix that the best fits to the original matrix X are near the centre-bottom of the matrix corresponding to the summer of 1960. The overall contraction could be judged from the ratio defined by, error energy/un-truncated signal energy

$$= J/(P_2*(X^T P_1 X)) = \text{trace}(\Lambda - \bar{\Lambda})/\text{trace} = 3.31\ 10^{-2} \text{ for } k = 5$$

This ratio can be considered as the square of the noise/signal ratio and since it is very low, further truncation is possible within a slight increased penalty. The direction ratios are given by

$$\text{trace } E^T P_1 E/\text{trace } X^T P_1 X = 3.39\ 10^{-2}$$

$$\text{trace } E P_2 E^T/\text{trace } X P_2 X^T = 4.12\ 10^{-2}$$

and it is evident that the column direction is penalized more than the row direction.

Table 2

$X =$

2.8	11.8	9.9	8.3	-1.8	-5.8	-10.8	-10.7	-3.7	-1.5	-1.2	-0.3
-5.5	8.9	6.4	1.9	-10.4	-6.3	-5.1	-4.9	7.3	0.2	-2.0	9.5
-1.8	7.3	14.0	0.9	7.0	-11.2	-14.3	-14.1	0.4	0.2	4.6	7.0
1.1	14.8	3.1	-5.7	-8.1	-1.0	-15.9	-4.7	-3.5	3.6	8.4	10.0
1.9	7.4	19.1	20.6	10.8	-7.2	-18.0	-9.8	-5.8	-3.0	-5.9	-9.2
-2.1	-12.3	4.7	-0.7	2.3	-1.7	2.5	-6.3	1.2	1.8	4.5	5.8
-2.9	-5.1	-6.7	-1.6	-5.4	-0.8	8.0	-8.7	5.6	3.9	1.1	12.7
4.9	1.8	0.6	0.2	-0.3	9.0	1.5	-6.1	0.8	-6.2	-1.7	-0.6
-2.7	-7.9	0.9	-3.0	-2.3	12.4	3.2	5.5	6.5	-3.4	-1.0	-8.1
11.4	-1.4	-5.2	-15.0	-6.5	11.4	13.4	27.8	-7.1	-3.4	-5.5	-18.9
-5.4	-17.1	-6.9	-12.1	4.5	-4.3	11.1	22.5	0.9	-0.4	5.2	4.9
6.3	-8.2	-40.0	7.3	10.2	5.5	25.2	9.6	-5.7	8.1	-5.5	-12.8

Table 3

$\tilde{X} =$

2.3	11.2	9.3	7.4	-3.1	-3.6	-12.4	-11.0	-1.7	-0.4	-1.1	0.5
-1.8	6.3	5.8	-2.7	-6.2	-2.3	-7.2	-8.1	2.7	0.6	3.4	9.5
-3.9	6.1	13.7	5.8	3.1	-11.2	-15.0	-11.5	0.0	0.5	3.5	8.9
2.1	16.8	5.2	-5.2	-9.9	-5.4	-12.7	-4.7	-1.8	2.4	4.3	9.8
2.8	8.4	19.1	18.8	12.2	-7.3	-17.7	-10.9	-5.4	-3.6	-5.5	-10.4
-5.4	-9.1	3.0	4.9	2.6	-1.8	1.8	-4.2	4.9	-0.6	2.6	6.3
-4.5	-4.5	-8.3	-2.1	-3.9	1.0	5.2	-8.6	6.3	3.1	3.7	11.8
1.5	4.2	1.4	1.8	-3.2	7.2	1.5	-5.4	2.9	-2.3	-2.8	-2.8
-4.1	-7.8	3.1	-3.6	-2.9	14.5	5.5	4.4	4.0	-4.9	-2.9	-6.4
8.9	-0.9	-6.9	-14.2	-6.0	13.0	11.4	28.8	-6.0	-3.9	-5.2	-19.0
-7.7	-18.0	-6.4	-13.4	4.4	-4.9	11.7	21.7	0.4	1.0	5.2	4.1
6.8	-8.3	-39.5	7.4	9.9	5.1	25.8	9.5	-6.2	8.0	-5.1	-12.4

The reconstructed data process (including the steady states) is almost identical in form to the representation illustrated in Fig.1.

Table 4 - Optimized Values of Error Functions

k	trace $E^T P_1 E$	$J = P_1 * (E P_2 E^T)$	trace $E P_2 E^T$
12	0.	0.	0.
11	0.	0.	0.
10	$1.27\ 10^{-1}$	$8.96\ 10^{-2}$	$1.52\ 10^{-1}$
9	4.58	3.40	6.40
8	$3.03\ 10^{1}$	$2.36\ 10^{1}$	$3.69\ 10^{1}$
7	$7.66\ 10^{1}$	$6.07\ 10^{1}$	$8.95\ 10^{1}$
6	$1.81\ 10^{2}$	$1.45\ 10^{2}$	$2.04\ 10^{2}$
5	$3.23\ 10^{2}$	$2.71\ 10^{2}$	$4.25\ 10^{2}$
4	$7.78\ 10^{2}$	$6.53\ 10^{2}$	$9.51\ 10^{2}$
3	$1.41\ 10^{3}$	$1.21\ 10^{3}$	$1.72\ 10^{3}$
2	$2.75\ 10^{3}$	$2.15\ 10^{3}$	$2.95\ 10^{3}$
1	$4.24\ 10^{3}$	$3.50\ 10^{3}$	$4.43\ 10^{3}$

trace $X^T P_1 X = 9.52\ 10^3$ trace $X P_2 X^T = 1.03\ 10^4$

$P_1 * (X P_2 X^T) = P_2 * (X^T P_1 X) = 8.19\ 10^3 =$ trace Λ

6. Conclusion. A two-dimensional K-L-type expansion has been proposed assuming that there are two spectra or covariance matrices which can be associated with regression in horizontal (row) and vertical (column) directions. If the experiments in both directions are equally likely and certain, this expansion reduces to the singular value decomposition of a rectangular matrix.

The technique will have application for the analysis of temporal and spatially-located data which exist in many disciplines. It is also being developed and extended for two-dimensional curve fitting and prediction[17] (see Chapter 3), which will have application for the forecasting of a wide range of industrial and socio-economic system data. Extensions to multidimensional processes are also under consideration.

7. References

1. AHMED, N., and RAO, K.R.: 'Orthogonal transforms for digital signal processing', Springer-Verlag, Berlin, 1975.
2. FU, K.S.: 'Sequential methods in pattern recognition and machine learning', Academic Press, NY, 1968.
3. ANDREWS, H.C.: 'Computer techniques in image processing', Academic Press, NY, 1969.
4. ANDREWS, H.C.: 'Introduction to mathematical methods in pattern recognition', Wiley, NY, 1972.
5. ULLMANN, J.R.: 'Pattern recognition techniques', Butterworths, London, 1973.
6. MENDEL, J.M., and FU, K.S.: 'Adaptive, learning and pattern recognition systems', Academic Press, NY, 1970.
7. FUKUNGA, K.: 'An introduction to pattern recognition', Academic Press, NY, 1973.
8. NICHOLSON, H.: 'Sequential least-squares prediction based on spectral analysis', Int. Jl. Compr. Math., Sect. B, 1972, 3, pp.257-270.
9. ALBERT, G.E.: 'Statistical methods in prediction, filtering and detection problems', Jl. SIAM Appl. Math., 1960, 8, 4, pp.640-653.
10. SELIN, I.: 'The sequential estimation and detection of signals in normal noise, I', Inf. and Control, 1964, 7, pp.512-534.
11. MATTHEWMAN, P.D., and NICHOLSON, H.: 'Techniques for load prediction in the electricity-supply industry', Proc.IEE, 1968, 115, 10, pp.1451-1457.
12. STERLING, M.J.H., and ANTCLIFFE, D.J.: 'A technique for the prediction of water demand from past consumption data only', Jl. Inst. of Water Engineers, 1974, 28, 8, pp.413-420.

13. NICHOLSON, H., and SWANN, C.D.: 'The prediction of traffic flow volumes based on spectral analysis', Transpn. Res., 1974, 8, pp.533-538.

14. FERNANDO, K.V.M., and NICHOLSON, H.: 'The double-sided least-squares problem', Electronics Letters, 1979, 15, 20, pp.624-625.

15. WILLSKY, A.S.: 'Relationships between digital signal processing and control and estimation', Proc.IEEE, 1978, 66, 9, pp.996-1017.

16. GOLUB, G.H., and REINSCH, C.: 'Singular value decomposition and least squares solutions', Numer. Math., 1970, 14, pp.403-420. Also in Handbook for Automatic Computation: Linear algebra, Vol.2, pp.134-151, Edited by Wilkinson, J.H., and Reinsch, C., Springer-Verlag, Berlin, 1971.

17. FERNANDO, K.V.M., and NICHOLSON, H.: 'Two-dimensional curve fitting and prediction using spectral analysis', submitted for publication, Proc.IEE, 1982, 192D, in press.

18. MACDUFFEE, C.C.: 'The theory of matrices', Chelsea, NY, 1946.

19. BROWN, R.G.: 'Smoothing, forecasting and prediction', Prentice Hall, NJ, 1962.

20. CLIFF, A.D., and ORD, J.K.: 'Spatial autocorrelation', Pion Ltd, London, 1973.

21. ALBERT, A.: 'Regression and the Moore-Penrose pseudoinverse', Academic Press, NY, 1972.

22. BRIER, G.W., and MELTESEN, G.T.: Eigenvector analysis for prediction of time series', J. Applied Meterology, 1976, 15, 12, pp.1307-1312.

23. NICHOLSON, H.: 'Structure of interconnected systems', Peter Peregrinus, IEE, London, 1978.

24. BOX, G.E.P., and JENKINS, G.M.: 'Time series analysis, forecasting and control', Holden-Day, San Francisco, 1971.

25. LARIMORE, W.E.: 'Statistical inference on stationary random fields', Proc.IEEE, 1977, 65, 6, pp.961-970.

26. SAITO, O., and TAKEDA, H.: 'Two-stage predictor of air pollution adapted to daily fluctuation', Proc.IEE, 1979, 126, 1, pp.107-112.

27. FERNANDO, K.V.M., and NICHOLSON, H.: 'The Karhunen-Loève expansion with reference to singular value decomposition and separation of variables', to be submitted for publication.

28. PETERSON, G.E., CARNEVALE, A., and KURKJIAN, O.R.: 'Singular-value decomposition and boron spectra in glass', J. of non-crystalline solids, 1977, 23, pp.243-259.

29. SAWCHUK, A.A., and PEYROVIAN, M.J.: 'Restoration of astigmatism and curvature of field', J. Optical Soc. of Amer., 1975, 65, 6, pp.712-715.

30. SAHASRABUDHE, S.C., and KULKARNI, A.D.: 'Shift variant image degradation and restoration using SVD', Computer Graphics & Image Proessing, 1979, 9, pp.203-212.

31. ANDREWS, H.C., and PATTERSON, C.L.: 'Outer product expansions and their uses in digital image processing', Amer. Math. Monthly, 1975, 82, pp.1-13.

32. TOU, J.T., and GONZALEZ, R.C.: Pattern Recognition Principles , Addison-Wesley, Reading, MA, 1974.

33. LOÈVE, M.: Probability Theory, Van Nostrand, Princeton, NJ, 1963.

8. Appendix

Removal of the 'steady states'

In row formulation of the expansion, it is required that the expected value of the row sums to be zero.

$$\sum_{j=1}^{n} x_i(j) = a_i \quad , \qquad E[a_i] = 0 \quad \text{for all } i$$

Similarly, in the column representation of the expansion, the column sums have to be zero.

$$\sum_{i=1}^{m} x_i(j) = b_j \quad , \qquad E[b_j] = 0 \quad \text{for all } j$$

If the signal matrix has large deviations from the expected values, then removal of steady states or the mean levels would be required. From the computational point of view, this would be desirable to avoid large numerical values. Steady states removal is sometimes known as entropy reduction[32].

The following numerical scheme could be used to remove the steady state matrix X^s from the signal matrix X^d to give the matrix X with required properties.

$$X^d = X^s + X$$

The rank one matrix X^s is defined as

$$X^s = \frac{1}{d} ab^T \qquad \text{if } d \neq 0$$

where $a = (a_1, \ldots, a_m)^T$

$b = (b_i, \ldots, b_n)^T$

$$d = \sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$$

The matrix X^s can also be represented in the format

$$X^s = c_s \bar{a}\bar{b}^T = \lambda_s^{\frac{1}{2}} \bar{a}\bar{b}^T$$

where $\bar{a}$ and $\bar{b}$ are normalized vectors of a and b, respectively, and

$$c_s = \lambda_s^{\frac{1}{2}} = (a^T a)^{\frac{1}{2}} (b^T b)^{\frac{1}{2}}/d$$

Then λ_s could be interpreted as the 'energy value' associated with the steady states.

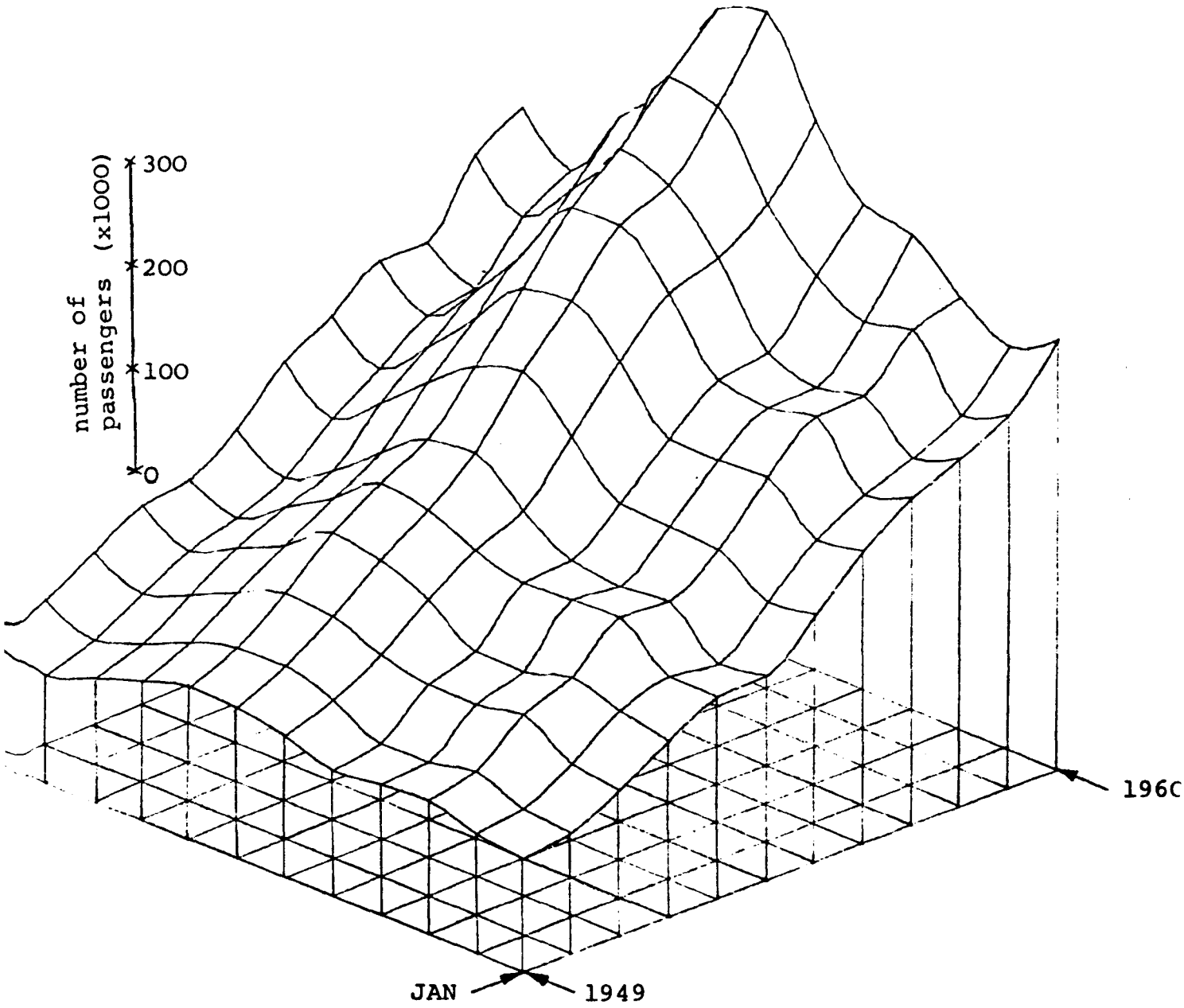


Fig. 1. Number of passengers carried on scheduled international airlines

CHAPTER 4

Two-dimensional curve-fitting and prediction using spectral analysis

Abstract: A curve-fitting or prediction problem for two-dimensional or cyclic processes is defined and solved using spectral techniques. The assumed statistical model is structurally similar to the Karhunen-Loève expansion and the technique can be implemented using the singular value decomposition. Two examples using published data illustrate the feasibility of the method and the peculiarities associated with the problem.

1. Introduction

One-dimensional techniques have often been used for the processing of data or signals which are essentially two-dimensional or cyclic. These include, for example, load demands on utilities, meteorological conditions, biological cycles, consumer demands and economic indicators. In one-dimensional methods, as in conventional ARMA time series analysis[17], which is generally not suitable for long-term prediction and for seasonal or cyclic processes[15], two-dimensional properties such as long-term trend and cyclic effects are removed or neglected, and thus such techniques do not utilize available information optimally. However, the two-dimensional and cyclic nature of processes has been appreciated recently[1-3], and is leading to new research activity in this area.

Two-dimensional properties of data can be exploited using the Karhunen-Loève expansion[20] (KLE), which is a fundamental expansion for random processes. This is equivalent to the well-known technique of principal component analysis[16], and at an abstract level to separation of variables methods in functional analysis and mathematical physics[4] (Chapter 2). It is also known as the method of characteristics[11].

Contraction and smoothing of data using the double-sided KLE based on the singular value decomposition[18] of a matrix have been studied previously[3,4] (Chapters 2,3). Contraction and smoothing is a problem of interpolation and in this particular case it is equivalent to least-square interpolation in tensor product spaces. We now extend this to extrapolation using past data which can be viewed as either prediction or curve-fitting depending on the type of data being handled.

The data or the signal matrix could consist of discretized values along temporal or spatial dimensions or a combination of the two. Causality

of data is not required or assumed and purely spatial two-dimensional processes can also be studied. If the data is cyclic then one of the dimensions could represent short-term behaviour within the cycle. The other dimension could then represent long-term aspects including any trends and other non-stationary effects.

The technique of extrapolation of data using KLE has been used in the forecasting of load demands on power systems[7-11] and water distribution networks[12], and of traffic flows[13], internal pressure during brain hemorrhage[11] and air pollution[14], among other applications. The object of this chapter is to develop a model to extrapolate data which can be considered as two-dimensional or cyclic. The structure of the model is similar to the KLE and can be implemented using the singular value decomposition technique.

2. The problem

It is assumed that the discretized data are in matrix form, with the two-dimensions of the data corresponding to the row and column directions. Without loss of generality, the submatrix X_{22} is taken as the unknown data to be predicted which is imbedded in the m,n dimensional composite matrix X. Thus,

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} \quad , \quad X_{ij} \in M_{m_i,n_j} \quad , \quad i,j = 1,2$$

$$m = m_1 + m_2 \quad , \qquad n = n_1 + n_2$$

The unknown matrix X_{22} could represent future data to be forecasted, as in load prediction problems, or it could correspond to inaccessible data in a hostile environment or to lost records due to instrumentation failure.

3. The Karhunen-Loève extrapolation method[7-14]

Forecasting of future data using the Karhunen-Loève expansion is implemented usually in two stages. The first stage involves the formation of a covariance matrix, the spectral decomposition of that matrix to obtain the "modes" or the eigenvector structure and then contraction of the data to eliminate the "noise" or the insignificant modes of the system. This can be considered as the training or learning stage of the method. The actual extrapolation using the modes or basis "pattern" functions is represented by the second stage.

We assume that the data matrix $Y_1 = [X_{11} \quad X_{12}]$ is formed using m_1 experiments each with n discrete sampled values. The data record denoted by the row vector y^i contains the information about the ith experiment and it is assumed that the expected value of y^i is zero, and that this second-order process has a covariance matrix R_y,

$$E[y^i] = 0 \quad , \qquad y^i \in M_{1,n} \quad , \qquad Y_1 \in M_{m_1,n}$$

$$E[(y^i)^T(y^i)] = R_y$$

where $E[\cdot]$ denotes the expectation operator.

If the covariance matrix R_y is already known (say, using analytical modelling or experimental evaluation) then we proceed to the decomposition of that matrix. However, if it is not known, then it has to be estimated using the information in the matrix Y_1. This can be achieved using Bayesian or any other standard learning method[20]. The most simple method is given by the asymptotic approximation,

$$R_y \approx \frac{1}{m_1} \sum_{i=1}^{m_1} (y^i)^T(y^i) = \frac{1}{m_1} Y_1{}^T Y_1$$

provided m_1 is large.

The spectral expansion of the covariance matrix R_y can be written in the spectral format,

$$R_y = V D_y^2 V^T \quad , \quad V , D_y^2 \in M_{n,n}$$

where the matrix V contains the n orthogonal set of eigenvectors and D_y^2 is the diagonal eigenvalue matrix ordered with decreasing magnitude. The KLE for the process can then be expressed as,

$$Y_1 = A_1 V^T \quad , \quad A_1 \in M_{m_1,n}$$

where A_1 is a random coefficient matrix, given by

$$A_1 = Y_1 V$$

The most important property of this matrix is that (the expected value of) the columns of the matrix are orthogonal, with

$$E\left[A_1^T A_1\right] = D_y^2$$

Because of this diagonal form, the modes of the process can be decoupled individually. All the optimal properties ascribed to this expansion are due to this absence of off-diagonal terms.

The diagonal values of the matrix D_y^2 may contain zero or near zero entries and they can be neglected without loss of information. If the expansion is truncated using only k modes, then it can be written in the format,

$$\bar{Y}_1 = \bar{A}_1 \bar{V}^T \quad , \quad \bar{Y}_1 \in M_{m_1,n} \quad , \quad \bar{A}_1 \in M_{m_1,k}$$

$$, \quad \bar{V} \in M_{n,k} \quad , \quad \bar{D}_y^2 \in M_{k,k}$$

where $\bar{Y}_1$ denotes the data matrix reconstructed from the most significant k modes and $\bar{A}_1$ and $\bar{V}$ are truncated matrices of A_1 and V, respectively, containing only the first k columns.

Prediction or extrapolation is based on the assumption that the remaining experiments as specified by the matrix $Y_2 = [X_{21} \quad X_{22}]$ has the same truncated eigenvector structure. The matrix Y_2 can then be expressed using the KLE, as

$$Y_2 = \bar{A}_2 \bar{V}^T + E_2 \quad , \quad Y_2, E_2 \in M_{m_2,n} \quad , \quad \bar{A}_2 \in M_{m_2,k}$$

where $\bar{A}_2$ is the truncated coefficient matrix and E_2 is a possible error matrix. The above expansion can be decomposed into two parts.

$$X_{21} = \bar{A}_2 \bar{V}_1^T + E_{21} \quad , \quad \bar{V}_1 \in M_{n_1,k} \quad , \quad E_{21} \in M_{m_2,n_1}$$

$$X_{22} = \bar{A}_2 \bar{V}_2^T + E_{21} \quad , \quad \bar{V}_2 \in M_{n_2,k} \quad , \quad E_{22} \in M_{m_2,n_2}$$

where $$V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \quad , \qquad E_2 = [E_{21} \quad E_{22}]$$

If n_1 is greater than the number of modes k, then the random coefficient matrix $\bar{A}_2$ can be obtained by solving the following least-squares problem.

$$X_{21} = \bar{A}_2 \bar{V}_1^T + E_{21} \quad ,$$

$$\underset{\hat{A}}{\text{minimize}} \text{ trace } (E_{21} E_{21}^T) .$$

The least-squares estimate $\hat{\bar{A}}_2$ is then given by,

$$\hat{\bar{A}}_2 = X_{21} \bar{V}_1 (\bar{V}_1^T \bar{V}_1)^{-1}$$

The estimate $\hat{X}_{22}$ of the unknown matrix X_{22} can be computed as,

$$\hat{X}_{22} = X_{21} \bar{V}_1 (\bar{V}_1^T \bar{V}_1)^{-1} \bar{V}_2^T$$

Since the extrapolation was performed with the matrix X_{21} in the horizontal or row direction using the eigenvector structure of the covariance matrix R_y, it can be called the horizontal or row prediction method (Fig.1).

The assumption that the data matrix is due to m experiments each with n sampled values is quite arbitrary in most applications, and it could be conveniently assumed that the data is from n experiments each with m sampled data values[3,4] (Chapters 2,3). A vertical or column predictor (Fig.2) can then be obtained by extrapolating the matrix X_{12} in the column direction by employing the eigenvector structure of the covariance matrix R_z,

$$R_z = UD_z^2U^T$$

and with analogous relationships as defined for the horizontal case.

If z^j is a column vector of Z_1, where,

$$Z_1 = \begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix}$$

then an estimate of R_z can be obtained by using the approximation,

$$R_z \approx \frac{1}{n_1} \sum_{j=1}^{n_1} (z^j)(z^j)^T = \frac{1}{n_1} Z_1 Z_1^T$$

4. A case for two-dimensional extrapolation

As mentioned in the introduction, one-dimensional techniques have often been used for the processing of data or signals which are essentially two-dimensional or approximately cyclic. However, often the solution is then a series of continuous (smooth) curves rather than a continuous surface, and such extrapolation methods do not utilize all available data optimally. Even if such methods provide very acceptable one-dimensional

fitting, it will be at the expense of the fitting in the complementary direction. Thus, true two-dimensional extrapolation has to be good in both directions and should not have a bias to one direction or the other (unless of course, such a bias is desired).

In practical problems with limited data points, the assumption that the process has a zero mean (i.e. $E[y^i] = 0$) is highly restrictive. Often, the experiments are not stationary and it is required to remove the trend or cyclic or other effects from the data before the asymptotic covariance matrix is formed. The conventional practice is to remove the trend effects by differencing of data. However, differencing is akin to differentiation and such techniques are anathema to most engineers. It is also possible to use polynomial curve-fitting to remove the trend or similar effects. Such methods artificially introduce a secondary curve-fitting process. One of our objectives is to avoid such secondary methods and to use the trend and other similar terms to our own advantage. This is achieved by considering the process to be two-dimensional and accepting such trend effects in the model.

One possible way of devising a two-dimensional extrapolation method is to exploit the eigenstructure of the matrix V as well as that of the matrix U. If both these matrices are used, most of the information in the data matrix is abstracted, and thus it is reasonable to assume that the prediction will be more consistent than if only one matrix is used.

Another objective is to avoid matrix inversion (or equivalently solution of linear matrix equations) by exploiting the orthogonal structure of the eigenvector matrices U and V. Such methods will also minimize the roundoff errors due to the reduced computational effort, and the technique will be numerically stable due to the inherently well-conditioned orthogonal matrices.

5. A two-dimensional statistical model

The element x_{ij} of the random matrix (field) is assumed to be of the form,

$$x_{ij} = u_{ik} d_k v_{jk}$$

where u_{ik} and v_{jk} are random variables with the following second order properties.

$$E\left[u_{ip} u_{iq}\right] = \delta_{pq} \quad , \quad E\left[v_{jp} v_{jq}\right] = \delta_{pq}$$

The function δ_{pq} is the Kronecker delta.

For an integer pair p,q, the sum of the product $x_{ip} x_{iq}$, $i = 1, m_1$ is given by

$$\sum_{i=1}^{m_1} x_{ip} x_{iq} = \sum_{r=1}^{k} \sum_{s=1}^{k} v_{pr} v_{qs} d_r d_q \sum_{i=1}^{m_1} u_{ir} u_{is}$$

If m_1 is large, we may assume the asymptotic approximation,

$$\frac{1}{m_1} \sum_{i=1}^{m_1} u_{ir} u_{is} \approx E\left[u_{ir} u_{is}\right] = \delta_{rs}$$

which gives the result,

$$\sum_{i=1}^{m_1} x_{ip} x_{iq} \approx m_1 \sum_{r=1}^{k} v_{pr} v_{qr} d_r^2$$

Equivalently, the above relation can be written in the matrix format,

$$Y_1^T Y_1 \approx m_1 \bar{\bar{V}} \bar{\bar{D}}^2 \bar{\bar{V}}^T \qquad (1)$$

Similarly, by considering the column direction, the following approximation can be obtained.

$$Z_1 Z_1^T \approx n_1 \bar{\bar{U}} \bar{\bar{D}}^2 \bar{\bar{U}}^T \qquad (2)$$

Equations (1) and (2) are very similar to the results obtained using the conventional Karhunen-Loève extrapolation method, except that there are two equations corresponding to both row and column extrapolation. To determine the "patterns" or the matrices $\bar{U}$ and $\bar{V}$, we again use asymptotic approximations.

If n is large, the following asymptotic approximation is valid,

$$\frac{1}{n} \sum_{i=1}^{n} v_{ir} v_{is} \approx E\left[v_{ir} v_{is}\right] = \delta_{rs}$$

which implies that

$$\bar{V}^T \bar{V} \approx I \qquad (3)$$

Similarly,

$$\bar{U}^T \bar{U} \approx I \qquad (4)$$

Thus, the matrices $\bar{U}$ and $\bar{V}$ are approximately orthonormal matrices. This property can be used to obtain estimates of these matrices by considering them to be the orthonormal eigenvector matrices of $Y_1^T Y_1$ and $Z_1 Z^T$, respectively. This is evident from equations (1) and (2).

In the next sections, extrapolation schemes to obtain the unknown matrix X_{22} are developed.

6. Extrapolation using the two-dimensional model

Since the "pattern" matrices $\bar{U}$ and $\bar{V}$ can be calculated using equations (3) and (4), the problem of extrapolation reduces to estimating the matrix $\bar{D}$. In the conventional approach, the actual extrapolation is achieved by considering the matrix X_{21} (the row method) or the matrix X_{12} (the column method). However, in the two-dimensional approach, we may use all three known matrices X_{12}, X_{21}, and X_{11}.

6.1 Extrapolation with three known matrices

Using the assumed two-dimensional model, the matrices Y_1 and Z_2 can be written in the format,

$$[X_{11} \quad X_{12}] = Y_1 = \bar{\bar{U}}\bar{\bar{D}}\bar{\bar{V}}_1^T + E_1$$

$$\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} = Z_1 = \bar{U}_1\bar{\bar{D}}\bar{\bar{V}}^T + E_2$$

where E_1 and E_2 are possible error matrices. Since, the matrices Y_1, Z_1, $\bar{U}$, $\bar{V}$, $\bar{U}_1$, and $\bar{V}_1$ are known, the diagonal matrix $\bar{D}$ can be estimated by using a least-squares technique.

If the quadratic error function for minimization is taken as,

$$\text{trace}(E_1E_1^T) + \text{trace}(E_2E_2^T)$$

the least-squares estimate of the diagonal matrix $\bar{D}$ is given by[6] (Appendix 2)

$$\hat{\bar{D}} = ((\bar{U}^T\bar{U}) * (\bar{V}_1^T\bar{V}_1) + (\bar{U}_1^T\bar{U}_1) * (\bar{V}^T\bar{V}))^{-1}c$$

where the elements of the vector c are given by

$$c_i = (u^i)^T Z_1 v_1^i + (u_1^i)^T Y_1 v^i$$

u^i and v^i are the column vectors of the matrices $\bar{U}$ and $\bar{V}$, respectively, and * denotes the Hadamard product or the Schur product defined as the element by element product of any two matrices of equal dimensions.

Using orthogonality conditions to simplify the solution, the elements of the matrix $\hat{\bar{D}}$ can be written in the form,

$$\hat{\bar{d}}_i = c_i/((v_1^i)^T v_1^i + (u_1^i)^T u_1^i)$$

Thus, inversion of matrices is not required in this method.

An estimate of the unknown matrix X_{22} can be computed as,

$$\hat{X}_{22} = \bar{U}_2 \hat{\bar{D}} \bar{V}_2^T$$

Since all three known matrices are used in the extrapolation, the estimate will be less sensitive to any unrepresentative values in the known matrices. In the conventional approach, only one known matrix (X_{12} or X_{21}) is used, and such values tend to have large effects on the predicted values. In the context of load prediction in the power and water industries, such unexpected demands occur due to freak weather conditions and bank holiday weekends.

6.2 Extrapolation using an arbitrary number of known matrices

We have so far assumed that the unknown block is the matrix X_{22} and that there are only three known blocks. However, such assumptions are not necessary and we can extend this to any arbitrary number of blocks, and the unknown block does not have to be the corner block.

The least-squares problem is then defined by

$$X_{k\ell} = \bar{U}_k \bar{\bar{D}} \bar{V}_\ell^T + E_{k\ell}$$

$$\text{minimize} \sum_k \sum_\ell \text{trace}\,(E_{k\ell} E_{k\ell}^T) \quad , \qquad k \, , \ell \in \text{known blocks}$$

Then the estimate $\bar{D}$ is given by

$$\hat{\bar{d}} = \left(\sum_k \sum_\ell (\bar{U}_k^T \bar{U}_k) * (\bar{V}_\ell^T \bar{V}_\ell) \right)^{-1} c$$

$$c_i = \sum_k \sum_\ell (u_k^i)^T X_{k\ell} v_\ell^i$$

The estimate of the unknown block X_{st} is given by

$$\hat{X}_{st} = \bar{U}_s \hat{\bar{D}} \bar{V}_t^T$$

By judicious choice of the known blocks for extrapolation, inversion of matrices may be avoided.

7. Formation of the data matrix for cyclic processes

For a cyclic process of period N (e.g. N = 12 months or N = 52 weeks if the cycle is taken as a year; N = 24 hours if the cycle is a day), the data matrix X can be formed in the following way[9].

$$X = \left(\begin{array}{ccc|ccc} x_1(1) & \dots & x_1(N') & x_1(N'+1) & \dots & x_1(N) \\ \dots & & \dots & \dots & & \dots \\ x_M(1) & \dots & x_M(N') & x_M(N'+1) & \dots & x_M(N) \\ \hline x_{M+1}(1) & \dots & x_{M+1}(N') & x_{M+1}(N'+1) & \dots & x_{M+1}(N) \end{array}\right) = \left(\begin{array}{c|c} X_{11} & X_{12} \\ \hline X_{21} & X_{22} \end{array}\right)$$

The suffix i in $x_i(j)$ refers to the cycle. It is assumed that data upto and including the Mth cycle and the data points 1 to N'+1 in the M+1 cycle are known. The unknown data to be predicted are the data points N'+1 to N in the M+1 cycle.

Using the above method, predictions can be made only of a fraction of a cycle. If however, full cycle ahead predictions are required, the following 'doubling up' procedure can be used.

$$X = \left(\begin{array}{ccc|ccc} x_1(1) & \dots & x_1(N) & x_2(1) & \dots & x_2(N) \\ \dots & & \dots & \dots & & \dots \\ x_{M-1}(1) & \dots & x_{M-1}(N) & x_M(1) & \dots & x_M(N) \\ \hline x_M(1) & \dots & x_M(N) & x_{M+1}(1) & \dots & x_{M+1}(N) \end{array}\right) = \left(\begin{array}{c|c} X_{11} & X_{12} \\ \hline X_{21} & X_{22} \end{array}\right)$$

This method has the advantage of having the continuity preserved from data points N to N+1 (e.g. December to January) in the row direction.

8. Computational procedure

Computation of covariance matrices $Y_1^T Y_1$ and $Z_1 Z_1^T$ can be avoided if the singular value decomposition of Y_1 and Z_1, respectively, are used in obtaining the spectral decompositions. The following procedure is used in implementing the prediction method.

(a) Compute the singular value decompositions[18]

$$Y_1 = U_y D_y V^T \quad , \quad Z_1 = U D_z V_z^T$$

(b) Truncate the eigenvector matrices U and V to give $\bar{U}$ and $\bar{V}$, respectively.

(c) Solve the least squares problems defined in section 6 to give the matrix $\bar{D}$.

(d) Compute the prediction submatrix $\hat{X}_{22} = \bar{U}_2 \hat{\bar{D}} \bar{V}_2^T$.

9. Example 1

To illustrate the method, one cycle ahead predictions equivalent to twelve months were attempted for data corresponding to the number of international airline passengers per month (entering and leaving the United States) in the years 1949 to 1960. This airline data is seasonal or cyclic with a trend and has been widely analysed in the literature[3,17].

Figure 3 shows the actual passenger levels and the predicted values. The predictions were obtained using five years of immediate past data ($M = 5$) and two modes ($k = 2$). We arrived at the following conclusions after extensive numerical experiments with this data.

(a) Reasonable one-year ahead predictions can be obtained using this method as is obvious from Fig. 3. Efficiency of the method was gauged using the mean error and the mean squared error for each year.

(b) It was found that if more than five years of past data were used, the errors marginally increased. This may be due to long-term nonstationary effects which cannot be taken into account. Similarly, if M was taken as less than five, predictions were poor which may be due to insufficient statistics.

(c) Generally, acceptable predictions were obtained using only one mode (k = 1). For k = 2, corresponding to two modes, slightly better (which might not be statistically significant) predictions can be made. It was found that the ratio of 'energy' in the first mode to the second mode $\hat{d}_1^2/\hat{d}_2^2$ is high, and this could explain the marginal differences between the predictions using one and two modes.

(d) The elements of the matrices $Y_1^T Y_1$ and $Z_1 Z_1^T$ are positive and thus, the first modes v_1 and u_1 correspond to Perron-Frobenius eigenvectors which have positive elements. The vector v_1 contained most of the seasonal information and the vector u_1, the trend. Due to orthogonality conditions, u_2, v_2 and other higher-order modes, contain both positive and negative elements and thus can be considered as completely oscillatory modes.

(e) Since only past data is used, it is not possible to take into account other factors which affect the number of passengers such as the high number of airline accidents[19] in the United States in 1958. This might perhaps explain the large prediction errors for that year.

10. Example 2

As our second example, we have chosen the power demand on the Hydro-Quebec system from Monday 15 November to Sunday 21 November 1971 as published by Srinivasan and Pronovost[21]. We have attempted 12 hour predictions at hourly intervals at mid-day and mid-night. Since the

data is approximately cyclic with a period of 24 hours, the doubling up procedure detailed in section 7 is not required for 12 hour predictions.

Figure 4 shows the actual realized load levels and the predictions using four days of past data (M = 4) and two modes (k = 2).

It can be seen from Figure 4 that the prediction errors are within reasonable limits except in the periods 1 to 12 hours on the Saturday and the Sunday. This is due to the different patterns of load which exist in weekdays and weekends. Since, the "patterns" or the "modes" for prediction are based on the weekdays, such errors are to be expected. This indicates that power load data has three-dimensional properties rather than two-dimensional with hourly, daily, and weekly correlations, which has been well-recognized by power engineers for some time[21], although the significance of higher-dimensional data have not been fully appreciated yet in the literature[1-2].

11. Conclusions

A curve-fitting or prediction problem for two-dimensional or cyclic processes has been formulated and solved. The assumed statistical model is structurally similar to the Karhunen-Loève expansion, and is implemented using the singular value decomposition.

The method can be used for temporal, spatial or mixed data in engineering, socio-economics, bio-medicine and other fields for general two-dimensional curve-fitting and long-term prediction. A typical application of this method could be in prediction of load demands in power and other utilities. In such problems, enough data is usually available for formation of the asymptotic covariances and the data is cyclic. Furthermore, it is reasonable to assume that such systems are excited by noise rather than by deterministic inputs which conforms to our strictly statistical model.

11. References

1. Willsky, A.S.: 'Digital signal processing and control and estimation theory - points of tangency, areas of intersection. and parallel directions', M.I.T. Press, Cambridge, MA, 1979.
2. Strintzis, M.G.: 'Dynamic representation and recursive estimation of cyclic and two-dimensional processes', IEEE Trans. Automatic Control, 1978, AC-23, pp.801-809.
3. Fernando, K.V.M., and Nicholson, H.: 'Discrete double-sided Karhunen-Loève expansion', Proc.IEE, part D, 1980, 127, pp.155-160.
4. Fernando, K.V.M., and Nicholson, H.: 'Karhunen-Loève expansion with reference to singular-value decomposition and separation of variables', Proc.IEE, part D, 1980, 127, pp.204-206.
5. Fernando, K.V.M., and Nicholson, H.: 'Double-sided least-squares problem', Electronics Letters, 1979, 15, (20), pp.624-625.
6. Fernando, K.V.M., and Nicholson, H.: 'Double-sided least-squares problem with diagonal constraints', Electronics Letters, 1980, 16, (3), pp.82-83.
7. Farmer, E.D., and Potten, M.J.: 'Development of online load prediction techniques with results from trials in the south-west region of the CEGB', Proc.IEE, 1968, 115, pp.1549-1558.
8. Matthewman, P.D., and Nicholson, H.: 'Techniques for load prediction in the electricity supply industry', Proc.IEE, 1968, 115, pp.1451-1457.
9. Nicholson, H.: 'Structure of interconnected systems', Peter Peregrinus, London, 1978.

10. Belik, D.D., Nelson, D.J., and Olive, D.W.: 'Use of Karhunen-Loève expansion to analyse hourly load requirements for a power utility', Proc.IEEE PES Winter Meeting, New York, January 1978.

11. Ivakhnenko, A.G., and Lapa, V.G.: 'Cybernetics and forecasting techniques', Elsevier, New York, 1967.

12, Sterling, M.J.H., and Antcliffe, D.J.: 'A technique for the prediction of water demand from past consumption data only', Jl. Inst. of Water Engineers, 1974, 28, pp.413-420.

13. Nicholson, H., and Swann, C.D.: 'The prediction of traffic flow volumes based on spectral analysis', Transpn. Res., 1974, 8, pp.533-538.

14. Saito, O., and Takeda, H.: 'Two-stage predictor of air pollution adapted to daily fluctuation', Proc.IEE, 1979, 126, pp.107-112.

15. Chatfield, C., and Prothero, D.L.: 'Box-Jenkins seasonal forecasting: problems in a case-study', J. R. Statist. Soc. A, 1973, 136, pp.295-336.

16. Brillinger, D.R.: 'Time series: data analysis and theory', Holt, Reinhart and Winston, New York, 1975.

17. Box, G.E.P., and Jenkins, G.M.: 'Time series analysis, forecasting and control', Holden-Day, San Francisco, 1970.

18. Garbow, B.S., Boyle, J.M., Dongarra, J.J., and Moler, C.B.: 'Matrix eigensystem routines: EISPACK guide extensions', Lecture notes in computer science, vol-51, Springer-Verlag, Berlin, 1977.

19. Lombardo, T.G.: 'The federal aviation administration under scrutiny', IEEE Spectrum, 1980, 17, (11), pp.53-56.

20. Fu, K.S.: 'Sequential methods in pattern recognition and machine learning', Academic Press, New York, 1968.

21. Srinivasan, K., and Pronovost, R.: 'Short term load forecasting using multiple correlation models', IEEE Trans. Power Apparatus and Systems, 1976, PAS-94, (5), pp.1854-1858.

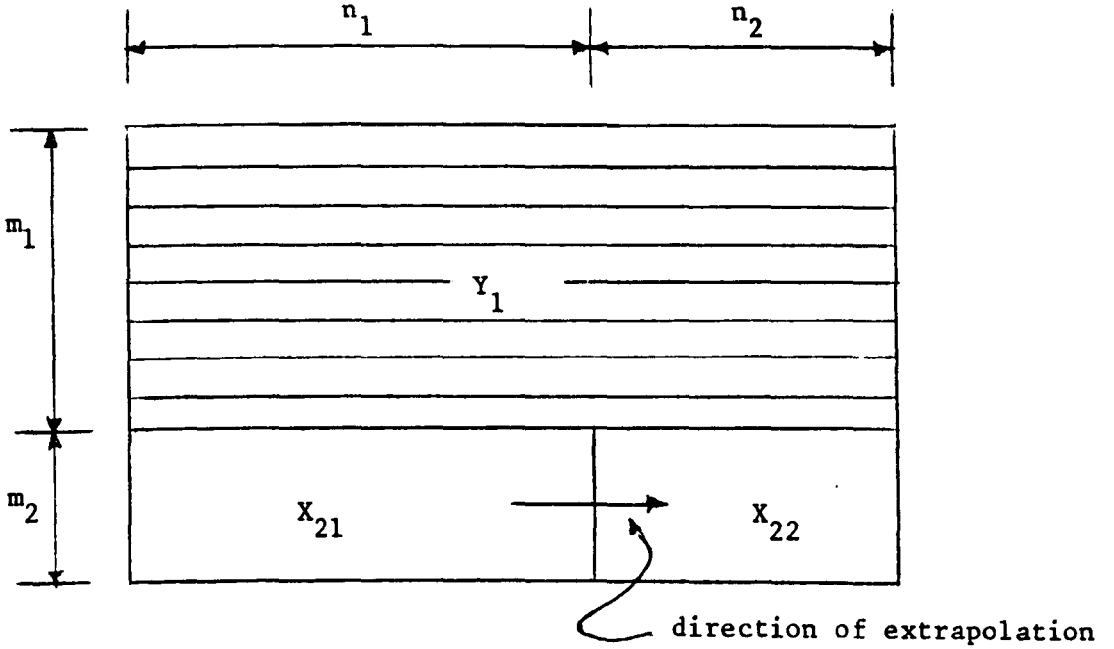


Fig. 1: Conventional horizontal (row) extrapolation

Hatched area indicates the data required to form the covariance matrix

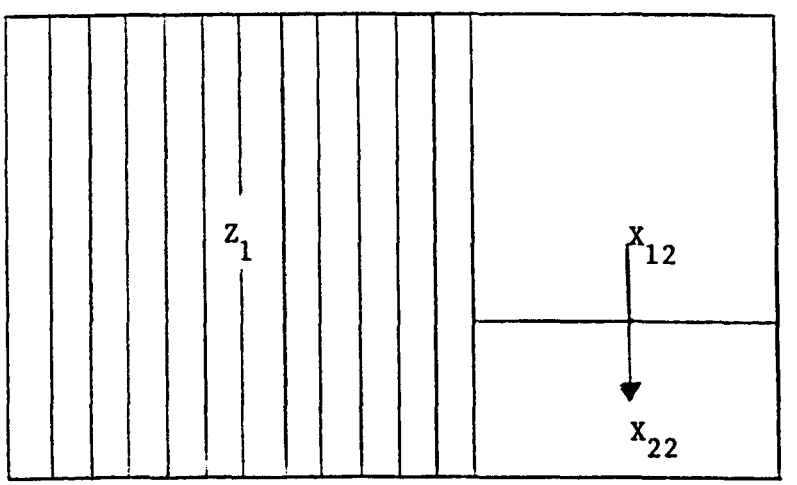


Fig. 2: Conventional vertical (column) extrapolation

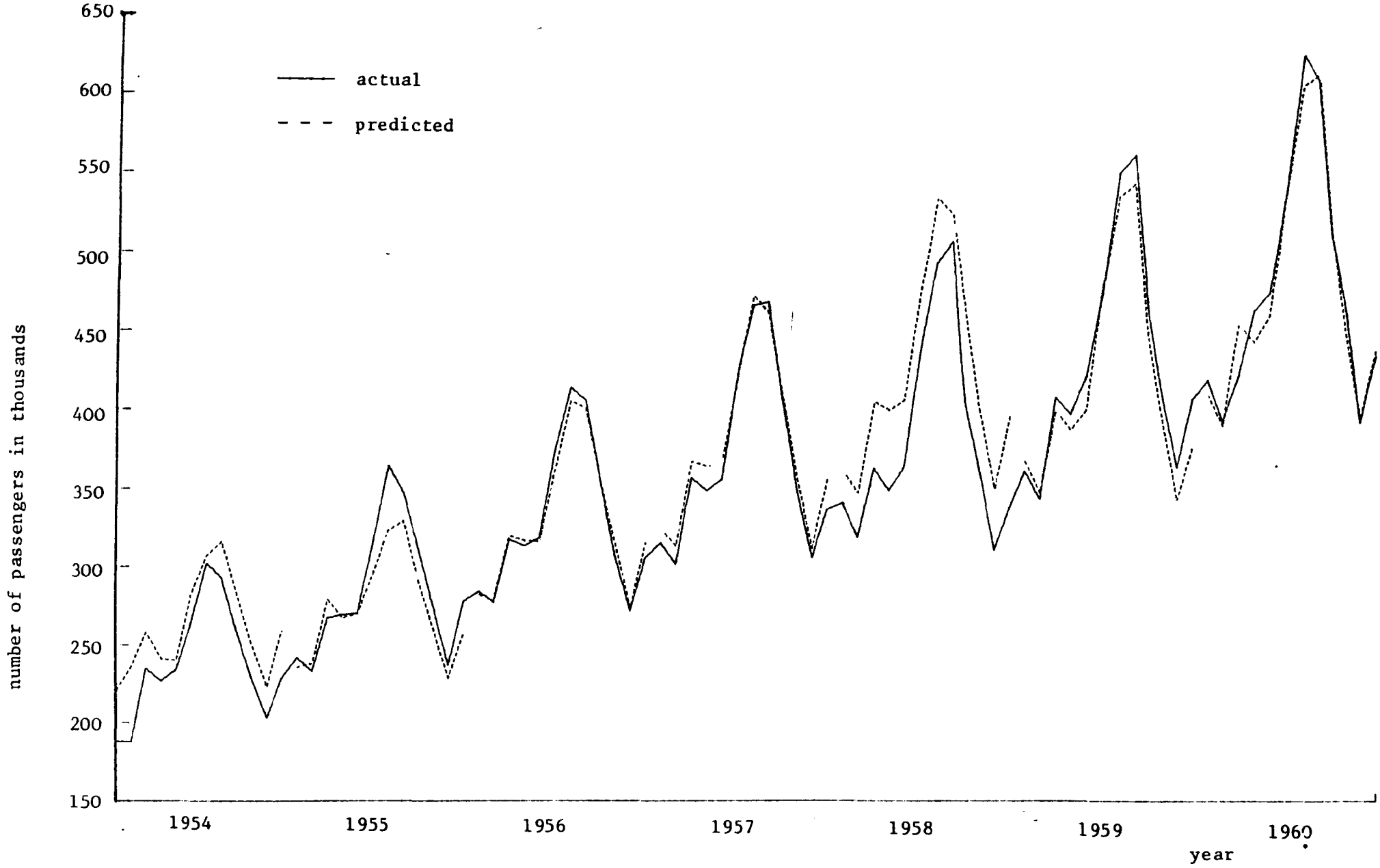


Fig. 3: The international passengers leaving and entering the USA per month

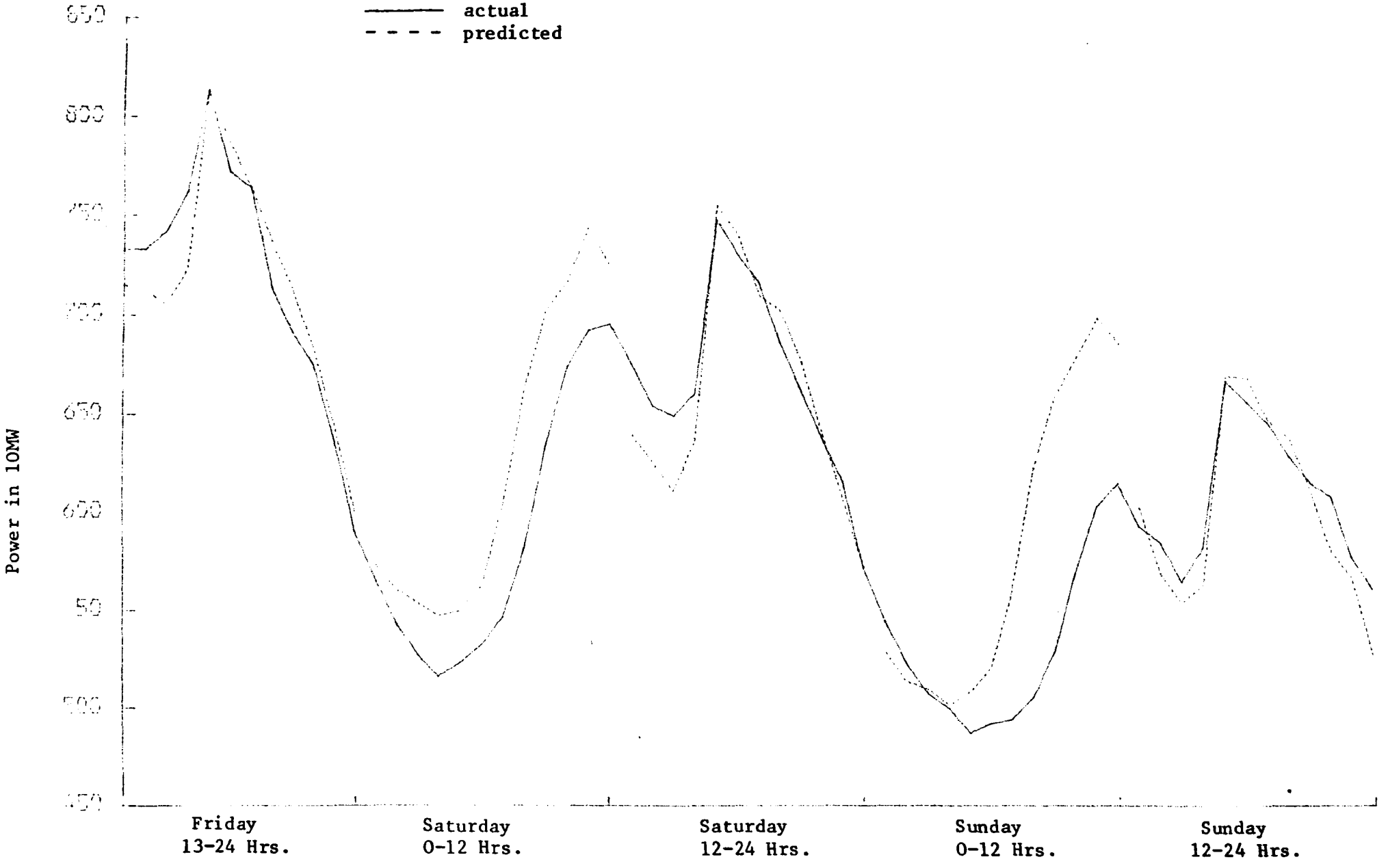


Fig. 4. Load on the Hydro-Quebec System

PART 3

Singular Perturbational Model Order Reduction of Balanced Systems

CHAPTER 5

Singular Perturbational Model Reduction of Continuous-time Balanced Systems

Abstract: The balanced representations of linear systems due to Moore are shown to be natural and convenient media for singular perturbational model reduction.

1. Introduction

The linear stable time-invariant system S(A,B,C) defined by

$$\dot{x} = Ax + Bu \quad , \quad y = Cx$$

$$x \in M_{n,1} \quad , \quad u \in M_{m,1} \quad , \quad y \in M_{r,1}$$

$$A \in M_{n,n} \quad , \quad B \in M_{n,m} \quad , \quad C \in M_{r,n}$$

is an input-output balanced system, if the controllability Gramian W_c and the observability Gramian W_o are diagonal and equal[6-8]. In this case

$$W_c = W_o \triangleq W \quad , \quad W_c, W_o W \; M_{n,n}$$

where $$W_c = \int_o^\infty e^{At} B B^T e^{A^T t} dt$$

$$W_o = \int_o^\infty e^{A^T t} C^T C e^{At} dt$$

and A,B,C represent balanced system matrices.

A linear system can be balanced using similarity transformations[6] which is the basic operation used in principal component analysis (and is also equivalent to the Karhunen-Loève expansion/transform method[4] (see Chapter 2)) of linear systems[6-8].

In the balanced system approach, the reduced order model is obtained by direct elimination of 'weak' subsystems whose contribution to the impulse response of the system is negligible. However, in the perturbational approach, a 'boundary layer' correction is used to account for the eliminated subsystem[1,2,12]. Our aim is to show that such boundary layer corrections can be accommodated in balanced models and such models are natural representations for singular perturbational reduction.

2. Model Reduction using Balanced Systems

If the linear system S(A,B,C) is balanced[6-8], then the diagonal Gramian matrix W satisfies the Lyapunov equations,

$$WA^T + AW = -BB^T \qquad (1)$$

$$WA + A^TW = -C^TC \qquad (2)$$

Model reduction is achieved by considering a partition of the balanced system matrices of the form

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right), \quad B = \left(\begin{array}{c} B_1 \\ \hline B_2 \end{array}\right), \quad C = \left[\begin{array}{c|c} C_1 & C_2 \end{array}\right]$$

where the matrices A_{11} and A_{22} are square. The two subsystems $S(A_{ii},B_i,C_i)$ are also input-output balanced and the Gramians for the subsystems are given by

$$W_i A_{ii}^T + A_{ii} W_i = - B_i B_i^T$$

$$W_i A_{ii} + A_{ii}^T W_i = - C_i^T C_i \qquad , \qquad i = 1,2$$

where $$W = \left(\begin{array}{c|c} W_1 & 0 \\ \hline 0 & W_2 \end{array} \right)$$

The diagonal Gramians also satisfy the following equations which relate the cross-coupling between the subsystems.

$$W_i A_{ji}^T + A_{ij} W_j = - B_i B_j^T$$

$$W_i A_{ij} + A_{ji}^T W_j = - C_i^T C_j \qquad , \qquad i,j = 1,2$$

If $S(A_{22},B_2,C_2)$ is a weak subsystem, then the diagonal elements of the W_2 Gramian will be small in comparison with that of W_1. By eliminating the weaker subsystem $S(A_{22},B_2,C_2)$, we obtain the reduced order model $S(A_{11},B_1,C_1)$.

3. Model Reduction using the Singular Perturbation Method

In the singular perturbation method[1,2], the linear system defined by

$$\begin{bmatrix} \dot{x}_1 \\ \mu\dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \quad , \quad y = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is approximated by

$$\dot{\bar{x}} = \bar{A}\bar{x} + \bar{B}u \qquad , \qquad \bar{y} = \bar{C}\bar{x}$$

where $$\bar{A} = A_{11} - A_{12}A_{22}^{-1}A_{21}$$

$$\bar{B} = B_1 - A_{12}A_{22}^{-1}B_2$$

$$\bar{C} = C_1 - C_2A_{22}^{-1}A_{21}$$

The reduced-order state matrix $\bar{A}$ is the Schur complement of the matrix A. The Schur complement of a partitioned matrix is fundamental in

model reduction problems and also in data contraction[10], the Kron method of tearing[3], and elsewhere[7].

The zeroth order perturbational approximation is exact for $\mu = 0$, and frequently used[12,13] for the case $\mu = 1$. However, the perturbational parameter μ does not explicitly appear in the approximation and such simplifications can be considered as approximate aggregation[11].

4. Weak and Fast Subsystems

The diagonal values of the state matrix A are related directly to the elements of the Gramian matrices. Thus,

$$2w_{ii}a_{ii} = -\sum_{j=1}^{m} b_{ij}^2$$

where the lower case letters indicate the elements of the respective matrices. The relatively small elements w_{ii} in weak systems can occur in two ways:

(a) with large values of a_{ii} which will correspond to high damping (real parts of the eigenvalues) which is a property of fast systems,

(b) with small values of b_{ij} which means that the states x_i are not strongly excited by impulses.

We observe that not all systems are suitable for singular perturbational reduction particularly if mechanism (b) dominates. However, if the eigenvalues cluster into disjoint regions, singular perturbational reduction is possible and for such systems, mechanism (a) will prevail. Thus, fast subsystems and weak subsystems have common attributes in certain classes of systems.

5. Model Reduction of Balanced Systems using the Singular Perturbation Method

The principal component approach is elegant from the point of view of minimal realization, although, as pointed out by Moore[6], model reduction through subsystem elimination is not a well understood operation. Model reduction using singular perturbational methods has achieved a certain amount of maturity, and thus, it is natural to investigate whether the singular perturbational approach is compatible with the balanced representation. The following proposition answers this affirmatively.

Proposition: If the system $S(A,B,C)$ is balanced, then the singular perturbational approximation $S(\bar{A},\bar{B},\bar{C})$ also defines a balanced system. Further, the diagonal Gramian for the reduced-order system is given by the matrix W_1. Thus, the Lyapunov equations for the reduced-order system are of the form,

$$W_1\bar{A}^T + \bar{A}W_1 = -\bar{B}\bar{B}^T \qquad (3)$$

$$W_1\bar{A} + \bar{A}^TW_1 = -\bar{C}^T\bar{C} \qquad (4)$$

Proof: The Lyapunov equation (3) can be obtained by premultiplying and postmultiplying (1) by T_1 and T_1^T, respectively, where

$$T_1 = \left[\, I \;\vdots\; -A_{12}A_{22}^{-1} \,\right]$$

Similarly, (4) can be obtained by premultiplying and postmultiplying (2) by T_2 and T_2^T, respectively, where

$$T_2 = \left[\, I \;\vdots\; -A_{21}^T(A_{22}^T)^{-1} \,\right] \quad , \quad T_1 \, , \, T_2 \in M_{r,n}$$

Since the Gramian matrix W_1 is diagonal, the reduced-order system is also balanced and this completes the proof.

In the subsystem elimination method using balanced systems, the retention of the dominant second-order modes (i.e. the diagonal elements of the Gramian) is considered as a criterion for model reduction. Since we arrive at the same Gramian matrix W_1, this condition is naturally satisfied in the perturbational approach. For single-input single-output systems, a shared Gramian W_1 also signifies the fact that the two systems have a common first moment (i.e. the dc gain). This can be proved using results in reference 7.

6. Numerical Procedure

To obtain a reduced-order model for the system $S(\tilde{A},\tilde{B},\tilde{C})$

(a) use similarity transformations[6] to give the balanced system $S(A,B,C)$

(b) partition the system $S(A,B,C)$ to give the strong subsystem $S(A_{11},B_1,C_1)$ and the weak subsystem $S(A_{22},B_2,C_2)$

(c) check whether the weak subsystem is fast by calculating the eigenvalues of the matrices A_{11} and A_{22}

(d) if (c) is true, calculate the reduced-order representation

$$\bar{A} = A_{11}-A_{12}A_{22}^{-1}A_{21} \quad , \quad \bar{B} = B_1-A_{12}A_{22}^{-1}B_2 \quad , \quad \bar{C} = C_1-C_2A_{22}^{-1}A_{21}$$

7. Example

To illustrate the procedure and the peculiarities of singular perturbational reduction of balanced systems, we have chosen the fourth-order model for the longitudinal dynamics of an F-8 aircraft without the wind disturbances[5].

An internally balanced representation of that model is given by

$$A = \begin{bmatrix} -6.560\text{x}10^{-3} & -7.577\text{x}10^{-2} & 7.390\text{x}10^{-4} & 3.564\text{x}10^{-3} \\ 7.577\text{x}10^{-2} & -8.383\text{x}10^{-3} & 9.204\text{x}10^{-4} & 4.445\text{x}10^{-3} \\ -9.171\text{x}10^{-4} & 1.142\text{x}10^{-3} & -9.219\text{x}10^{-2} & -3.086 \\ 3.597\text{x}10^{-3} & -4.486\text{x}10^{-3} & 3.136 & -1.816 \end{bmatrix}, \quad B = \begin{bmatrix} -4.713 \\ 4.831 \\ -3.293\text{x}10^{-1} \\ 1.292 \end{bmatrix}$$

$$C = \begin{bmatrix} 5.530\text{x}10^{-4} & 1.231\text{x}10^{-3} & -1.951\text{x}10^{-1} & -1.743\text{x}10^{-1} \\ 4.713 & 4.831 & -2.653\text{x}10^{-1} & -1.281 \end{bmatrix}$$

with $W = \text{diag}\,(1.693\text{x}10^{3} \quad 1.392\text{x}10^{3} \quad 5.882\text{x}10^{-1} \quad 4.597\text{x}10^{-1}$

If the system is partitioned into 2x2 subsystems in the natural order, the eigenvalues of the subsystems are given by

$$\lambda(A_{11}) = -7.472\text{x}10^{-3} \pm j\ 7.577\text{x}10^{-2}$$

$$\lambda(A_{22}) = -9.543\text{x}10^{-1} \pm j\ 2.989$$

Thus, the subsystem $S(A_{22},B_2,C_2)$ is considerably weak and fast compared with the subsystem $S(A_{11},B_1,C_1)$.

The balanced reduced-order model is then given by,

$$\bar{A} = \begin{bmatrix} -6.562\text{x}10^{-3} & -7.577\text{x}10^{-2} \\ 7.577\text{x}10^{-2} & -8.380\text{x}10^{-3} \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} -4.713 \\ 4.831 \end{bmatrix}$$

$$\bar{C} = \begin{bmatrix} 8.510\text{x}10^{-4} & 8.593\text{x}10^{-4} \\ 4.713 & 4.831 \end{bmatrix}$$

8. Conclusions

The singular perturbational model reduction technique is accommodated in the framework of internally balanced systems which gives a new unified technique which retains the advantages of both methods.

In some physical systems, singular perturbational parameters can be explicitly identified[1]. However, in general, this may not be possible and difficulties can be encountered in identifying the subsystems to be eliminated. Usually, this requires solution of (quadratic) Riccati equations to determine the fast systems[12]. By using balanced systems, we have avoided such difficulties and the balancing operation requires only solution of (linear) Lyapunov equations. The second order modes of the system (i.e. the diagonal elements of the matrix W) in someway act as the perturbational parameters. However, further research is necessary to investigate such aspects.

In this Chapter, we have exploited the Schur complement in perturbational reduction. However, it appears in a wider context in aggregation and in the Kron method of tearing among other areas, and thus it is reasonable to speculate that the balanced approach of Moore is applicable in these areas as well.

9. References

1. Kokotovic, P.V., O'Malley, R.E., and Sannuti, P.: 'Singular perturbations and order reduction in control theory - An overview', Automatica, 1976, 12, pp.123-132.
2. Sandell, Jr., N.R., Varaiya, P., Athans, M., and Safonov, M.G.: 'Survey of decentralized control methods for large scale systems', IEEE Trans. Automatic Control, 1978, AC-23, pp.108-128.
3. Nicholson, H.: 'Structure of Interconnected Systems', Peregrinus, London, 1978.
4. Fernando, K.V.M., and Nicholson, H.: 'Karhunen-Loève expansion with reference to singular-value decomposition and separation of variables', Proc.IEE, 1980, 127D, pp.204-206.
5. Teneketzis, D., and Sandell, Jr., N.R.: 'Linear regulator design for stochastic systems by a multiple time-scales method', IEEE Trans. Automatic Control, 1977, AC-22, pp.615-621.
6. Moore, B.C.: 'Principal component analysis in linear systems: Controllability, observability and model reduction', IEEE Trans. Automatic Control, 1981, AC-26, pp.17-32.
7. Fernando, K.V., and Nicholson, H.: 'On the structure of balanced and other principal representations of SISO systems', IEEE Trans. Automatic Control, 1983, AC-28, (3), (tentative).
8. Pernobo, L., and Silverman, L.M.: 'Balanced systems and model reduction', Proc.IEEE Conf. Decision and Control, Fort Lauderdale, Florida, pp.865-867, December 1979.
9. Cottle, R.W.: 'Manifestations of the Schur complement', Linear Algebra and its Applications, 1974, 8, pp.189-211.

10. Brillinger, D.R.: 'Time Series: Data Analysis and Theory', Holt, Rinehart and Winston, New York, 1974.

11. Aoki, M.: 'Control of large-scale dynamic systems by aggregation', IEEE Trans. Automatic Control, 1968, AC-13, pp.246-253.

12. O'Malley, R.E., Anderson, L.R.: 'Singular perturbations, order reduction, and decoupling of large-scale systems', in "Numerical Analysis of Singular Perturbation Problems", (Hemker, P.W., and Miller, J.J.H., eds.), Academic Press, London, 1979.

13. Calise, A.J.: 'Singular perturbation methods for variational problems in aircraft flight', IEEE Trans. Automatic Control, 1976, AC-21, pp.345-353.

CHAPTER 6

Singular Perturbational Approximations for Discrete-time Balanced Systems

Abstract: A new zero-order singular perturbational type approximation is developed for model reduction of discrete-time linear systems. This approximation is also suitable for model reduction of systems which have fast subsystems and which are represented in the internally balanced format. Subsystem elimination as suggested by Moore for continuous-time systems does not generalise for the discrete-time case and the singular perturbational approximation gives a particular solution to the discrete-time internally balanced model reduction problem.

1. Introduction

Principal axis realizations for discrete-time linear systems were first introduced by Mullis and Roberts[1,2] in the synthesis of minimum roundoff noise fixed-point digital filters. The problem under consideration was to find the optimum word length necessary in registers to optimize the storage and quantization efficiencies, simultaneously. These results were extended by Moore[3] for continuous-time linear systems. The storage and quantization effects can be translated into controllability and observability properties, respectively, if the control systems terminology is used. The best trade-off between high controllability with low observability and low controllability with high observability is provided by internally balanced principal axis state-space representations, which contain equal amounts of information about controllability and observability. Such balanced representations are convenient media for model reduction since equal amounts of information about controllability and observability can be neglected without causing any imbalance in controllability or observability properties.

For continuous-time balanced systems, Moore[3] proposed direct elimination of weak subsystems which are characterized by small second-order modes as a first approximation in the model reduction problem. This is a convenient and a very acceptable technique since the resultant reduced-order model is also internally balanced and retains the dominant second-order modes of the original system. However, this technique does not generalize for the discrete-time case in the sense that the reduced-order system is neither internally balanced nor contains the dominant second-order modes. The object of this Chapter is to demonstrate that if a zero-order singular perturbational approximation is used in model reduction of discrete-time systems which have fast subsystems, internally balanced reduced-order representations can be obtained which retain the dominant second-order modes.

2. Internally Balanced Discrete-time Systems

For the asymptotically stable, controllable and observable discrete-time linear system S(A,B,C) described in the format,

$$x(k+1) = Ax(k) + Bu(k) \quad , \quad y(k) = Cx(k)$$

the controllability Gramian matrix $W_c(p)$ and the observability Gramian matrix $W_o(p)$ are defined by

$$W_c(p) = \sum_{k=0}^{p} A^k BB^T (A^T)^k$$

$$W_o(p) = \sum_{k=0}^{p} (A^T)^k C^T C A^k$$

If the system S(A,B,C) is a principal axis representation, then the controllability and the observability Gramians are diagonal. If a system is represented by some other canonical form, then it can be brought into

a principal axis representation by using similarity transformations[1,2]. If a similarity transformation is used to bring them to be equal and diagonal, then the system is said to be internally balanced[3,4,9].

For the infinite time definition with $p \to \infty$, the Gramian matrices $W_c \triangleq W_c(\infty)$ and $W_o \triangleq W_o(\infty)$ may be computed by solving the discrete-time equivalents of the Lyapunov equations,

$$W_c - AW_cA^T = BB^T$$

$$W_o - A^TW_oA = C^TC$$

For the internally balanced system $S(\tilde{A},\tilde{B},\tilde{C})$, the matrix equations are given by

$$W - \tilde{A}W\tilde{A}^T = \tilde{B}\tilde{B}^T \qquad (1)$$

$$W - \tilde{A}^TW\tilde{A} = \tilde{C}^T\tilde{C} \qquad (2)$$

where W is a diagonal matrix. The assumed asymptotic stability, controllability and observability properties ensure that the diagonal values of W which are called the second-order modes are positive. We assume, without loss of generality, that these are ordered in the decreasing order of magnitude, and that they are distinct.

If the diagonal element w_{ii} of the Gramian matrix W is small in comparison with other elements, then the state x_i of the controllable system

$$x(k+1) = \tilde{A}x(k) + \tilde{B}u(k)$$

contributes marginally to the impulse response. Similarly, the contribution towards the impulse response of the observable system

$$x^d(k+1) = \tilde{A}^Tx^d(k) + \tilde{C}^Tv(k)$$

by the dual state x_i^d is small.

If the Gramian matrix W of dimension n,n, is partitioned in the format,

$$W = \left(\begin{array}{c|c} W_1 & 0 \\ \hline 0 & W_2 \end{array}\right)$$

where W_1 and W_2 are diagonal matrices of dimensions r,r and n-r,n-r respectively, then the system $S(\tilde{A},\tilde{B},\tilde{C})$ can be partitioned to conform to the above as,

$$\tilde{A} = \left(\begin{array}{c|c} \tilde{A}_{11} & \tilde{A}_{12} \\ \hline \tilde{A}_{21} & \tilde{A}_{22} \end{array}\right), \quad \tilde{B} = \left(\begin{array}{c} \tilde{B}_1 \\ \hline \tilde{B}_2 \end{array}\right), \quad \tilde{C} = [\tilde{C}_1 \,|\, \tilde{C}_2]$$

We assume that the diagonal elements of W_2 are appreciably smaller than those of W_1 and following Moore[3], we call the subsystem $S(\tilde{A}_{11},\tilde{B}_1,\tilde{C}_1)$ the strong subsystem and the subsystem $S(\tilde{A}_{22},\tilde{B}_2,\tilde{C}_2)$ the weak subsystem. However, we emphasise that the subsystems $S(A_{ii},B_i,C_i)$, $i = 1,2$ are not internally balanced as in the case of the continuous-time equivalent and the submatrices W_i, $i = 1,2$ are not balanced Gramians of the subsystems. These conditions preclude direct extension of the continuous-time results of Moore[3] to discrete-time systems.

3. Singular Perturbational Reduction of Discrete-time Systems

In a survey of large-scale systems[5], it was pointed out that singular perturbational results for discrete-time systems are not widely available. This situation has been remedied to some extent in the literature (see for example references 6-8). However, the published results are not directly suitable for our problem.

For continuous-time systems, a subsystem is said to be fast if the eigenvalues of the subsystem are large (i.e. $s \to -\infty$) in the complex plane, and a subsystem is said to be slow if its eigenvalues are near the origin

(i.e. $s \approx 0$). We carry this definition to the discrete-time case where a subsystem is said to be fast if the eigenvalues are near the origin (i.e. $z \approx 0$) and slow if they are near $z = 1$ in the complex plane. This extension is consistent with the well known sampled-data approximation $z = e^{sT}$ where T is the sampling time.

It is usually assumed in models of singular perturbation that the fast states rapidly approach a linear combination of the slow states. However, in references 6-8, it has been assumed that the slow states can be approximated using the fast states which is the converse of the usual assumption. Such a hypothesis leads to approximations in fast-time rather than in slow-time.

In this section, we develop a slow-time approximation for discrete-time systems. We recall that a continuous-time system $S_\mu(F,G,H)$ can be written in the singular perturbational format,

$$\dot{x}_1(t) = F_{11}x_1(t) + F_{12}x_2(t) + G_1u(t) \qquad \text{(slow-time)}$$

$$\mu\dot{x}_2(t) = F_{21}x_1(t) + F_{22}x_2(t) + G_2u(t) \qquad \text{(fast-time)}$$

where μ is a positive small perturbational parameter. By considering the approximation $\dot{x}(t) \approx [x(k+1)-x(k)]/T$, we propose the analogous singular perturbational discrete-time model $S_\mu(A,B,C)$ in the format,

$$x_1(k+1) - x_1(k) = (A_{11}-I)x_1(k) + A_{12}x_2(k) + B_1u(k) \qquad \text{(slow-time)}$$

$$\mu[x_2(k+1) - x_2(k)] = A_{21}x_1(k) + (A_{22}-I)x_2(k)+B_2u(k) \qquad \text{(fast-time)}$$

As in the continuous-time case, we assume that the matrices A_{12} and A_{21} are small and that the subsystem matrix A_{22} is fast. With $\mu \to 0$, we obtain the approximation

$$x_1(k+1) = [A_{11}+A_{12}(I-A_{22})^{-1}A_{21}]x_1(k) + [B_1+A_{12}(I-A_{22})^{-1}B_2]u(k)$$

(slow-time)

$$x_2(k) = (I-A_{22})^{-1}A_{21}x_1(k) + (I-A_{22})^{-1}B_2u(k) \quad \text{(fast-time)}$$

Similarly, by considering the observable system, the reduced-order approximation in slow-time can be obtained as $S(\hat{A},\hat{B},\hat{C})$, with

$$\hat{A} = A_{11} + A_{12}(I-A_{22})^{-1}A_{21}$$

$$\hat{B} = B_1 + A_{12}(I-A_{22})^{-1}B_2$$

$$\hat{C} = C_1 + C_2(I-A_{22})^{-1}A_{21}$$

We observe that the inverse of the matrix $(I-A_{22})$ always exists under our assumptions. The above approximation can be derived independently without considering the controllable and the observable systems separately [11] (see Appendix 3). To our knowledge, these reduced-order approximations have not appeared in the control systems literature. However, structurally similar "aggregations" have been used by Leontief in econometric problems[12].

4. Weak and Fast Subsystems

If fast subsystems are present in the system, then the transients associated with these subsystems will vanish quickly and thus the overall contribution towards the impulse response of the system will be small. Thus, fast subsystems will be characterized by relatively small diagonal elements of the matrix W corresponding to weak subsystems. However, we do not imply that all weak subsystems are fast and thus, this property should be checked at each instant of application of the proposed method of reduction.

Numerical experience[10] (see Chapter 5) with continuous-time systems which are suitable for perturbational reduction indicates that fast subsystems are substantially weak and this property also can be expected in discrete-time systems.

5. Singular Perturbational Reduction of Balanced Systems

If the weak subsystem is also a fast subsystem, then it can be removed using singular perturbational approximations. The following proposition shows a means of achieving this.

Proposition: If the system $S(\tilde{A},\tilde{B},\tilde{C})$ is internally balanced, then the singular perturbational approximation $S(\hat{A},\hat{B},\hat{C})$ also defines an internally balanced system. Further, the diagonal Gramian for the reduced-order system is given by the matrix W_1. Thus, the matrix equations which describe the Gramian are of the form,

$$W_1 - \hat{A}W_1\hat{A}^T = \hat{B}\hat{B}^T \qquad (3)$$

$$W_1 - \hat{A}^TW_1\hat{A} = \hat{C}^T\hat{C} \qquad (4)$$

Proof: Equation (3) can be obtained by premultiplying and postmultiplying (1) by T_1 and T_1^T, respectively, where

$$T_1 = [I \;\vdots\; \tilde{A}_{12}(I-\tilde{A}_{22})^{-1}]$$

Similarly, (4) can be obtained by premultiplying and postmultiplying (2) by T_2 and T_2^T, respectively, where

$$T_2 = [I \;\vdots\; \tilde{A}_{21}^T(I-\tilde{A}_{22}^T)^{-1}]$$

Since the Gramian matrix W_1 is diagonal, the reduced-order system is also internally balanced and this completes the proof.

The next proposition indicates that the reduced model has very desired properties.

Proposition: The reduced-order model $S(\hat{A},\hat{B},\hat{C})$ is asymptotically stable, controllable, and observable.

Proof: The positive definiteness of the diagonal matrix W_1 guarantees these properties[4].

We observe that the results given in this section are algebraically correct whether the system conforms to the perturbational model developed in Section III or not. However, the reduced-order model $S(\hat{A},\hat{B},\hat{C})$ cannot be considered as a good approximation of the original system $S(A,B,C)$ due to retention of the dominant singular values alone. Thus, the supporting evidence from the singular perturbational model is required to justify our approximation.

6. Numerical Procedure

To obtain a reduced-order model for the system $S(A,B,C)$

(a) use similarity transformations[4,9] to give the balanced system $S(\tilde{A},\tilde{B},\tilde{C})$

(b) partition the system $S(\tilde{A},\tilde{B},\tilde{C})$ to give the strong subsystem $S(\tilde{A}_{11},\tilde{B}_1,\tilde{C}_1)$ and the weak subsystem $S(\tilde{A}_{22},\tilde{B}_2,\tilde{C}_2)$ by inspecting the diagonal values of the Gramian matrix W

(c) check whether the weak subsystem is fast by calculating the eigenvalues of the matrix $\tilde{A}_{22}$ and check the smallness of A_{12} and A_{21}

(d) if (c) is true, calculate the first-order singular perturbational approximation

$$\hat{A} = \tilde{A}_{11} + \tilde{A}_{12}(I-\tilde{A}_{22})^{-1}\tilde{A}_{21} \quad , \quad \hat{B} = \tilde{B}_1 + \tilde{A}_{12}(I-\tilde{A}_{22})^{-1}\tilde{B}_2$$

$$\hat{C} = \tilde{C}_1 + \tilde{C}_2(I-\tilde{A}_{22})^{-1}\tilde{A}_{21}$$

7. Conclusions

We have demonstrated the feasibility of singular perturbational model reduction of systems which are represented in internally balanced formats. This is particularly significant, since the subsystem elimination method proposed by Moore[3] breaks down for the discrete case. However, the perturbational approximations are valid for the discrete case as well as for the continuous case[10] (Chapter 5).

Due to this inherent consistency of perturbational reduction in internally balanced systems, it may be possible to use other techniques available in the singular perturbational approach for the design and analysis of balanced systems. Since the internally balanced representations are well-conditioned with respect to controllability and observability, which are fundamental in control and other systems studies, balanced representations could be used to give a more formal structure with respect to controllability and observability in singular perturbation analysis.

8. References

1. Mullis, C.T., and Roberts, R.A.: 'Synthesis of minimum roundoff noise fixed point digital filters', IEEE Trans. Circuits and Systems, 1976, CAS-23, (8), pp.551-561.
2. Mullis, C.T., and Roberts, R.A.: 'Roundoff noise in digital filters: Frequency transformations and invariants', IEEE Trans. Acoustics, Speech and Signal Processing, 1976, ASSP-24, pp.538-550.
3. Moore, B.C.: 'Principal component analysis in linear systems: Controllability, observability, and model-reduction', IEEE Trans. Automatic Control, 1981, AC-276, (1), pp.17-32.
4. Pernobo, L., and Silverman, L.M.: 'Balanced systems and model reduction', Proc. IEEE Conf. on Decision and Control, Fort Lauderdale, Florida, pp.865-867, Dec. 1979.
5. Sandell, Jr., N.R., Varaiya, P., Athans, M., and Safonov, M.G.: 'Survey of decentralized control methods for large scale systems', IEEE Trans. Automatic Control, 1978, AC-23, pp.108-128.
6. Phillips, R.G.: 'Reduced modelling and control of two-time scale discrete systems', Int. J. Control, 1980, 31, pp.765-780.
7. Rajagopalan, P.K., and Naidu, S.: 'A singular perturbed difference equation in optimal control problems', Int. J. Control, 1980, 32, pp.925-936.
8. Blankenship, G.: 'Singularly perturbed difference equations in optimal control problems', IEEE Trans. Automatic Control, 1981, AC-26, (4), pp.911-917.
9. Kung, S.-Y., and Lin, D.W.: 'A state-space formulation for optimal Hankel-norm approximations', 1981, AC-26, (4), pp.942-946.

10. Fernando, K.V., and Nicholson, H.: 'Singular perturbational reduction of balanced systems', IEEE Trans. Automatic Control, 1982, AC-27, (32), pp.466-468.

11. Fernando, K.V., and Nicholson, H.: 'Singular perturbational model reduction in the frequency domain', IEEE Trans. Automatic Control, 1982, AC-27, (4).

12. Leontief, W.: 'An alternative to aggregation in input-output analysis and national accounts', The Review of Economics and Statistics, 1967, 49, pp.412-419.

CHAPTER 7

On Balanced Model-Order Reduction of Discrete-Time Systems and their Continuous-Time Equivalents

Abstract: The model-order reduction method for continuous-time systems based on subsystem elimination as proposed by Moore does not generalise to the discrete-time case. In this Chapter, we develop the discrete-time equivalent of the continuous-time technique using the Cayley transformation between discrete-time and continuous-time equivalents. The new reduced-order approximations based on this method are exactly balanced and retain the dominant modes of the system. The suitability of this result is verified using a singular perturbational model.

1. Introduction

The balanced model-order reduction method of Moore[1] is based on removal of 'modes' which are weak corresponding to joint controllability and observability (i.e. the minimality) as characterised by the controllability Gramian matrix W_c and the observability Gramian matrix W_o. The technique can be implemented by realization of the system in the balanced format, in which case the two Gramian matrices are equal and diagonal. The diagonal values are called the second order modes of the system. The robust part of the system is represented by relatively high second-order modes and the weak part (if any) by low values. The direct removal of the weak subsystem gives a robust approximation of the original system. Such approximate representations retain the dominant second-order modes of the original system and are balanced.

Direct generalization of this technique to the discrete-time case is not possible. If subsystem elimination is used for discrete-time systems, the resultant approximation neither retains the dominant second-order modes nor is it balanced[2,3].

In this Chapter, we develop the discrete-time equivalent of the continuous-time balanced technique of Moore using the Cayley transformation. This method gives reduced-order balanced representations which retain the dominant modes of the original system. The suitability of the approximation is investigated using a singular perturbational model. Furthermore, it is shown that this approximation can be considered as a 'generalized singular perturbational' technique[6] (Appendix 3). This solution is also complementary to the singular perturbational approximation for balanced discrete-time systems[5] (Chapter 6). For completeness, we summarize the balanced approximations for discrete-time systems and their equivalents for continuous-time systems.

2. Preliminaries

For the continuous-time asymptotically stable time-invariant system S(A,B,C)

$$\dot{x}(t) = Ax(t) + Bu(t) \quad , \quad y(t) = Cx(t)$$

the controllability Gramian matrix W_c and the observability Gramian matrix W_o can be obtained as the solution of the Lyapunov equations given by

$$W_c A^T + A W_c = - BB^T$$

$$W_o A + A^T W_o = - C^T C$$

If the system S(A,B,C) is a balanced representation[1], then the Gramian matrices are diagonal and equal.

$$W_c = W_o \triangleq W$$

The diagonal values of the Gramian matrix W are called the second-order modes, and we assume that they are ordered in the non-increasing order of magnitude. We may partition the matrix W in the format,

$$W = \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix}$$

where the diagonal matrix W_1 contains the dominant modes of the system and the matrix W_2 contains the non-dominant values. Conforming to the above partition, the system S(A,B,C) can be decomposed in the format

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \quad C = [C_1 \quad C_2]$$

By elimination of the weak subsystem $S(A_{22},B_2,C_2)$, we may obtain the reduced-order approximation of the original system as the subsystem $S(A_{11},B_1,C_1)$, which is a balanced representation with the Gramian matrix W_1.

For the discrete-time asymptotically stable time-invariant system S(F,G,H)

$$x(k+1) = Fx(k) + Gu(k) \quad , \quad y(k) = Hx(k)$$

the Gramian matrices can be obtained as the solution of the matrix equations given by

$$W_c - FW_cF^T = GG^T$$
$$W_o - F^TW_oF = H^TH$$

If the system is balanced, then

$$W_c = W_o = W$$

where W is a diagonal matrix ordered in the non-increasing order of magnitude. As in the continuous-time case, we may partition the discrete-time system in the format

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}, \quad G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}, \quad H = \begin{bmatrix} H_1 & H_2 \end{bmatrix}$$

The subsystem $S(F_{11},G_1,H_1)$ has been suggested[2,3] as a reduced-order approximation of the original system. However, this subsystem is not a balanced representation and does not retain the dominant second-order modes of W_1. Thus, this approximation method cannot be considered as the generalization of the balanced method of Moore[1].

3. Balanced Model Order Reduction

It is well known[7] that the Gramian matrices are identically equal if the continuous-time system S(A,B,C) and the discrete-time system S(F,G,H) are related by the Cayley transformation defined by

$$A = -(I+F)^{-1}(I-F)$$

$$B = \pm \sqrt{2}\,(I+F)^{-1}G$$

$$C = \pm \sqrt{2}\,H(I+F)^{-1}$$

This also corresponds to transformations between immittance matrix and scattering matrix descriptions in network theory[7,10].

If the weak subsystem $S(A_{22},B_2,C_2)$ corresponding to the balanced Gramian matrix W_2 is removed, then the balanced approximation

$S(A_{11},B_1,C_1)$ is given by

$$A_{11} = -(I+\bar{F})^{-1}(I-\bar{F})$$

$$B_1 = \pm \sqrt{2}\,(I+\bar{F})^{-1}\bar{G}$$

$$C_1 = \pm \sqrt{2}\,\bar{H}(I+\bar{F})^{-1}$$

where

$$\bar{F} = F_{11} - F_{12}(I+F_{22})^{-1}F_{21}$$

$$\bar{G} = G_1 - F_{12}(I+F_{22})^{-1}G_2$$

$$\bar{H} = H_1 - H_2(I+F_{22})^{-1}F_{21}$$

This result may be verified using the well known matrix lemma[10] for inversion of partitioned matrices. The reduced-order system $S(A_1,B_1,C_1)$ has the balanced Gramian matrix W_1 and the equivalent discrete-time system $S(\bar{F},\bar{G},\bar{H})$ also has the same balanced Gramian matrix.

Thus, the required approximation corresponding to subsystem elimination in continuous-time is given by the discrete-time subsystem $S(\bar{F},\bar{G},\bar{H})$ if the Cayley transformation is assumed as the criterion of equivalence between continuous-time and discrete-time systems.

4. The Singular Perturbational Interpretation

It has been demonstrated previously[4,5] (Chapters 5,6 and Appendix 3) that balanced model-order reduction is consistent with the singular perturbational technique. We now describe the derived reduced-order model $S(\bar{F},\bar{G},\bar{H})$ using a singular perturbational argument. To obtain this approximation, we define the singular perturbational

model $S_\mu(F,G,H)$ in the form,

$$x_1(k+1) + x_1(k) = (F_{11}+I)x_1(k) + F_{12}x_2(k) + G_1u(k) \quad \text{(fast-time)}$$

$$\mu[x_2(k+1) + x_2(k)] = F_{21}x_1(k) + (F_{22}+I)x_2(k)+G_2u(k) \quad \text{(slow-time)}$$

where μ denotes a small perturbational parameter. As usual, we assume that the matrices F_{12} and F_{21} which define the interaction between subsystems are small. Then, with $\mu=0$, we may obtain the required approximations as

$$x_1(k+1) = \bar{F}x_1(k) + \bar{G}u(k) \quad \text{(fast-time)}$$

$$x_2(k+1) = -(I+F_{22})^{-1}F_{21}x_1(k) - (I+F_{22})^{-1}G_2u(k) \quad \text{(slow-time)}$$

It is obvious that the approximation is valid only if the eigenvalues of the matrix sum $(I+F_{22})$ are large. This is equivalent to the requirement that the eigenvalues of the matrix F_{22} are near $z = 1$ in the complex plane. In addition, we assume that the eigenvalues of the matrix F_{11} are away from $z = 1$ and thus the subsystem $S(F_{11},G_1,H_1)$ is a 'fast' subsystem.

Similarly, by considering the observable system, it is possible to derive the complete singular perturbational approximation in 'fast-time' as $S(\bar{F},\bar{G},\bar{H})$.

5. Generalized Singular Perturbational Balanced Approximations

Fernando and Nicholson[6] (Appendix 3) demonstrated that balanced approximations are special cases of 'generalized singular perturbational' approximations. There are two possible balanced approximations for continuous-time systems with equivalent results in discrete-time. For completeness, we review these results and indicate the applicability and suitability of each reduced-order representation.

The generalized approximation $S(A_g,B_g,C_g)$ for continuous-time systems are of the form,

$$A_g(s_0) = A_{11} + A_{12}(s_0 I - A_{22})^{-1} A_{21}$$

$$B_g(s_0) = B_1 + A_{12}(s_0 I - A_{22})^{-1} B_2$$

$$C_g(s_0) = C_1 + C_2(s_0 I - A_{22})^{-1} A_{21}$$

where s_0 is the dominant frequency of the robust subsystem $S(A_{11},B_1,C_1)$ and the non-dominant frequency of the weak subsystem $S(A_{22},B_2,S_2)$. Similarly, for discrete-time systems, they are given by

$$F_g(z_0) = F_{11} + F_{12}(z_0 I - F_{22})^{-1} F_{21}$$

$$G_g(z_0) = G_1 + F_{12}(z_0 I - F_{22})^{-1} G_2$$

$$H_g(z_0) = H_1 + H_2(z_0 I - F_{22})^{-1} F_{21}$$

If the frequency s_0 is negative infinite corresponding to a 'fast' subsystem $S(A_{11},B_1,C_1)$ and a 'slow' subsystem $S(A_{22},B_2,C_2)$, then the result obtained by Moore through subsystem elimination can be derived.

$$S(A_{11},B_1,C_1) = S(A_g(-\infty),\ B_g(-\infty),\ C_g(-\infty))$$

The corresponding discrete-time equivalent with $z_0 = -1$ is given by

$$S(\bar{F},\bar{G},\bar{H}) = S(F_g(-1),\ G_g(-1),\ H_g(-1))$$

If the frequency s_0 is zero, corresponding to a 'slow' subsystem $S(A_{11},B_1,C_1)$ and a fast subsystem $S(A_{22},B_2,C_2)$, then the result obtained by Fernando and Nicholson[4] (Chapter 5) manifests,

$$S(A_{11}-A_{12}A_{22}^{-1}A_{21},\ B_1-A_{12}A_{22}^{-1}A_{21}B_2,\ C_1-C_2A_{22}^{-1}A_{21})$$

$$= S(A_g(0),\ B_g(0),\ C_g(0))$$

which is the traditional singular perturbational approximation given by the Schur complement. The corresponding approximation in discrete-time with $z_0 = 1$ is then given by

$$S(F_{11}+F_{12}(I-F_{22})^{-1}F_{21},\ G_1+F_{12}(I-F_{22})^{-1}G_2,\ H_1+H_2(I-F_{22})^{-1}F_{21})$$

$$= S(F_g(1),\ G_g(1),\ H_g(1))$$

which was derived by Fernando and Nicholson[5] (Chapter 6).

Thus, it is possible to present a unified approach for balanced model-order reduction using the generalized singular perturbational approach.

6. Conclusions

We have developed the equivalent of the subsystem elimination method of Moore in discrete-time and have shown that it does not generalise in the manner suggested by Moore. Instead, balanced model-order reduction which retains the dominant second-order modes is best explained using singular perturbational arguments.

For completeness, we have demonstrated that there are two continuous-time generalized approximations which can give balanced reduced-order models. The 'fast' and 'slow' approximations have equivalent counterparts in discrete-time systems through the Cayley transformation. In network theory, this relationship can be considered as the equivalence between immittance matrix and scattering matrix descriptions.

Perhaps the feasibility of singular perturbational model-order approximations in the framework of balanced systems was first

suggested by Verriest and Kailath[9] but was not elaborated by those authors. We have demonstrated conclusively that the singular perturbational technique is the central theme in balanced model-order reduction in both continuous-time and discrete-time systems.

7. References

1. Moore, B.C.: 'Principal component analysis in linear systems: Controllability, observability, and model reduction', IEEE Trans. Automatic Control, 1981, AC-26, (1), pp.17-32.
2. Kung, S.: 'A new identification and model reduction algorithm via singular value decomposition', 12th Asilomar Conf. Circuits, Systems, and Computers, Nov. 1978.
3. Silverman, L.M., and Bettayeb, M.: 'Optimal approximations of linear systems', Proc. JACC, paper FA8-A, August 1980.
4. Fernando, K.V., and Nicholson, H.: 'Singular perturbational reduction of balanced systems', IEEE Trans. Automatic Control, 1982, AC-27, (2), pp.466-468.
5. Fernando, K.V., and Nicholson, H.: 'Singular perturbational approximations for discrete-time systems', IEEE Trans. Automatic Control, 1983, AC-28, (1).
6. Fernando, K.V., and Nicholson, H.: 'Singular perturbational model-order reduction in the frequency domain', IEEE Trans. Automatic Control, 1982, AC-27, (4).
7. Anderson, B.D.O., and Vongpanitlerd, S.: 'Network Analysis and Synthesis', Prentice-Hall, Englewood Cliffs, NJ, 1973.
8. Pernebo, L., and Silverman, L.M.: 'Model reduction via balanced state space representations', IEEE Trans. Automatic Control, 1982, AC-27, (2), pp.382-387.

9. Verriest, E., and Kailath, T.: 'On generalized balanced realizations', Proc. 19th IEEE Conf. Decision and Control, Albuquerque, NM, pp.504-505, Dec. 1980.

10. Nicholson, H.: 'Structure of Interconnected Systems', Peregrinus, London, 1978.

CHAPTER 8

Reciprocal Transformations in Balanced Model-Order Reduction

Abstract: Direct elimination of weak subsystems and singular perturbational approximations given by the Schur complement have been suggested for model-order reduction of balanced linear systems. These two approaches are dual to each other and such reciprocal approximations are well known in other model-order reduction techniques. A standard example is used to illustrate the two reciprocal approaches.

1. Introduction

Moore[1] was able to interpret in an elegant manner, the minimal realization problem and the model-order reduction problem in linear systems theory from the point of view of 'signal injection'. Instead of relying on the classical parameters of the system, which may be susceptible to structural instabilities, Moore based the realization on second-order averages of the controllable part of the system and the observable part which are excited by impulse inputs. These averages, given by the controllable Gramian matrix and the observable Gramian matrix are central to the realization of internally balanced models. Using this approach, Moore was able to show that 'nearly optimal' reduced-order representations which have approximately the same impulse responses as the original system can exist. This is achieved by removing the 'weak' subsystems and

retaining the robust part of the system corresponding to the 'strong' second-order modes. These second-order modes are given by Gramian matrices which are diagonal and equal for balanced realizations.

Fernando and Nicholson[4] (Chapter 5) considered the special case where the weak subsystem corresponds to fast dynamics and the strong subsystem to slow dynamics of the system. Most physical systems behave in this manner, and the concept has been exploited in modal methods of model-order reduction. Singular perturbational approximations are feasible for this case and a low-frequency approximation at $s = 0$ is given by the usual Schur complement result. Fernando and Nicholson[5] (Appendix 3) also demonstrated that if the weak subsystem and the strong subsystem are due to slow and fast dynamics, respectively, then a 'generalized singular perturbational' approximation is possible at $s = -\infty$. This result is given by direct elimination of the weak subsystem which was first proposed by Moore as a 'balanced model-order reduction'. However, Moore derived this result using reasonable but heuristic arguments and no conventional explanation was given. In both these approaches, the resultant reduced-order representations are balanced and retain the dominant second-order modes of the original system.

The object of this Chapter is to show that the direct elimination method of Moore and the usual perturbational result are related by a reciprocal transformation. Similar transformations are well known in other model-order reduction techniques[2,3,10] (Appendix 4). We demonstrate the viability of the two approaches and compare the results using a standard example.

2. Internally Balanced Models and their Reciprocals

For the time-invariant, asymptotically stable linear system S(A,B,C) described by

$$\dot{x}(t) = Ax(t) + Bu(t) \quad , \quad y(t) = Cx(t)$$

the controllability Gramian matrix W_c^2 and the observability Gramian matrix W_o^2 may be defined in the infinite interval by

$$W_c^2 = \int_o^\infty (e^{At}B)(e^{At}B)^T dt$$

$$W_o^2 = \int_o^\infty (e^{A^T t}C^T)(e^{A^T t}C^T)^T dt$$

These Gramian matrices can also be obtained as the solution of the following Lyapunov equations

$$W_c^2 A^T + AW_c^2 = - BB^T$$

$$W_o^2 A + A^T W_o^2 = - C^T C$$

The state-space representation S(A,B,C) is said to be internally balanced if the Gramian matrices W_c^2 and W_o^2 are equal and diagonal[1].

$$W_c^2 = W_o^2 = W^2$$

Internally balanced representations can be obtained using similarity transformations[1] and without loss of generality, we assume that the system S(A,B,C) is internally balanced with the diagonal Gramian matrix W^2. The diagonal elements of the Gramian matrix W^2 are called the second-order modes of the system and we assume that they are ordered as a non-increasing sequence.

We define the 'reciprocal system' S(F,G,H) of S(A,B,C) by

$$F = A^{-1} \quad , \quad G = A^{-1}B \quad , \quad H = CA^{-1}$$

The system S(F,G,H) is also asymptotically stable and the Lyapunov equations are given by

$$W^2F^T + FW^2 = -GG^T$$

$$W^2F + F^TW^2 = -H^TH$$

The reciprocal system S(F,G,H) is also controllable and observable and has the same diagonal Gramian matrix W^2.

It is easily seen that the reciprocal system S(F,G,H) is given by the system S(A,B,C) and thus these two systems are dual to each other. This dual relationship is also reflected in the Markov parameters and the moments of the system. The kth moment of the system S(A,B,C) is equal to the (k+3)rd Markov parameter of the system S(F,G,H)

$$CA^{-k}B = HF^{k+3-1}G \quad , \qquad k \geq 1$$

Similarly, the kth moment of the system S(F,G,H) is given by the (k-1)th Markov parameter of the system S(A,B,C)

$$HF^{-k}G = CA^{k-1-1}B$$

These relations indicate that a reciprocal system has reciprocal properties to the original system. We observe particularly that if a system has dominant high-frequency behaviour, then the reciprocal system will have dominant low-frequency behaviour and vice-versa.

Similar reciprocal relationships are well known in the model-reduction literature and have been widely used to eliminate inherent frequency biases in the Routh approximation method[2,3] and elsewhere[10].

3. Model-Order Reduction of Balanced Systems and their Reciprocals

The nth order internally balanced system S(A,B,C) can be partitioned in the format,

$$A = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right), \quad B = \left(\begin{array}{c} B_1 \\ \hline B_2 \end{array}\right), \quad C = [C_1 \mid C_2]$$

where $S(A_{11},B_1,C_1)$ and $S(A_{22},B_2,C_2)$ are rth order and (n-r)th order subsystems, respectively. If the diagonal Gramian matrix W^2 is partitioned in the format

$$W^2 = \left(\begin{array}{c|c} W_1^{\,2} & 0 \\ \hline 0 & W_2^{\,2} \end{array}\right)$$

the diagonal Gramians $W_i^{\,2}$, i = 1,2 can be associated with the internally balanced subsystems $S(A_{ii},B_i,C_i)$, i = 1,2. Without loss of generality, we assume that the diagonal values of the matrix $W_1^{\,2}$ are large in comparison with the elements of $W_2^{\,2}$. Following Moore[1], the subsystem $S(A_{11},B_1,C_1)$ is called the 'strong' subsystem and the subsystem $S(A_{22},B_2,C_2)$, the 'weak' subsystem.

Moore advocated, as a first approximation, direct elimination of the weak subsystem and thus, the reduced-order model is simply given by the strong subsystem $S(A_{11},B_1,C_1)$.

We now investigate the structure of the model if reciprocal transformations are used in this model-order reduction process. Instead of eliminating the weak subsystem of the system S(A,B,C), it is first transformed into the reciprocal format S(F,G,H) and the weak subsystem of S(F,G,H) is eliminated to give $S(F_{11},G_1,H_1)$. The

reduced-order model $S(\hat{A}_{11},\hat{B}_1,\hat{C}_1)$ is obtained by computing the reciprocal of $S(F_{11},G_1,H_1)$. Thus,

$$S(A,B,C) \xrightarrow{\text{reciprocal}} S(F,G,H) \xrightarrow[\text{subsystem}]{\text{eliminate weak}} S(F_{11},G_1,H_1) \xrightarrow{\text{reciprocal}} S(\hat{A}_{11},\hat{B}_1,\hat{C}_1)$$

Again, we point out that similar intermediate reciprocal transformations are common in the model-order reduction literature[2,3,10].

If this procedure is followed, then the reduced-order model $S(\hat{A}_{11},\hat{B}_1,\hat{C}_1)$ is given by the usual singular perturbational approximation,

$$A_{11} = A_{11} - A_{12}A_{22}^{-1}A_{21} \quad , \quad B_1 = B_1 - A_{12}A_{22}^{-1}B_2$$

$$C_1 = C_1 - C_2A_{22}^{-1}A_{21}$$

It can be shown[4] that reduced-order model $S(A_{11},B_1,C_1)$ is also an internally balanced system with the balanced Gramian matrix W_1^2.

The above low-frequency singular perturbational approximation can be derived by considering the inverse of the matrix A in the format of the well known lemma for inversion of partitioned matrices which is also sometimes known as the K-partitioned inverse[8].

$$A^{-1} = \left[\begin{array}{c|c} F_{11} & -F_{11}A_{12}A_{22}^{-1} \\ \hline -A_{22}^{-1}A_{21}F_{11} & A_{22}^{-1}+A_{22}^{-1}A_{21}F_{11}^{-1}A_{12}A_{22}^{-1} \end{array}\right]$$

with $$F_{11}^{-1} = A_{11} - A_{12}A_{22}^{-1}A_{21}$$

Thus, a direct duality exists between direct and reciprocal elimination of subsystems through the reciprocal transformation.

We recall that for low-frequency singular perturbational approximation to be valid, the weak subsystem $S(A_{22},B_2,C_2)$ has to be a fast subsystem. However, the high-frequency approximation given by $S(A_{11},B_1,C_1)$ is valid only if the weak subsystem $S(A_{22},B_2,C_2)$ is slow[5]. Thus, these two methods are complementary to each other.

4. An Illustrative Example

To illustrate the two different approaches in the balanced model-order reduction problem, consider the system described by the transfer-function,

$$\frac{(s+4)}{(s+1)(s+3)(s+5)(s+10)}$$

which has been previously studied by Meier and Luenberger[6], Wilson[7], and recently by Moore[1] and Fernando and Nicholson[9] (Chapter 11). The system can be realized in the internally balanced format[1],

$$A = \left[\begin{array}{cc|cc} -0.4378 & 1.168 & 0.4143 & 0.05098 \\ -1.168 & -3.135 & -2.835 & -0.3289 \\ \hline 0.4143 & 2.835 & -12.48 & -3.249 \\ -0.05098 & -0.3289 & 3.249 & -2.952 \end{array}\right], \quad B = \left[\begin{array}{c} -0.1181 \\ -0.1307 \\ \hline 0.05634 \\ -0.006875 \end{array}\right]$$

$$C = \left[\begin{array}{cc|cc} -0.1181 & 0.1307 & 0.05634 & 0.006875 \end{array}\right]$$

and the balanced Gramian matrix W^2 is given by,

$$W^2 = \mathrm{diag}\left[\begin{array}{cc|cc} 0.01954 & 0.272\times10^{-2} & 0.1272\times10^{-3} & 0.8006\times10^{-5} \end{array}\right]$$

The system can be decomposed into 2x2 subsystems, in the natural order, as shown. The subsystem $S(A_{11},B_1,C_1)$ is the strong subsystem and it is the reduced-order balanced approximation of Moore.

If the reciprocal method of reduction is used, the perturbational result is given by

$$\hat{A}_{11} = \begin{bmatrix} -0.4249 & 1.257 \\ -1.257 & -3.735 \end{bmatrix} , \quad \hat{B}_1 = \begin{bmatrix} -0.1164 \\ -0.1427 \end{bmatrix}$$

$$\hat{C}_1 = \begin{bmatrix} -0.1164 & 0.1427 \end{bmatrix}$$

We observe that the perturbational correction is not very prominent for this example. The eigenvalues of the subsystem matrices A_{11} and A_{22} are given by

$$\lambda(A_{11}) = -1.113 , \quad -2.460$$

$$\lambda(A_{22}) = -11.20 , \quad -4.232$$

Thus, the weak subsystem is faster than the strong subsystem, although the effect is not very pronounced.

The transfer functions associated with the reduced models, together with their error ratios are given in Table 1. If the scalar h(t) denotes the impulse response of the original system and $\bar{h}(t)$ is a reduced-order representation, then the error ratio is defined as,

$$\frac{\left(\int_0^\infty h_e^2(t)\, dt\right)^{\frac{1}{2}}}{\left(\int_0^\infty h^2(t)\, dt\right)^{\frac{1}{2}}} \qquad \text{where} \quad h_e(t) = h(t) - \bar{h}(t)$$

For comparison purposes, the optimal solution with respect to the least-squares criterion[6,7],

$$\text{minimize} \int_0^\infty h_e^2(t)\, dt$$

is also included in Table 1. This was obtained by direct numerical optimization.

It can be seen from Table 1 that the reduced-order models obtained through direct elimination and reciprocal elimination (singular perturbational result) are nearly optimal. If the optimality is paramount, the direct elimination result is the preferred solution for this problem. However, the near optimality has been achieved by having a badly positioned numerator zero. In the result based on the reciprocal approach, the zero is most favourably positioned. This can be an advantage in certain applications and design procedures.

Figure 1 gives the frequency response of the systems. Both reduced-order models represent the original transfer-function adequately. As expected, at high frequencies, the direct solution is marginally better than the reciprocal approach.

Figure 2 gives the phase behaviour of the original system and the reduced-order representations. Even at high frequencies, the phase behaviour of the reduced-order representation obtained through reciprocal transformations is better than the direct approach.

5. Conclusions

We have demonstrated that a duality exists between direct elimination of subsystems as proposed by Moore and the low-frequency singular perturbational-type approximation as given by the Schur complement, and the two approximations are reciprocal to each other.

A standard numerical example is used to illustrate the form of the reduced-order solutions. There are certain minor advantages and disadvantages in these two approaches and we have compared these with the optimal solution.

If the system under consideration has a clearly defined weak subsystem which is fast (as in the example in reference 4), then the singular perturbational approach seems to give the best solution. Alternatively, if the weak subsystem is slow then the direct elimination method would be preferred. However, if the weak subsystem is either marginally fast or slow, then both these approaches should be investigated to determine the appropriate approximation which can be based on optimality, frequency response or other criteria.

Table 1: The reduced-order models and their error ratios

method	transfer function	error ratio
direct elimination (after Moore[1])	$\frac{-0.003127(s-23.14)}{(s+1.113)(s+2.460)}$	0.03938
reciprocal elimination	$\frac{-0.006808(s-12.29)}{(s+1.003)(s+3.158)}$	0.06931
'optimal'	$\frac{-0.003222(s-22.66)}{(s+1.099)(s+2.511)}$	0.03929

6. References

1. Moore, B.C.: 'Principal component analysis in linear systems: Controllability, observability, and model reduction', IEEE Trans. Automatic Control, 1981, AC-26, (1), pp.17-32.
2. Hutton, M.F., and Friedland, B.: 'Routh approximations for reducing order of linear, time-invariant systems', IEEE Trans. Automatic Control, 1975, AC-20, (3), pp.329-337.
3. Fernando, K.V., and Nicholson, H.: 'On the applicability of Routh approximations and allied methods in model-order reduction', IEEE Trans. Automatic Control, submitted for publication.
4. Fernando, K.V., and Nicholson, H.: 'Singular perturbational reduction of balanced systems', IEEE Trans. Automatic Control, 1982, AC-27, (2), pp.466-468.
5. Fernando, K.V., and Nicholson, H.: 'Singular perturbational model-order reduction in the frequency domain', IEEE Trans. Automatic Control, 1982, AC-27, (4).
6. Meier, III, L., and Luenberger, D.G.: 'Approximation of linear constant systems', IEEE Trans. Automatic Control, 1967, AC-12, pp.585-588.
7. Wilson, D.A.: 'Optimum solution of model-reduction problem', Proc.IEE, 117, (6), pp.1161-1165, 1970
8. Nicholson, H.: 'Structure of Interconnected Systems', Peregrinus, London, 1978.
9. Fernando, K.V., and Nicholson, H.: 'On the Cauchy index of linear systems', IEEE Trans. Automatic Control, 1983, AC-28, (3).
10. Gutman, P.O., Mannerfelt, C.F., and Molander, P.: 'Contributions to the model reduction problem', IEEE Trans. Automatic Control, 1982, AC-27, (2), pp.454-455.

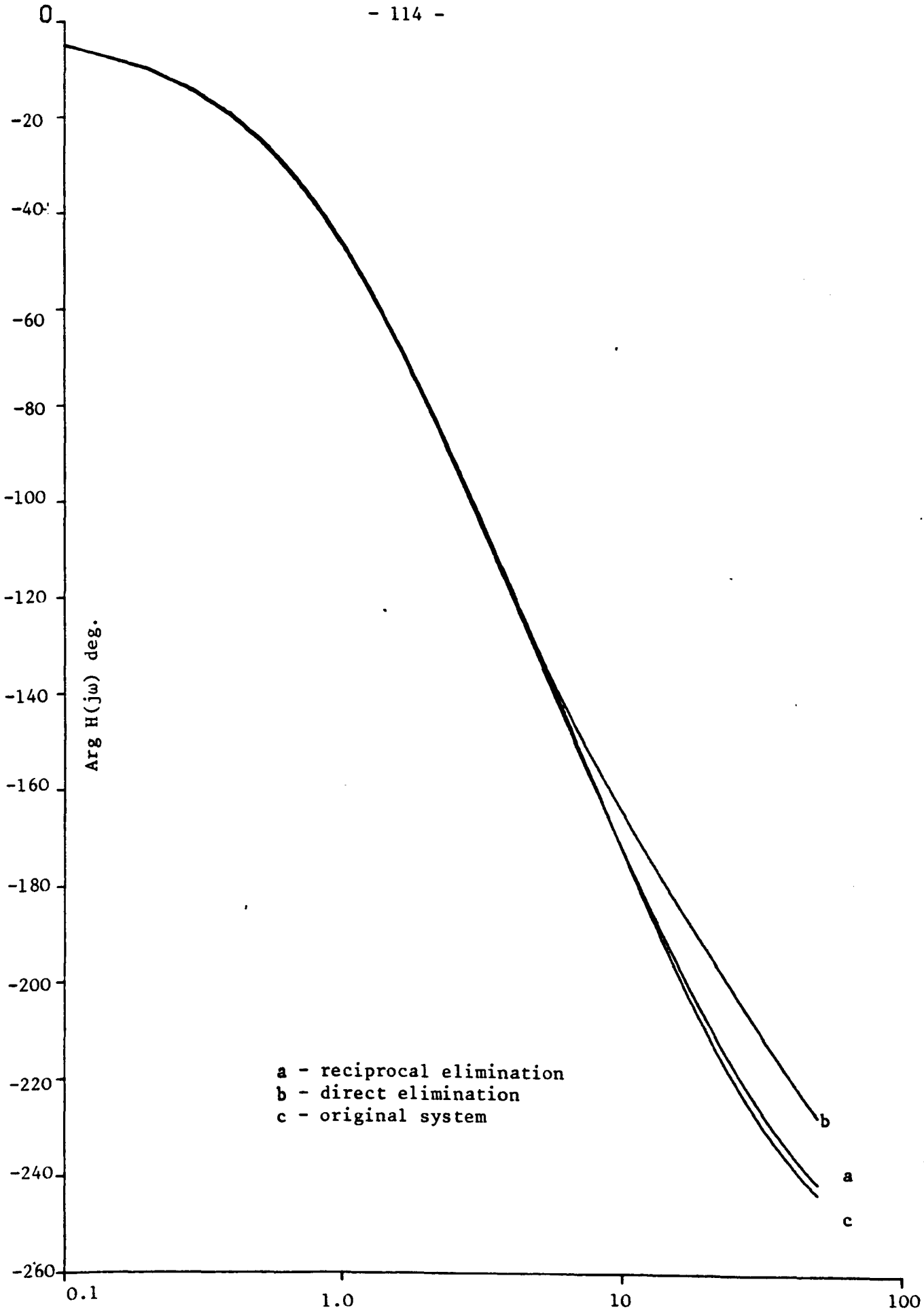


Fig. 2. Frequency Response : argument of H(jω)

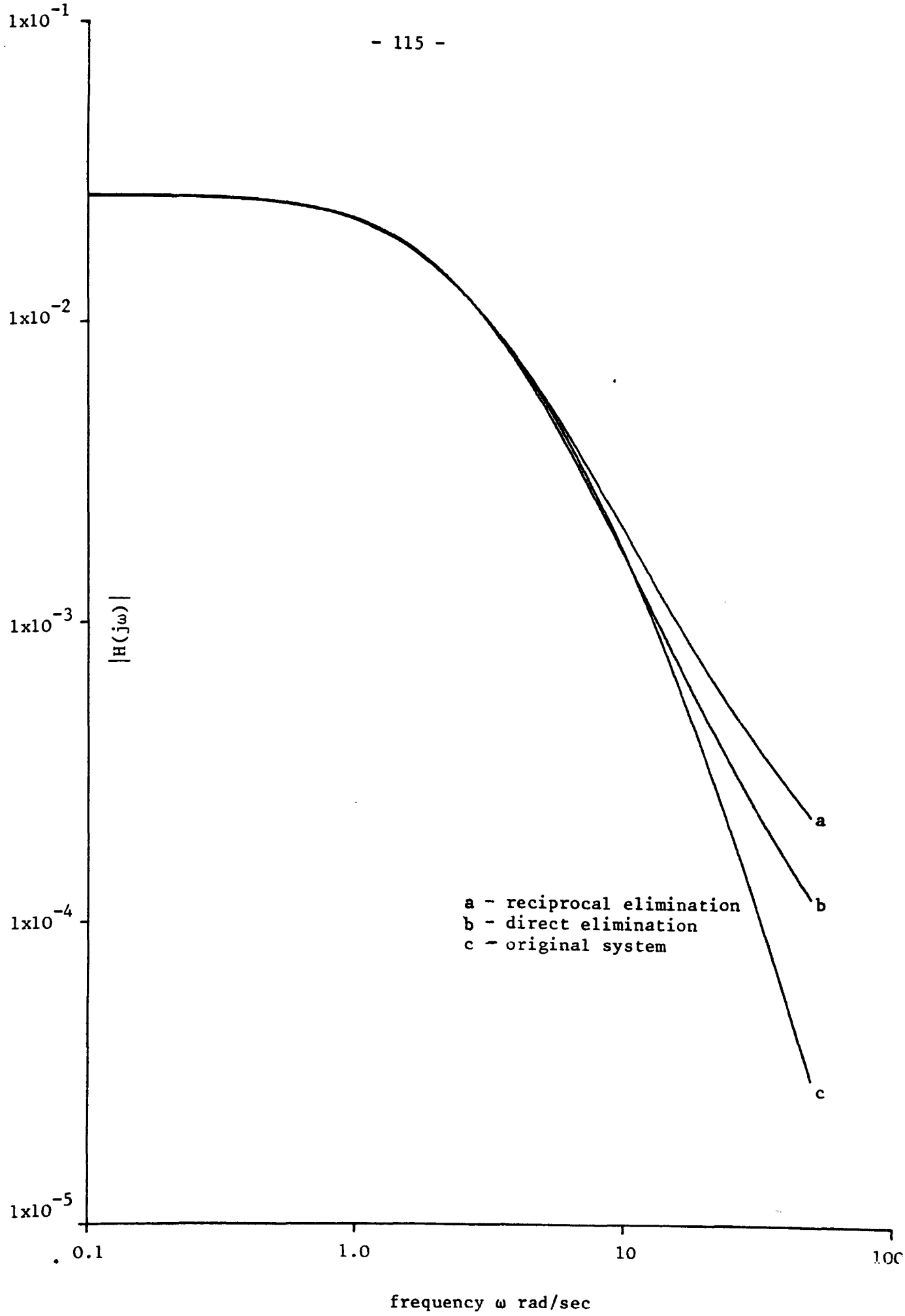


Fig. 1. Frequency Response : magnitude of H(jω)

PART 4

The Cross-Gramian Matrix W_{co} in Linear Systems Theory

CHAPTER 9

On the Structure of Balanced and Other Principal Representations of SISO Systems

Abstract: A new matrix W_{co} which can be considered as a cross-Gramian matrix which contains information about both controllability and observability is defined for single-input, single-output linear systems. Using this matrix, the structural properties of linear systems are studied in the context of principal component analysis. The matrix W_{co} can be used in obtaining balanced and other principal representations without computation of the controllability and the observability Gramians. The importance of this matrix in model-order reduction is highlighted.

1. Introduction

Moore[1] used concepts from the principal component analysis of Hotelling to investigate the controllability and observability of linear systems and also as a tool for model-order reduction. The technique is based essentially on simultaneous diagonalization of the controllability Gramian W_c[2] and the observability Gramian W_o[2] using appropriate similarity transformations. It was shown by Moore that it is somewhat inadequate and sometimes misleading to study controllability or observability, individually, and combined investigations are required.

In this Chapter, we define a new matrix W_{co} which can be considered as a cross-Gramian matrix and which carries information pertaining to controllability and observability, and which is directly connected to both controllability and observability Gramians. Thus, this new

matrix is a natural candidate for the study of combined investigations of controllability and observability and is used to expound the structure of SISO linear systems in the framework of principal component analysis.

Our main results are based on the absolute value symmetry of the state matrix under balanced conditions. However, the analysis could be carried out using more general principal (axis) representations[3] in which more specific balanced representations also belong. The role of principal representations in model-order reduction is also investigated.

The spectral structure of the matrix W_{co} is paramount in our analysis and in fact the absolute values of the spectrum are given by the singular values of the system. We also show the relationship between the singular values and the dc gain of the system and the importance of that result as an alternative criterion for model-order reduction.

2. Preliminaries

For the linear nth order single-input, single-output asymptotically stable time-invariant system S(A,b,c) described by

$$\dot{x}(t) = Ax(t) + bu(t) \quad , \qquad y(t) = cx(t)$$

the controllability Gramian matrix W_c^2 is defined as

$$W_c^2 = \int_0^T (e^{At}b)(e^{At}b)^T \, dt$$

where the term $e^{At}b$ represents the impulse response of the states of the system. We assume that the system is controllable and thus, W_c^2 is a positive definite matrix. Similarly, the observability Gramian matrix W_o^2 can be defined by considering the impulse response of the dual system.

$$\dot{x}_d(t) = A^T x_d(t) + c^T u(t)$$

with $$W_o^2 = \int_0^T (e^{A^T t_c T})(e^{A^T t_c T})^T \, dt$$

We also assume that the system is observable, resulting in a positive definite observability Gramian.

If the time interval of interest (0,T) is taken as infinite, then the Gramian matrices can be obtained by solving the following Lyapunov equations.

$$W_c^2 A^T + A W_c^2 = - bb^T \qquad (1)$$

$$W_o^2 A + A^T W_o^2 = - c^T c \qquad (2)$$

In the principal component analysis approach of Moore[1], the system S(A,b,c) is transformed into normalized and balanced forms which have one of the following constraints.

Input normal form : $W_c^2(P) = I$, $W_o^2(P) = \Sigma^4$

Output normal form : $W_c^2(P) = \Sigma^4$, $W_o^2(P) = I$

Internally balanced : $W_c^2(P) = W_o^2(P) = \Sigma^2$

The matrix P denotes the similarity transformation

$$S(A,b,c) \rightarrow S(P^{-1} AP, P^{-1} b, cP)$$

required to bring the Gramian matrices to normal or the balanced formats (see table 1). The matrix Σ^2 is diagonal and the diagonal positive elements are called the singular values of the system.

$$\Sigma^2 = \text{diag}\,(\sigma_1^2, \ldots, \sigma_n^2)$$

Table 1

The system under the similarity transformation P

original	transformed
A	$P^{-1}AP$
b	$P^{-1}b$
c	cP
W_c^2	$P^{-1}W_c^2(P^{-1})^T$
W_o^2	$P^TW_o^2P$
$W_c^2W_o^2$	$P^{-1}W_c^2W_o^2P$
W_{co}	$P^{-1}W_{co}P$

We assume that they are ordered in the non-increasing order of magnitude.

We observe that the normal and the internally balanced forms differ only by a diagonal similarity transformation. A more general format which encompasses the normal and the balanced forms can be defined in the following manner, which is called a principal axis representation [3].

Principal (axis) representation : $W_c^2(P) = \Sigma_c^2$, $W_o^2(P) = \Sigma_o^2 = \Sigma^4\Sigma_c^{-2}$

where Σ_c^2 and Σ_o^2 are positive diagonal matrices.

We note that one of the diagonal matrices Σ_c^2 or Σ_o^2 is arbitrary but not both. We denote internally balanced and principal representations by $S(\hat{A},\hat{b},\hat{c})$ and $S(\tilde{A},\tilde{b},\tilde{c})$, respectively.

3. The Symmetry in Internally Balanced Systems

Moore[1] referred to the absolute value symmetry of the state matrix A in single-input, single-output internally balanced systems. We present that property as a lemma.

Lemma 1: If the system $S(\hat{A},\hat{b},\hat{c})$ is internally balanced, then

(a) $\hat{b}_i = \pm\hat{c}_i$

(b) either $\hat{a}_{ij} = \hat{a}_{ji}$ or $\sigma_i^2 = \sigma_j^2$ if $\hat{b}_i\hat{b}_j = \hat{c}_i\hat{c}_j \neq 0$

(c) $\hat{a}_{ij} = -\hat{a}_{ji}$ if $\hat{b}_i\hat{b}_j = -\hat{c}_i\hat{c}_j \neq 0$

(d) either $\hat{a}_{ij} = -\hat{a}_{ji}$ and $\sigma_i^2 = \sigma_j^2$ or $\hat{a}_{ij} = \hat{a}_{ji} = 0$ if $\hat{b}_i\hat{b}_j = \hat{c}_i\hat{c}_j = 0$

where $\hat{a}_{ij}$ denotes the i,jth element of the matrix $\hat{A}$.

Proof: Since the Gramian matrices $W_c^2(P)$, $W_o^2(P)$ are diagonal and equal to Σ^2 the diagonal elements of (1) and (2) are of the form,

$$2\hat{a}_{ii}\sigma_i^2 = -\hat{b}_i^2 = -\hat{c}_i^2 , \quad \text{for } i = 1,n$$

and part (a) of the lemma is true.

The i,jth elements of (1) and (2) are given by

$$\sigma_i^2\hat{a}_{ji} + \hat{a}_{ij}\sigma_j^2 = -\hat{b}_i\hat{b}_j \tag{3}$$

$$\sigma_i^2\hat{a}_{ij} + \hat{a}_{ji}\sigma_j^2 = -\hat{c}_i\hat{c}_j \tag{4}$$

and the difference and the sum of (3) and (4) are of the form,

$$(\sigma_j^2 - \sigma_i^2)(\hat{a}_{ij} - \hat{a}_{ji}) = -\hat{b}_i\hat{b}_j + \hat{c}_i\hat{c}_j \tag{5}$$

$$(\sigma_j^2 + \sigma_i^2)(\hat{a}_{ij} + \hat{a}_{ji}) = -\hat{b}_i\hat{b}_j - \hat{c}_i\hat{c}_j \tag{6}$$

If $\hat{b}_i\hat{b}_j = \hat{c}_i\hat{c}_j \neq 0$, then from (5), the element $\hat{a}_{ij}$ appears symmetrically

in the matrix $\hat{A}$ or $\sigma_i^2 = \sigma_j^2$. If the second possibility $\hat{b}_i\hat{b}_j = -\hat{c}_i\hat{c}_j \neq 0$ is satisfied, then from (6), the element $\hat{a}_{ij}$ appears skew-symmetrically.

If the remaining possibility, $\hat{b}_i\hat{b}_j = \hat{c}_i\hat{c}_j = 0$ is true, then from (3) and (4) one possibility is $\hat{a}_{ij}^2 = \hat{a}_{ji}^2$. However, since σ_i^2 and σ_j^2 are positive, the condition $\hat{a}_{ij} = \hat{a}_{ji}$ is not admissible and hence $\hat{a}_{ij} = -\hat{a}_{ji}$ and $\sigma_i^2 = \sigma_j^2$. The other possibility is $\hat{a}_{ij} = \hat{a}_{ji} = 0$. □

The appearance of non-symmetrical or non-skewsymmetrical elements according to part (b) of lemma 1 is non-generic. The following result indicates that internally balanced formats can be found with absolute value symmetry.

Lemma 2: Let $S(\hat{A},\hat{b},\hat{c})$ be a balanced representation with non-distinct singular values of multiplicity two. If the state matrix $\hat{A}$ is not absolute value symmetric, then an orthogonal transformation can be found which transforms the representation to another balanced format with absolute value symmetry.

Proof: Without loss of generality assume that the first two singular values of the system $S(\hat{A},\hat{b},\hat{c})$ are non-distinct and thus, the two elements $\hat{a}_{12}$ and $\hat{a}_{21}$ do not appear symmetrically or skew-symmetrically. Now consider the orthogonal transformation Q where

$$Q = M \oplus I_{n-2} \quad , \quad M = \frac{1}{\sqrt{m^2+1}} \begin{pmatrix} 1 & m \\ -m & 1 \end{pmatrix}$$

The symbol $\oplus$ denotes the direct sum.

Under the orthogonal transformation the representation becomes,

$$S(\hat{A},\hat{b},\hat{c}) \rightarrow S(\hat{Q^T AQ},\hat{Q^T b},\hat{cQ}) = S(\bar{A},\bar{b},\bar{c})$$

It is easily seen that $S(\bar{A},\bar{b},\bar{c})$ is still internally balanced since the diagonal Gramian matrices are invariant under this transformation. $Q^T\Sigma^2Q = \Sigma^2$.

Again, lemma 1 guarantees the absolute value symmetry of the matrix $\bar{A}$ except the elements $\bar{a}_{12}$ and $\bar{a}_{21}$. These elements are given by

$$(1+m^2)\bar{a}_{12} = \hat{a}_{12} - (\hat{a}_{22}-\hat{a}_{11})m - \hat{a}_{21}m^2$$

$$(1+m^2)\bar{a}_{21} = \hat{a}_{21} - (\hat{a}_{22}-\hat{a}_{11})m - \hat{a}_{12}m^2$$

If we constrain these two elements to be skewsymmetric, then m should satisfy the quadratic equation given by

$$(\hat{a}_{12}+\hat{a}_{21})m^2 + 2(\hat{a}_{22}-\hat{a}_{11})m - (\hat{a}_{12}+\hat{a}_{21}) = 0$$

A real solution for m always exists for the above equation and the existence of a real orthogonal transformation with required properties is proved. □

Remark 1: Since, $\bar{a}_{21} = -\bar{a}_{12}$, then from (3) and (4)

$$\bar{b}_1 = \bar{c}_1 = 0 \quad \text{or} \quad \bar{b}_2 = \bar{c}_2 = 0$$

This shows that non-absolute value symmetry as considered in lemma 1 part (b) can be considered in part (d) of lemma 1, after the orthogonal transformation. □

Remark 2: According to part (b) of lemma 1, either $\hat{a}_{12} = \hat{a}_{21}$ or $\sigma_1^2 = \sigma_2^2$. However, a pathological case can occur where $\hat{a}_{12} = \hat{a}_{21}$ and $\sigma_1^2 = \sigma_2^2$. Then, under the orthogonal transformation,

$$\bar{a}_{12} = \bar{a}_{21} = 0 \quad , \text{ and either } \bar{a}_{11} \text{ or } \bar{a}_{22} \text{ is zero.}$$

If the system is second order, then it is not an asymptotically stable system which contradicts our assumption about stability. This argument seems to be correct even for higher order systems. □

In the remainder of this chapter, we assume that the internally balanced representation possesses the absolute value symmetry and that the system is not pathological (as defined in remark 2).

4. The Cross Gramian Matrix W_{co}

Using the impulse responses of the controllable system and the observable system, we define a matrix W_{co} as

$$W_{co} \triangleq \int_0^\infty (e^{At}b)(e^{A^T t}c^T)^T dt = \int_0^\infty e^{At}bce^{At}dt \qquad (7)$$

which we call the cross-Gramian matrix of the system. To our knowledge, this matrix W_{co} has not appeared previously in the control literature. Perhaps the closest analog is the cross-covariances in statistical analysis where the usual Gramian matrices could be considered as auto-covariances under appropriate white noise inputs instead of the usual impulse inputs.

It is easily seen that the matrix W_{co} can be computed by solving the linear matrix equation,

$$W_{co}A + AW_{co} = -bc \qquad (8)$$

Since the state matrix A is assumed to be stable, a unique solution matrix W_{co} exists. Standard algorithms are available for obtaining this solution[2].

It is intuitively clear that the matrix W_{co} carries information about both controllability and observability. This contrasts with the Gramian W_c[2] which contains controllability data only and the Gramian W_o[2] which contains observability data only.

It was shown by Moore[1] that it is inadequate and sometimes misleading to study controllability and observability, individually, and a combined approach is required in the analysis of dynamical systems. We will demonstrate that this cross-Gramian matrix W_{co} is consistent with the philosophy advocated by Moore and also fundamental to it.

It is easily seen from table 1 that the eigenvalues of the matrix W_{co} are invariant under similarity transformations of the system. We denote these eigenvalues by λ_i, $i = 1,n$ and form the diagonal eigenvalue matrix Λ as

$$\Lambda = \text{diag}\ (\lambda_1,\ldots,\lambda_n)$$

The following result shows the importance of the matrix W_{co} in the principal component analysis of SISO linear systems.

Theorem 1: If the system $S(\hat{A},\hat{b},\hat{c})$ is internally balanced with absolute value symmetry in the state matrix $\hat{A}$, then

(a) the corresponding cross-Gramian matrix $W_{co}(P)$ for the balanced system is diagonal and the diagonal elements are given by

$$\lambda_i = \sigma_i^2 \quad \text{if} \quad c_i = b_i \neq 0$$

$$\lambda_i = -\sigma_i^2 \quad \text{if} \quad c_i = -b_i$$

(b) the square of the matrix $W_{co}(\bar{P})$ is equal to the product $W_c^2(\bar{P})W_o^2(\bar{P})$ under any arbitrary similarity transformation $\bar{P}$. That is,

$$W_{co}^2 = W_c^2 W_o^2$$

Proof: Form the diagonal signature matrix U such that

$$u_i = 1 \quad \text{if} \quad \hat{c}_i = \hat{b}_i \neq 0$$

$$u_i = -1 \quad \text{if} \quad \hat{c}_i = -\hat{b}_i$$

If the system is internally balanced, (2) can be written in the form,

$$\Sigma^2\hat{A} + \hat{A}^T\Sigma^2 = -\hat{c}^T\hat{c}$$

and by premultiplying the above by U, we obtain

$$(U\Sigma^2)\hat{A} + (U\hat{A}^TU)(U\Sigma^2) = -\hat{b}\hat{c} \qquad (9)$$

The i,jth element of the matrix $U\hat{A}^TU$ is equal to $u_iu_j\hat{a}_{ji}$ and we consider all four possibilities.

$$\left.\begin{array}{l}\hat{c}_i = \hat{b}_i\ ,\ \hat{c}_j = \hat{b}_j \\ \hat{c}_i = -\hat{b}_i\ ,\ \hat{c}_j = -\hat{b}_j\end{array}\right\} \quad \hat{a}_{ij} = \hat{a}_{ji}\ ,\ u_iu_j\hat{a}_{ji} = \hat{a}_{ij}$$

$$\left.\begin{array}{l}\hat{c}_i = -\hat{b}_i\ ,\ \hat{c}_j = \hat{b}_j \\ \hat{c}_i = \hat{b}_i\ ,\ \hat{c}_j = -\hat{b}_j\end{array}\right\} \quad \hat{a}_{ij} = -\hat{a}_{ji}\ ,\ u_iu_j\hat{a}_{ji} = \hat{a}_{ij}$$

Thus, $U\hat{A}^TU = \hat{A}$, and by comparing (9) with (8), we obtain the required result (part a)

$$W_{co}(P) = U\Sigma^2 = \Lambda \qquad (10)$$

To prove part b, from (10) $W_{co}^2(P) = \Sigma^4$. However, $W_c^2(P)W_o^2(P) = \Sigma^4$ and hence

$$W_c^2(P)W_o^2(P) = W_{co}^2(P)$$

It is easily seen that this result is true even under any arbitrary similarity transformation $\bar{P}$, which completes the proof. □

The following result extends the above theorem concerning balanced systems to more general principal representations.

Corollary 1: If the system $S(\tilde{A},\tilde{b},\tilde{c})$ is a principal representation, then the corresponding matrix $W_{co}(P)$ is diagonal and the diagonal elements are given by

$$\lambda_i^2 = \sigma_i^2 \quad \text{if} \quad \text{sign}(\tilde{c}_i) = \text{sign}(\tilde{b}_i) \quad , \quad \tilde{b}_i = \tilde{c}_i \neq 0$$
$$\lambda_i^2 = -\sigma_i^2 \quad \text{if} \quad \text{sign}(\tilde{c}_i) = \text{sign}(-\tilde{b}_i)$$

Proof: We observe that any principal representation differs from the internally balanced format by a diagonal similarity transformation. However, the matrix $W_{co}(P)$ is equal to Λ when the system is internally balanced and it does not vary under diagonal similarity transformations. The rule for obtaining the signatures are obvious. □

The following result confirms the validity of the converse of the theorem 1 for principal representations.

Corollary 2: If the matrix $W_{co}(\bar{P})$ is diagonal, then the corresponding system $S(\bar{A},\bar{b},\bar{c}) = S(\bar{P}^{-1}\bar{A}\bar{P},\bar{P}^{-1}b,c\bar{P})$ is a principal representation.

Proof: If the matrix $W_{co}(\bar{P})$ is diagonal, then it is equal to the eigenvalue matrix (assuming that the diagonal elements are ordered in the non-increasing order of absolute value magnitude). It can be shown by substitution that

$$W_c^2(\bar{P}) = -\text{diag}(\ldots, \bar{b}_i^2/2\bar{a}_{ii},\ldots)$$
$$W_o^2(\bar{P}) = -\text{diag}(\ldots, \bar{c}_i^2/2\bar{a}_{ii},\ldots) \quad \text{for } \bar{a}_{ii} \neq 0$$

thus satisfying the conditions for principal representations. □

The following result indicates that the matrix W_{co} is well-conditioned even when there are non-distinct singular values.

Corollary 3: The eigenvalues of the matrix W_{co} are always distinct provided the multiplicity of singular values are at most two.

Proof: If σ_i^2 and σ_j^2 are equal, then due to lemma 2,

$$\hat{a}_{ij} = -\hat{a}_{ji}$$

However, due to theorem 1, $\lambda_i = \sigma_i^2$, $\lambda_j = -\sigma_i^2$

(or vice versa) and thus, the eigenvalues are distinct.

5. Model-order Reduction using Principal Representations

Moore[1] used internally balanced representations in model-order reduction based on subsystem elimination. However, as the following result indicates, we may use principal representations instead of internally balanced representations in model-order reduction and obtain the same reduced-order model.

Theorem 2: Let $S(\hat{A},\hat{b},\hat{c})$ be the internally balanced representation and $S(\tilde{A},\tilde{b},\tilde{c})$ be a principal representation with

$$S(\hat{A},\hat{b},\hat{c}) = S(D^{-1}\tilde{A}D, D^{-1}\tilde{b}, \tilde{c}D)$$

where D is an arbitrary diagonal matrix which defines all possible principal representations for that system (assuming that the matrix Λ is ordered in the non-increasing order of absolute value magnitude).

If the internally balanced representation and the principal representation are partitioned in the format,

$$\hat{A} = \begin{pmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{pmatrix}, \quad \hat{b} = \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \end{pmatrix}, \quad \hat{c} = \begin{bmatrix} \hat{c}_1 & \hat{c}_2 \end{bmatrix}$$

$$\tilde{A} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{pmatrix}, \quad \tilde{b} = \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{pmatrix}, \quad \tilde{c} = \begin{bmatrix} \tilde{c}_1 & \tilde{c}_2 \end{bmatrix}$$

with order of $\hat{A}_{11}$ = order of $\tilde{A}_{11}$ etc, then the balanced representation $S(\hat{A}_{11},\hat{b}_1,\hat{c}_1)$ and the principal representation $S(\tilde{A}_{11},\tilde{b}_1,\tilde{c}_1)$ describe the same reduced-order model.

Proof: If the matrix D is partitioned in the same format as,

$$D = D_1 \oplus D_2$$

then $$S(\hat{A}_{11},\hat{b}_1,\hat{c}_1) = S(D_1^{-1}\tilde{A}_{11}D_1, D_1^{-1}\tilde{b}_1, \tilde{c}_1D_1)$$

which completes the proof. □

If we decompose the matrices Σ^2 and Λ conforming to the partitions in theorem 2, then

$$\Sigma^2 = \Sigma_1^2 \oplus \Sigma_2^2 \quad , \quad \Lambda = \Lambda_1 \oplus \Lambda_2$$

Moore[1] used the trace of the diagonal Gramian matrix Σ_2^2 as a measure of error in model reduction. The trace of the matrix Σ^2, which is equal to the sum of the singular values, can be considered as the total "energy" of the system and relative error ratios can be computed in conjunction with this value.

The following result indicates that the trace of the matrix is related to the dc gain of the system.

Theorem 3: The sum of the eigenvalues λ_i, i = 1,n gives half the dc gain of the system. That is

$$\text{trace } \Lambda = -\tfrac{1}{2}\, cA^{-1}b$$

Proof: $W_{co} = \int_0^\infty e^{At}bce^{At}dt$

$$\text{trace } W_{co} = \int_0^\infty ce^{2At}b\, dt = -\tfrac{1}{2}\, cA^{-1}b$$

Since, trace W_{co} = trace Λ, the result follows. □

An alternative criterion for model reduction can be stated using the above theorem. Instead of the requirement that the trace of the matrix Σ_2 is "small", we may specify that the dc gain of the subsystem $S(A_{22},b_2,c_2)$ given by twice the trace of the matrix Λ_2 should be small. The smallness of the singular values guarantees low dc gains, however the converse is not necessarily true.

This is a significant result since in most established model-order reduction methods, the dc gain is one of the criteria for obtaining

reduced order models. Often, it is specified that the reduced order model and the original model should have the same dc gain and due to the direct relationship between the singular values and the dc gain, it may be possible to take such contraints into account in this approach.

6. Computation of Principal Representations

Since principal representations are equally useful as the balanced representation in model-order reduction, we may use principal instead of balanced representations. Computation of balanced representations normally require[1] the solution of Lyapunov equations (1) and (2) and a series of spectral decompositions and similarity transformations. However, for SISO systems we suggest the following numerical procedure which requires less computational effort.

(a) Compute the matrix W_{co} as the solution of (8) using any standard algorithm[2]

(b) Compute the real spectral decomposition of the matrix W_{co}, to give

$$W_{co} = V\Lambda V^{-1}$$

where V is an eigenvector matrix and Λ is the diagonal eigenvalue matrix.

(c) Compute the principal representation $S(\tilde{A},\tilde{b},\tilde{c})$ given by

$$S(\tilde{A},\tilde{b},\tilde{c}) = S(V^{-1}AV,\ V^{-1}b, cV)$$

This step is obvious from corollary 2.

(d) (Optional step) If a balanced representation is required, we may use the diagonal similarity transformation $\bar{P}$, defined by,

$$\bar{P} = \text{diag}(\ldots, \bar{p}_i, \ldots)$$

where $\bar{p}_i = (u_i \tilde{b}_i / \tilde{c}_i)^{\frac{1}{2}}$ if $\tilde{c}_i \neq 0$

$\bar{p}_i = 1$ if $\tilde{c}_i = 0$ □

We have used the above numerical procedure on a number of problems including the examples described by Moore[1] and identical results were obtained.

In references 3 and 4 and elsewhere, computation of principal representations is based on diagonalization of the matrix product $W_c^2 W_o^2$ which is equal to W_{co}^2. However, W_{co}^2 can have non-distinct eigenvalues while the eigenvalues of the matrix W_{co} are distinct (corollary 3). Thus, the spectral decomposition problem associated with W_{co} is always well-conditioned, which may not be the case with W_{co}^2. Also, our approach avoids formation of the product $W_c^2 W_o^2$ and is well-conditioned with respect to round-off errors.

7. Conclusions

We have defined a new matrix W_{co} which can be considered as a cross-Gramian matrix and which contains information pertaining to both controllability and observability. Using this matrix, the structure of SISO linear systems in the context of principal component analysis has been studied. It was shown that its properties can be used in model-order reduction in the framework of more general principal representations without computing the more specific balanced representations. However, both principal and balanced representations give the same reduced-order model.

Due to inherent signatures, the cross-Gramian matrix W_{co} contains more information than the controllability and observability Gramian matrices. In fact, it can be shown that W_{co} and the Hankel matrix associated with the same system, share common properties including the Cauchy index[5] (Chapter 11).

We have also proved the relationship between the singular values of the system and the dc gain. It was explained how this property can be used as an alternative criterion in model-order reduction.

8. References

1. Moore, B.C.: 'Principal component analysis in linear systems: Controllability, observability, and model reduction', IEEE Trans. Automatic Control, 1981, AC-26, (1), pp.17-32.
2. Bartels, R.H., and Stewart, G.W.: 'Solution of the matrix equation AX + XB = C', Comm. of the ACM, 1972, 15, (9), pp.820-826.
3. Mullis, C.T., and Roberts, R.A.: 'Synthesis of minimum roundoff noise fixed point digital filters', IEEE Trans. Circuits and Systems, 1976, CAS-23, (9), pp.551-561.
4. Silverman, L.M., and Bettayeb, M.: 'Optimal approximation of linear systems', in Proc. 1980 Joint Automatic Control Conference, San Francisco, CA, paper FA8-A, August 1980.
5. Fernando, K.V., and Nicholson, H.: 'On the Cauchy index of linear systems', IEEE Trans. Automatic Control, 1983, AC-28, (3).

CHAPTER 10

Minimality of SISO Linear Systems

Abstract: A new test for minimality of single-input single-output state-space realizations is proposed based on rank conditions of a cross-Gramian type matrix without computing the controllability and observability Gramian matrices.

1. Introduction

The problem of determining minimality of state-space representations of linear systems is of fundamental importance in modern control and general systems theory. Much effort (see for example Kailath[2]) has gone into investigations of minimality since the pioneering work of Kalman et al[1] and other workers. If a system representation is minimal, then the system is jointly controllable and observable and hence minimality tests are based on controllability and observability criteria.

Fernando and Nicholson[3] (see Chapter 9) defined a cross-Gramian matrix W_{co} which contains information pertaining to both controllability and observability. It was shown that this matrix W_{co} can be used in deriving internally balanced and other principal axis realizations of SISO systems without computing the controllability Gramian matrix W_c and the observability Gramian matrix W_o. The object of this Chapter is to show that if the matrix W_{co} is of full rank, then the realization is minimal.

2. Preliminaries

For the asymptotically stable SISO time-invariant system S(A,b,c) defined by

$$\dot{x}(t) = Ax(t) + bu(t) \quad , \quad y(t) = cx(t)$$

the controllability and observability Gramian matrices[1,2] can be defined in the infinite-time interval as,

$$W_c = \int_0^\infty (e^{At}b)(e^{At}b)^T \, dt \tag{1}$$

$$W_o = \int_0^\infty (ce^{At})^T(ce^{At}) \, dt \tag{2}$$

These Gramian matrices can be also obtained by solution of the Lyapunov equations given by,

$$W_c A^T + AW_c = - bb^T \tag{3}$$

$$W_o A + A^T W_o = - c^T c \tag{4}$$

If (3) and (4) are used as definitions of the Gramian matrices, instead of (1) and (2), then the assumptions regarding the asymptotic stability is not required provided that there are no eigenvalues of the state matrix A such that,

$$\lambda_i(A) + \lambda_j(A) = 0 \qquad \text{for} \qquad i = 1,n \ , \ j = 1,n$$

where n is the order of the system. The above condition also guarantees the uniqueness of the solutions W_c and W_o of (3) and (4).

Fernando and Nicholson[3] (see Chapter 9) defined the cross-Gramian matrix W_{co} as

$$W_{co} = \int_0^\infty (e^{At}b)(ce^{At}) \, dt$$

This matrix can be obtained as the solution of the matrix equation given by

$$W_{co}A + AW_{co} = - bc \qquad (5)$$

which can be considered as an equivalent and a more general definition. We observe that there are standard algorithms[4] for solution of (5).

3. The results

We present the following result which relates the three Gramian matrices.

Proposition 1: $W_c W_o = W_{co}^2$

Proof: For asymptotically stable systems, an indirect proof was provided in Chapter 9. However, a simple proof is possible provided the state matrix A is semi-simple. Using the eigenvector matrix U of the matrix A as a similarity transformation, we may obtain the equivalent system $S(\tilde{A},\tilde{b},\tilde{c})$ where $\tilde{A}$ is diagonal.

$$S(\tilde{A},\tilde{b},\tilde{c}) = S(U^{-1}AU, U^{-1}b, cU)$$

The representation $S(\tilde{A},\tilde{b},\tilde{c})$ is in general complex.

For the transformed system, the i,jth elements of the Gramian matrices are given by,

$$(\tilde{W}_c)_{ij} = - \frac{\tilde{b}_i \tilde{b}_j}{\tilde{a}_{ii} + \tilde{a}_{jj}}$$

$$(\tilde{W}_o)_{ij} = - \frac{\tilde{c}_i \tilde{c}_j}{\tilde{a}_{ii} + \tilde{a}_{jj}}$$

$$(\tilde{W}_{co})_{ij} = -\frac{\tilde{b}_i \tilde{c}_j}{\tilde{a}_{ii} + \tilde{a}_{jj}}$$

and thus, $$\tilde{W}_c \tilde{W}_o = \tilde{W}_{co}^2$$

which provides the required result through the similarity transformation.

If the controllability Gramian matrix W_c is of full rank, then it is well known that the system is completely controllable [1,2]. Similarly, if the observability Gramian matrix W_o is of full rank, then it guarantees complete observability. The following result is about joint controllability and observability.

Proposition 2: If the cross-Gramian matrix W_{co} is of full rank, then the system is completely controllable and observable, and hence a minimal realization. This condition is both necessary and sufficient.

Proof: This results as a direct consequence of the relationship between Gramian matrices as indicated in proposition 1.

Thus, proposition 2 provides a direct method of determining minimality of SISO realizations by checking the rank of the matrix W_{co}. There is considerable savings in computation since we have to solve only one matrix equation for W_{co} rather than two matrix equations for W_c and W_o. Furthermore, we have to check only one matrix for rank rather than two.

4. The Discrete-time Problem

The results in section 3, can be extended to the discrete-time system $S_d(A,b,c)$ defined by

$$x_{t+1} = Ax_t + bu_t \quad , \quad y_t = cx_t$$

Instead of equations (3), (4), and (5), the Gramian matrices are defined by equations (6), (7), and (8), respectively.

$$W_c - AW_cA^T = bb^T \tag{6}$$

$$W_o - A^TW_oA = c^Tc \tag{7}$$

$$W_{co} - AW_{co}A = bc \tag{8}$$

We also assume that there are no eigenvalues of the state matrix A such that

$$\lambda_i(A)\lambda_j(A) = 1 \quad \text{for} \quad i = 1,n \ , \ j = 1,n$$

It can be shown easily that propositions 1 and 2 are also true for the discrete-time case.

5. Conclusions

We have developed a simple test for minimality of SISO realizations based on joint controllability and observability as characterized by the cross-Gramian matrix W_{co} which can be obtained as the solution of single matrix equations.

6. References

1. Kalman, R.E., Ho, Y.C., and Narendra, K.S.: 'Controllability of linear dynamical systems', Contributions to differential equations, 1962, 1, (2), pp.189-213.

2. Kailath, T.: 'Linear Systems', Prentice-Hall, Englewood Cliffs, NJ, 1980.

3. Fernando, K.V., and Nicholson, H.: 'On the structure of balanced and other principal representations of SISO systems', IEEE Trans. Automatic Control, 1983, AC-28, (3), (tentative publication date).

4. Golub, G.H., Nash, S., Van Loan, C.: 'A Hessenberg-Schur method for the problem AX + XB = C', IEEE Trans. Automatic Control, 1979, AC-24, (4), pp.903-913.

CHAPTER 11

On the Cauchy Index of Linear Systems

Abstract: The Cauchy index for linear SISO systems is given by the signature of the cross-Gramian matrix W_{co} defined by Fernando and Nicholson[1]. This index is useful in the qualitative understanding of systems and the importance of this index in internally balanced model reduction is illustrated.

1. Introduction

The Cauchy index is one of the fundamental parameters available in the study of rational functions and thus naturally important in studies involving transfer functions of both continuous-time and discrete-time linear systems[2,3,5]. This index is given by the signature of the associated Hankel matrix of the rational function and is especially useful in the characterisation of systems with respect to the structure of poles and zeros. For a system with distinct poles, the Cauchy index is equal to the number of real poles with positive residues minus the number of poles with negative residues.

Fernando and Nicholson[1] (Chapter 9) defined a cross-Gramian matrix W_{co} which contains information about both controllability and observability. Due to this property, information contained in the controllability Gramian W_c^2 and the observability Gramian W_o^2 become redundant if the matrix W_{co} is known. The object of this Chapter is to show that the Cauchy index is given by the signature of the

cross-Gramian matrix W_{co}. We also demonstrate the similarities between the Hankel matrix and the cross-Gramian matrix.

The Cauchy index is particularly useful in system identification[2] and in model-order reduction. Structural changes occur in the model-order reduction process, and by knowing the Cauchy index at each level, the structural changes which accompany the reduction process can be predicted. A numerical example is included to illustrate the importance of the Cauchy index in internally balanced model-order reduction[4].

2. The Gramian and Hankel Matrices

For the linear nth order time-invariant asymptotically stable continuous-time system $S_c(A,b,c)$,

$$\dot{x}(t) = Ax(t) + bu(t) \quad , \quad y(t) = cx(t)$$

the controllability Gramian matrix W_c^2 and the observability Gramian matrix W_o^2 are defined as,

$$W_c^2(T) = \int_0^T (e^{At}b)(e^{At}b)^T dt \quad , \quad T \geq T_o \qquad (1)$$

$$W_o^2(T) = \int_0^T (ce^{At})^T(ce^{At})dt \quad , \quad T \geq T_o \qquad (2)$$

Fernando and Nicholson[1] (Chapter 9) demonstrated that the information in the controllability Gramian and the observability Gramian are essentially contained in the cross-Gramian matrix W_{co} defined by

$$W_{co}(T) = \int_0^T (e^{At}b)(ce^{At})dt \quad , \quad T \geq T_o \qquad (3)$$

In a stochastic formulation of the problem, this matrix can be considered as a cross-covariance between the controllable states and the observable states, while the controllable and the observable Gramians can be considered as auto-covariances.

For the discrete-time system $S_d(A,b,c)$,

$$x_{t+1} = Ax_t + Bu_t \quad , \quad y_t = Cx_t$$

the Gramian matrices are defined by,

$$W_c^2(p) = \sum_{k=0}^{p} (A^k b)(A^k b)^T \quad , \quad p \geq n \tag{4}$$

$$W_o^2(p) = \sum_{k=0}^{p} (cA^k)^T (cA^k) \tag{5}$$

In an analogous manner to the continuous-time case, we may define the W_{co} matrix as,

$$W_{co}(p) = \sum_{k=0}^{p} (A^k b)(cA^k) \tag{6}$$

The Gramian matrices are directly related to the controllability and observability matrices of Kalman, thus

$$W_o^2(p) = Q_o^T(p) Q_o(p) \tag{7}$$

$$W_c^2(p) = Q_c(p) Q_c^T(p) \tag{8}$$

and $$W_{co}(p) = Q_c(p) Q_o(p) \tag{9}$$

where $$Q_c(p) = \begin{bmatrix} b & Ab & \cdots & A^{p-1}b \end{bmatrix}$$

$$Q_o^T(p) = \begin{bmatrix} c^T & (cA)^T & \cdots & (cA^{p-1})^T \end{bmatrix}$$

The (symmetric) Hankel matrix H(p) can be defined using these controllability and observability matrices, and is given by

$$H(p) = Q_o(p)Q_c(p) \qquad (10)$$

The similarity between the Hankel matrix H(p) and the cross-Gramian matrix $W_{co}(p)$ is obvious. The Hankel matrix may be considered as an "outer product" of the observability and the controllability matrices while the cross-Gramian matrix manifests as an "inner product". This suggests that there is a direct duality between the Hankel matrix and the cross-Gramian matrix.

We observe that the definitions used in the discrete-time case are also valid in the continuous-time case and provide alternative but equivalent definitions for Gramians.

3. The Cauchy Index

For the proper rational transfer function

$$g(s) = \sum_{k=1}^{\infty} h_k s^{-k}$$

the associated Hankel matrix is defined as,

$$H(p) = \begin{pmatrix} h_1 & h_2 & \cdots\cdots & h_p \\ h_2 & h_3 & \cdots\cdots & h_{p+1} \\ h_p & h_{p+1} & \cdots\cdots & h_{2p-1} \end{pmatrix}$$

If this transfer function is of the dynamical system, S_c or S_d, then the Markov parameters h_k are given by,

$$h_k = cA^{k-1}b$$

We reproduce the fundamental theorem of Hermite and Hurwitz[2,3] which relates the Cauchy index with the signature of the Hankel matrix. (The signature of a matrix is the number of positive real eigenvalues minus the number of real negative eigenvalues).

Theorem 1: The Cauchy index of g(s) is equal to the signature of the associated Hankel matrix.

A simple "control theoretic" proof was provided by Brockett[2] which illuminates the essentials of the theorem.

The following result which is often quoted in minimal realization theory and circuit analysis is due to Anderson[5].

Theorem 2: For the system S_c or S_d, there exists a unique symmetric matrix R satisfying

$$RA = A^T R \quad , \qquad Rb = c^T$$

and the signature of R is equal to the Cauchy index of the system. □

4. The Cauchy Index and the Cross-Gramian Matrices

We propose the following two results which relate the Cauchy index to the signature of the cross-Gramian matrices.

Proposition 1: The Cauchy index of the system S_c or S_d is given by the signature of the cross-Gramian matrix $W_{co}(p)$.

Proof: It is obvious from (9) and (10) that all non-zero eigenvalues of the Hankel matrix H(p) are given by the eigenvalues of the cross-Gramian matrix $W_{co}(p)$. Thus, the signatures of these matrices are equal and hence the result. □

Proposition 2: The Cauchy index of the system S_c is given by the signature of the cross-Gramian matrix $W_{co}(T)$.

Proof: From theorem 2 (see 5 for details),

$$R\, Q_c(n) = Q_o^T(n)$$

As a direct consequence of the Cayley-Hamilton theorem

$$R\, e^{At}b = e^{A^T t}c^T$$

Post-multiplying the above by $b^T e^{A^T t}$ and integrating in the interval (0,T), we obtain

$$R\, W_c^2(T) = W_{co}^T(T)$$

Since, $W_c^2(T)$ is positive definite, the result is obvious.

We have included the cross-Gramian matrices $W_{co}(p)$ and $W_{co}(T)$ in the list of matrices which contain information about the Cauchy index. We now proceed to study the Cauchy index in relation to internally balanced representations.

5. The Cauchy Index and Balanced Representations

If the controllability Gramian matrix $\tilde{W}_c^2(T)$ and the observability Gramian matrix $\tilde{W}_o^2(T)$ of the system $S_c(\tilde{A},\tilde{b},\tilde{c})$ are diagonal and equal, then the system $S_c(\tilde{A},\tilde{b},\tilde{c})$ is said to be an internally balanced representation[4]. Such balanced representations can be obtained by using similarity transformations on any equivalent realization $S_c(A,b,c)$. The following result is due to Fernando and Nicholson[1] (Chapter 9).

Theorem 3: For the system $S_c(A,b,c)$, a balanced representation $S_c(\tilde{A},\tilde{b},\tilde{c})$ (with $T\to\infty$) exists such that

(a) the elements of the vectors $\tilde{b}$ and $\tilde{c}$ are equal in magnitude, that is

$$\tilde{b}_i = \tilde{c}_i \quad \text{or} \quad \tilde{b}_i = -\tilde{c}_i \qquad \text{for all } i.$$

(b) the cross-Gramian matrix $W_{co}(\infty) \triangleq \Lambda$ is diagonal.

(c) If $\tilde{b}_i = \tilde{c}_i \neq 0$, then $\lambda_i = \sigma_i^2$

If $\tilde{b}_i = -\tilde{c}_i$, then $\lambda_i = -\sigma_i^2$

where $\tilde{W}_{co}(\infty) = \Lambda = \text{diag}(\ldots\ldots, \lambda_i, \ldots\ldots)$

$\tilde{W}_c^2(\infty) = \tilde{W}_o^2(\infty) = \Sigma^2 = \text{diag}(\ldots\ldots, \sigma_i^2, \ldots\ldots)$

(d) If $\tilde{b}_i\tilde{b}_j = \tilde{c}_i\tilde{c}_j \neq 0$, then $\tilde{a}_{ij} = \tilde{a}_{ji}$

If $\tilde{b}_i\tilde{b}_j = -\tilde{c}_i\tilde{c}_j$, then $\tilde{a}_{ij} = -\tilde{a}_{ji}$ □

The following result gives an explicit solution for the matrix R (defined in Theorem 2) for balanced representations.

Proposition 3: For the system $S_c(\tilde{A},\tilde{b},\tilde{c})$, the matrix R is diagonal and the diagonal elements are given by

$$\tilde{r}_i = 1 \quad \text{if} \quad \tilde{b}_i = \tilde{c}_i$$
$$\tilde{r}_i = -1 \quad \text{if} \quad \tilde{b}_i = -\tilde{c}_i$$

Proof: Obvious from Theorem 3. □

Proposition 3 provides a convenient way of determining the Cauchy index of balanced representations by inspection. The validity of Proposition 2 is also obvious from the above result.

6. Application in Model Reduction: An Example

The following transfer function is used by Moore[4] in internally balanced model-order reduction.

$$g(s) = \frac{(s+4)}{(s+1)(s+3)(s+5)(s+10)}$$

$$= \frac{1}{24(s+1)} - \frac{1}{28(s+3)} - \frac{1}{40(s+5)} + \frac{2}{105(s+10)}$$

Thus, the Cauchy index for this system is zero which is one of the five possible values (±4, ±2, 0) for a fourth-order system.

If the system is realized in the internally balanced format, the controllability and the observability Gramians are diagonal, and equal, and given by[4] (for the infinite-time definition of Gramians)

$$\tilde{W}_c^2 = \tilde{W}_o^2 = \Sigma^2 = \text{diag}(0.0159,\ 0.272\text{x}10^{-2},\ 0.126\text{x}10^{-3},\ 0.8\text{x}10^{-5})$$

The cross-Gramian matrix W_{co} is also diagonal and equal to the balanced Gramians except for the signature (Theorem 3).

$$\tilde{W}_{co} = \Lambda = \text{diag}(0.0159,\ -0.272\text{x}10^{-2},\ 0.127\text{x}10^{-3},\ -0.8\text{x}10^{-5})$$

Since the signature of the matrix $\tilde{W}_{co}$ is zero, from Proposition 2, we know that the Cauchy index is zero.

In internally balanced model-order reduction, the reduced model is obtained by subsystem elimination and the dominant diagonal values (the second-order modes) are retained in the reduction process. Thus, by inspecting the diagonal values of $\tilde{W}_{co}$, the Cauchy indices of the reduced-order models can be inferred. Thus,

order	4	3	2	1
Cauchy index	0	1	0	1

We now consider the transfer function

$$\bar{g}(s) = \frac{1}{24(s+1)} + \frac{1}{28(s+3)} + \frac{1}{40(s+5)} + \frac{2}{105(s+10)}$$

which is equal to the transfer function g(s) except for the positions of the zeros.

The Cauchy index for this transfer function is four, and thus the signature of the cross-Gramian matrix $\bar{W}_{co}$ is also four. The Cauchy indices for the reduced-order models of the transfer function $\bar{g}(s)$ are given by,

order	4	3	2	1
Cauchy index	4	3	2	1

We observe that if the absolute value of the Cauchy number is equal to the system order[2]

(a) the system has real poles only

(b) the system is not non-minimal phase.

Thus, the reduced-order models of $\bar{g}(s)$ are not non-minimal phase. However, such inference cannot be used for the transfer function g(s), since the Cauchy index is zero. Thus, the occurrence of right-hand plane zeros (in the reduction process) cannot be overruled, although the transfer function g(s) is minimal phase. In fact, the second- and third-order reduced models obtained by Moore have right-hand plane zeros. However, if the second-order approximation is obtained by eliminating the second and the fourth states (instead of the third and fourth), the reduced model will not have right-hand plane zeros.

We have deduced this information by inspecting the partial fraction expansion of the transfer function $\bar{g}(s)$. However, in large-scale system studies, it is much more convenient to compute the matrix W_{co} and then by inspecting the signature of the matrix W_{co} we may obtain the same results.

7. Conclusions

We have related the Cauchy index to the cross-Gramian matrices and shown the important structural information given by the Cauchy index. The Cauchy index is classically associated with the Hankel matrix and we have proved that the information contained in the Hankel matrix is essentially present in the cross-Gramian matrices.

The wealth of structural information provided by the Cauchy index in the context of internally balanced model-order reduction has been amply demonstrated using a standard example.

8. References

1. Fernando, K.V., and Nicholson, H.: 'On the structure of balanced and other principal representations of SISO systems', IEEE Trans. Automatic Control, 1983, AC-28, (3).

2. Brockett, R.W.: 'Some geometric questions in the theory of linear systems', IEEE Trans. Automatic Control, 1976, AC-21, (4), pp.449-455.

3. Gantmacher, F.R.: 'Theory of matrices', vol. II, Chelsea, New York, 1976.

4. Moore, B.C.: 'Principal component analysis in linear systems: Controllability, observability, and model reduction', IEEE Trans. Automatic Control, 1981, AC-26, (1), pp.17-32.

5. Anderson, B.D.O.: 'On the computation of the Cauchy index', Quart. Appl. Math., 1972, 29, pp.577-582.

PART 5

Measures for Quantification of Controllability, and Observability, and Input-Output Behaviour

CHAPTER 12

The Degree of Controllability due to Individual Inputs

Abstract: Mahalanobis distance, which is an information theoretic metric measure, can be used as an index to investigate the effectiveness of individual inputs in multivariable control systems.

1. Introduction

Since the introduction of the concept of controllability for linear dynamical systems by Kalman two decades ago, much effort has been devoted to quantifying controllability. The natural candidate for such an index has been the controllability Gramian matrix (W-matrix),

$$W(\tau) = \int_0^\tau e^{At} B B^T e^{A^T t} dt$$

for the linear time-invariant dynamical system defined by

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

$$x \in M_{n,1} \quad , \quad u \in M_{m,1} \quad , \quad A \in M_{n,n} \quad , \quad B \in M_{n,m}$$

Kalman et al[1] suggested that scalar functions of the W-matrix, namely the determinant and the trace of the inverse of the W-matrix, are suitable for identifying the degree of controllability. These functions are related to the minimal energy problem and physical interpretations are possible[2].

Kalman et al also realized the importance of the W-matrix as an information theoretic measure and related it to the Shannon definition of information and the Fisher information matrix. Later, Mitra[3] investigated the W-matrix using information quantifiers due to Kolmogorov and Shannon in the context of model order reduction.

However, the basic weakness of this approach is that the W-matrix is not invariant under similarity transformations of the system (A,B). This was overcome by Friedland[4] by defining the degree of controllability as the condition number of the W-matrix given by

$$k(W) = \|W\| \cdot \|W^{-1}\|$$

If $\|\cdot\|$ is taken as the spectral norm of W, the condition number is then given by

$$k(W) = |\lambda_{max}| \; / \; |\lambda_{min}|$$

where λ_{max} and λ_{min} are the eigenvalues of W having the largest, and the smallest magnitudes, respectively. The basic notion for this definition comes from numerical algebra, however a physical interpretation can be given using Raleigh quotients[5,6]. Recently, Denham[7] used angles between subspaces, which is a non-metric information measure, for inter-relating inputs and outputs of large scale systems.

The objective of this Chapter is to develop measures to quantify the degree of controllability of individual inputs in multi-input linear systems. This is achieved by using a metric information measure known as the Mahalanobis distance[8,9].

2. The Mahalanobis distance[8,9]

It is well known that for a Gaussian vector z, the probability density function is given by

$$p(z) = \alpha \exp(-\tfrac{1}{2} \|z-\bar{z}\|^2_{\phi^{-1}})$$

where $\alpha = (2\pi)^{-n/2}(\det \phi)^{-\frac{1}{2}}$

$$\bar{z} = E[z] \quad , \qquad \phi = E[(z-\bar{z})(z-\bar{z})^T]$$

$E|\cdot|$ denotes the expectation operator. If we take the logarithm of the function, ignoring the constant scalar, we may define a new function

$$M^2(z) = \|z-\bar{z}\|^2_{\phi^{-1}} = (z-\bar{z})^T \phi^{-1} (z-\bar{z})$$

which is still a measure of probability associated with the random vector z. This measure is sometimes known as the Mahalanobis distance from the mean.

For a particular class of random vectors denoted by z^i, belonging to the class S^i, which is a subset of the general class S,

$$z^i \in S^i \quad , \quad z \in S \quad , \quad S^i \subseteq S \quad , \quad i = 1,m$$

the Mahalanobis distance $M(z^i,\bar{z}^i)$ defined by

$$M^2(z^i,\bar{z}^i) = \|z^i-\bar{z}^i\|^2_{\phi^{-1}}$$

is a measure of 'oscillatory energy' of the vector z^i. The loci of points of constant energy defines a hyper-ellipsoid with the principal axes in the directions of the eigenvectors of the covariance matrix ϕ.

The lengths of the semi-axes are given by square roots of the eigenvalues. Thus, the role of the matrix ϕ^{-1} is to de-weight heavily the eigenvector 'modes' which are powerfully represented in the set S, and lightly the 'modes' which are not. The expected value of this measure is given by

$$E\left[M^2(z^i,\bar{z}^i)\right] = \text{trace } \phi^{-1}\phi^i$$

where the matrix ϕ^i is the covariance of the vector z^i, thus

$$\phi^i = E\left[(z^i-\bar{z}^i)(z^i-\bar{z}^i)^T\right]$$

Although, this is a distance measure, due to the form of weighting by the matrix ϕ^{-1}, it is independent of the unit of measurement (i.e. dimensionless). Thus, it can be considered as the fraction of energy in the subset S^i with respect to the set S.

3. The degree of controllability due to individual inputs

We assume that the linear system (A,B) is asymptotically stable and fully controllable. These may be relaxed in a more formal study. The W-matrix can be considered as the Gramian matrix

$$W(\tau) = \int_0^\tau x(t)x^T(t)dt$$

for deterministic unit impulses at the inputs. For stochastic inputs of the form

$$E\left[u(t)\right] = 0 \quad , \quad E\left[u(t)u^T(\tau)\right] = I\delta(t-\tau)$$

we can take the W-matrix as the covariance

$$W = \lim_{t\to\infty} W(t) = \lim_{t\to\infty} E\left[x(t)x^T(t)\right]$$

Without loss of generality, it is assumed that $\tau \to \infty$, and thus the results for deterministic and stochastic approaches technically coincide. If there is an impulse signal at the i'th input with all other inputs being held zero, the response of the system is given by

$$x^i(t) = e^{At} b^i \quad , \quad x^i(t) \quad , \quad b^i \in M_{n,1}$$

where b^i is the i'th column of the matrix B.

Now we can compare the energy in the controllability subspace due to the i'th input only using the Mahalanobis distance. This is a measure of the degree of controllability of the state-space and we define this index by

$$d_i^2 = 1/n \text{ trace } W^{-1} W^i$$

where $$W^i = \lim_{t \to \infty} E\left[x^i(t)(x^i(t))^T\right] = \int_0^\infty e^{At} b^i (b^i)^T e^{A^T t} dt$$

4. Properties of the degree of controllability index

(a) The scalar d_i^2 is invariant under similarity transformations of the linear system (A,B). Under the similarity transformation T, the system becomes (TAT^{-1}, TB) and the distance measure is given by

$$(TWT^T)^{-1}(TW^iT^T) = (T^T)^{-1} W^{-1} W^i T^T$$

The result follows due to the invariancy of the trace under similarity transformations.

(b) The sum of the indexes are equal to unity, $\sum_{i=1}^{m} d_i^2 = 1$.

The matrix W^i is given by the solution of the Lyapunov equation,

$$W^i A^T + A W^i = -b^i (b^i)^T$$

The result is due to the fact that, $W = \sum_{i=1}^{m} W_i$.

(c) The indexes are always positive, and less than unity (equality holds for the case m = 1),

$$1 \geq d_i^2 > 0 \quad , \qquad i = 1,m$$

Since the system is asymptotically stable, the matrix W^i is positive semi-definite[5,11], and hence the result.

5. Extension

The definitions given for controllability for linear continuous systems can be extended to unstable systems, to systems having uncontrollable subspaces, and can be used to study the importance of various outputs using the concept of observability. They can also be extended to discrete systems and to include input signal statistics of the form $E\left[u(t)u^T(\tau)\right] = N\delta(t-\tau)$ where N is a positive diagonal matrix.

6. Conclusions

We have defined measures to quantify the degree of controllability of individual inputs based on the Mahalanobis distance. These measures, in some ways, are similar to the information theoretic measures proposed by Kalman et al[1], and can also be used to investigate the dissimilarity between controllable subspaces[14].

We believe that the definitions based on the Mahalanobis distance are very appropriate for linear systems analysis due to the direct connection with Gaussian processes. However, the Mahalanobis distance is not the only measure available from the fields of information theory and pattern recognition, and other measures[8-10,12,13] both metric and nonmetric (e.g. angles between subspaces[7]) are important in linear systems analysis and design.

7. References

1. Kalman, R.E., Ho, Y.C., and Narendra, K.S.: 'Controllability of linear dynamical systems', Contributions to Differential Equations, 1963, 1, pp.189-213.
2. Johnson, C.D.: 'Optimization of a certain quality of complete controllability and observability for linear dynamical systems', ASME J. Basic Engineering, June 1969, pp.228-238.
3. Mitra, D.: 'W matrix and the geometry of model equivalence and reduction', Proc.IEE, 1969, 116, pp.1101-1106.
4. Friedland, B.: 'Controllability index based on conditioning number', ASME J. of Dynamical Systems, Measurement & Control, December 1975, pp.444-445.
5. Lancaster, P.: 'Theory of Matrices', Academic Press, New York, 1969.
6. McCallion, H.: 'Vibration of linear mechanical systems', Longman, London, 1973.
7. Denham, M.J., and Mahil, S.S.: 'Determination of large scale system structures using principal component analysis', Proc. Int. Conf. on Control and Its Applications, University of Warwick, UK, March 1981.

8. Tou, J.T., and Gonzalez, R.C.: 'Pattern Recognition Techniques', Addison-Wesley, Reading, MA, 1974.

9. Rao, C.R.: 'Advanced Statistical Methods in Biometric Research', Wiley, New York, 1952.

10. Baram, Y., and Sandell, N.R.: 'An information theoretic approach to dynamical systems modelling and identification', IEEE Trans. Auto. Control, 1978, 23, 1, pp.61-66.

11. Barnett, S., and Storey, C.: 'Matrix Methods in Stability Theory', Nelson, London, 1970.

12. Fukunga, K.: 'Introduction to Statistical Pattern Recognition', Academic Press, New York, 1972.

13. Baram, Y.: 'Distance measures for stochastic models', Int. J. Control, 1981, 33, pp.149-157.

14. Fernando, K.V., and Nicholson, H.: 'On discrimination of inputs in multi-input systems', Under preparation.

CHAPTER 13

The Coherence Between System Inputs and Outputs

Abstract: Some measures, based on Gramian matrices and signal processing practice, are defined to quantify the inter-relationship between a particular system input and an output. Such quantifiers are useful in the analysis and design of linear multi-input multi-output multivariable systems. This approach is particularly significant for preliminary studies involving the input-output behaviour of large-scale or complex systems.

1. Introduction

The analysis of interaction between system inputs and outputs and the implications of such interacting behaviour constitutes an important branch in systems theory, especially in the context of large-scale systems. Although the ultimate aim of such analysis is usually to design a 'controller' for the system, the study of interaction itself poses a non-trivial problem to the system analyst. As an example, for a system with ten inputs and ten outputs (which is not an unusual situation in process control or econometrics), there could be up to a hundred possible forms of interaction. In such situations, physical considerations and intuitive reasoning may break down due to the dimensionality of the problem. Although most multivariable techniques utilise interaction effects in the design of controllers, such refined methods are usually not feasible in large-scale problems.

In process control, interactions between inputs and outputs are usually quantified using the relative gain array which is an extension

of the Bristol measure[15,16,22] based on the conditioning number related to the gain of the system. Since the steady behaviour is paramount in process control, this measure is usually based on the dc gain of the system. The Bristol measure has been used widely in the design of controllers for distillation columns and similar processes.

The study of interaction between inputs and outputs is also an important branch of econometrics. Input-output tables are widely used in the Leontief[11] analysis of national economies and other socio-economic behaviour. Using these tables, Leontief was able to disprove some popular but erroneous beliefs in economics.

From the control-theoretic point of view, the input and output behaviour is usually studied in the context of controllability and observability, respectively. Recently, there has been a resurgence of interest in quantifying controllability and observability. Friedland[4] defined an index to quantify the 'goodness' of controllability (and observability) using the conditioning number of the controllability Gramian matrix, which led Moore[1] to define balanced and other principal realizations. Moore has shown the advantages of working directly on signals and their statistics rather than on secondary objects such as model parameters. The principal component analysis of Moore led Denham et al[2] to define some measures to quantify the relationship between a particular input and an output. These measures can be determined by measuring the angles between some controllable and observable subspaces, and also by computing the second-order modes of the system.

In these different but related disciplines, namely econometrics, and control theory, the analysis of interaction between inputs and outputs has been considered as an important part of system analysis. Although these approaches differ conceptually, the underlying theme has been the problem of scaling the system to exemplify the interactions. As pointed out by Brockett et al[21], a unified scaling theory does not exist for the analysis of control systems. In other scientific disciplines, scaling is achieved using non-dimensional quantities, the well known Reynolds number in fluid mechanics being a typical example.

In estimation and other stochastic problems, cross-correlations and cross-covariances between signals, through which coherence functions can be defined, signify the presence or absence of relationships between these signals. Such coherence measures have been exploited fully in the literature, especially in input-output identification. Paradoxically, the applicability and the suitability of such measures for the analysis and design of multivariable systems have not been fully recognised.

The aim of this Chapter is to define similar measures for relating inputs to outputs based on concepts from signal processing practice and least-squares theory. These complement the measures already defined to quantify the degree of controllability/observability[3] (Chapter 12).

In common with the references 1-6, these measures are defined using Gramian matrices (W-matrices) which can be considered as (auto) covariances or second-order moments in a stochastic formulation of the problem, thus giving a direct link with least-squares theory.

However, the controllability Gramian W_c and the observability Gramian W_o are inadequate to compute coherence functions, since they do not give any information about the cross-effects between inputs and outputs, explicitly.

To account for such input-output interactions, a matrix W_{co} has been defined by Fernando and Nicholson[5,18,19] (Chapters 9-11), which is again a second-order statistical average and contains information about both the controllability and observability properties. The coherence measures are specified using scalar functions of these matrices.

2. The controllability and observability Gramians and extensions

For the nth order linear time-invariant, stable, controllable and observable system S(A,B,C),

$$\dot{x}(t) = Ax(t) + Bu(t) \quad , \quad y(t) = Cx(t)$$

the controllability Gramian matrix W_c^i with respect to the input u_i in the infinite interval is defined by

$$W_c^i \triangleq \int_0^\infty (e^{At}b^i)(e^{At}b^i)^T \, dt$$

The vector b^i denotes the ith column of the matrix B and $u_k(t)$ is the kth input.

In a statistical formulation, the matrix W_c^i can be considered as the (auto) covariance of the states $x^i(t)$, under the white noise input $u_i(t)$ with statistical averages of the form,

$$E|u_i(t)| = 0 \quad , \quad E|u_i(t)u_i(\tau)| = \delta(t-\tau)$$

where $E[\cdot]$ denotes the expectation operator, and $\delta(\cdot)$ is the Dirac delta. That is, $E[(x^i(t))(x^i(t))^T] = W_c{}^i$

Similarly, the observability Gramian matrix $W_o{}^j$ is defined by

$$W_o{}^j \triangleq \int_o^\infty (c_j e^{At})^T (c_j e^{At})\, dt$$

where c_j denotes the jth row of the matrix C. The dual system is characterized by,

$$\dot{x}_d(t) = A^T x_d(t) + C^T v(t) \quad , \quad y_d(t) = B^T x_d(t)$$

$$E[v_j(t)] = 0 \quad , \quad E[v_j(t) v_j(\tau)] = \delta(t-\tau)$$

$$v_k(t) = 0 \qquad \text{for } k \neq j \text{ and all } t$$

$$E[(x_d{}^j(t))(x_d{}^j(t))^T] = W_o{}^j$$

In studies involving inputs, the controllability Gramians $W_c{}^i$, $i = 1,n$ are paramount, and similarly, the observability Gramians $W_o{}^j$, $j = 1,n$ with respect to the outputs. However, these Gramians do not convey any direct information about the relationships between the inputs and the outputs of the system. It is well known that in similar problems, the cross effects between the two processes $x^i(t)$ and $x_d{}^j(t)$ are analysed using cross-covariances of the form[5,18,19] (Chapters 9-11),

$$E[(x^i(t)(x_d{}^j(t))^T] = W_{co}{}^{ij}$$

. With the same white noise input in the controllable and the observable systems,

$$u_i(t) = v_j(t) \quad ,$$

the matrix $W_{co}{}^{ij}$ is given by

$$W_{co}{}^{ij} = \int_0^\infty e^{At} b^i c^j e^{At} \, dt$$

The matrix $W_{co}{}^{ij}$ can be computed by solving the following Lyapunov type equation[5,18,19] (Chapters 9-11)

$$W_{co}{}^{ij} A + A W_{co}{}^{ij} = - (b^i)(c^j)$$

The fundamental relationship between the cross-Gramian matrix $W_{co}{}^{ij}$ and the controllability and observability Gramians $W_c{}^i$ and $W_o{}^j$ is given by[5,18,19]

$$(W_{co}{}^{ij})^2 = W_c{}^i W_o{}^j$$

3. Input-output relationships

For the system S(A,B,C), driven by white noise inputs of the form,

$$E[u(t)] = 0 \quad , \quad E[u(t)u^T(\tau)] = \delta(t-\tau)$$

the cross-covariance between the jth output and the ith input is given by

$$E[y_j(t)u_i(t)] = \int_0^\infty c^j e^{At} b^i dt = - c^j A^{-1} b^i$$

which is the negative of the first moment of the system or the dc gain. The moments are invariant under similarity transformations and do not depend on the particular state-space realization.

Information about the first moment is also contained in the cross-covariance matrix $W_{co}{}^{ij}$, and follows from the relation,

$$E[y_j(t)u_i(t)] = 2 \int_0^\infty c^j e^{At} e^{At} b^i dt = 2 \text{ trace } \int_0^\infty e^{At} b^i c^j e^{At} dt$$

$$= 2 \text{ trace } W_{co}{}^{ij}$$

Thus, the matrix W_{co}^{ij} carries information about the ith input and the jth output, which is not present in the Gramian matrices W_c^i and W_o^j, individually. The matrix W_{co}^{ij} may thus be considered as the carrier of information from the ith input to the jth output.

The dc gain is one of the fundamental measures in systems studies and its computation is almost routine in any design method. If the dc gain is zero in a system, then step-wise changes in the output cannot be obtained using step-wise inputs. If the dc gain is non-zero and if the transient response of the system is of no consequence, the long-range (infinite-time) objective of step-wise output control can be achieved by knowing the dc gain alone. This is also true for the multivariable case where all possible dc gains from each input to all the outputs will be required.

In fact, this is the type of control envisaged by Leontief[10,11] in the study of input-output economic and other models. The essentials of our argument are present in Leontief models, although the terminology is obviously different.

In practice, however, control of economic systems are based on heuristic and fuzzy rules which have been determined through past input-output correlations. Instead of using all the possible inputs to control all outputs (which might be required for the 'optimal' strategy) simplified laws are used to control a particular output (or a group of outputs) using a particular input (or a group of inputs). Such input-output pairings are well known and perhaps the control of inflation through control of money supply is a good typical example.

4. Possible measures for quantification of input-output relationships

We have shown the importance of the dc gain in control, and without any reservation it can be considered as the most important measure in relating inputs to outputs. For multivariable systems, we may predict the behaviour of the system by inspecting tabulated values of the dc gain.

As mentioned earlier, the dc gains are invariant under similarity transformations and in that sense they are absolute measures. However, the dc gains are not invariant under input or output scaling. Thus, if a particular input is measured in metres per second instead of kilometres per second, it will be reflected in the dc gain table as a thousand fold increase. If the physical knowledge of the system is limited, this may convey the false impression that the particular input is important because of the high gain path. Thus, for unbiased understanding of systems, we require measures which are also invariant under input and output scaling. We call such measures, structural measures.

One of the measures advocated by Denham et al[2] depends on the sum of the singular values of the system $S(A,b^i,c^j)$. According to our criterion, this quantity is not a structural measure, since it is not invariant under input and output scaling. However, Denham et al avoided some of the difficulties by using lateral arguments involving ratios. Such analysis is essentially Leontief[10,11] in nature!

Another possible quantifier could be based on the concept of output controllability[7] by using the value $c^j W_c{}^i (c^j)^T$ which is also equal to $(b^i)^T W_o{}^j b^i$. Again, this is not a structural measure.

Denham et al also defined quantifiers based on angles between observable subspaces and controllable subspaces. Such a measure is intuitively very appealing and is a true structural property. However, a firm control theoretic interpretation has not yet been found.

5. The coherence between inputs and outputs

It is well known that a coherence measure γ_{pq} to relate two scalar stochastic processes p and q can be defined by

$$\gamma_{pq} = f_{pq}^2 / f_{pp} f_{qq} , \qquad 0 \leq \gamma_{pq} \leq 1$$

where the scalar values f_{pp}, f_{qq} and f_{pq} are some second-order statistical averages of the processes. If the processes p and q are directly related to each other then the coherence measure tends to unity. If these are unrelated then the measure approaches zero.

Similarly, the coherence between the input i and the output j can be measured by considering the states $x^i(t)$ of the controllable system and the dual states $x_d{}^j(t)$ of the observable system. By using scalar functions of the auto-covariance matrices $W_c{}^i$ and $W_o{}^j$ and the cross-covariance matrix $W_{co}{}^{ij}$, we may define scalar coherence measures for vector processes. One possible definition for such a measure is given by

$$\gamma_{ij} = (\text{trace } W_{co}{}^{ij})^2 / \text{trace } (W_c{}^i W_o{}^j) = (\text{trace } W_{co}{}^{ij})^2 / \text{trace} (W_{co}{}^{ij})^2$$

where both the denominator and the numerator, and hence the measure, are invariant under similarity transformations of the form,

$$S(A, b^i, c^j) \rightarrow S(T^{-1}AT,\ T^{-1}b^i,\ c^jT)$$

It is also seen that the measure is invariant under input and output scaling of the form,

$$b^i \rightarrow \alpha\, b^i \quad , \qquad c^j \rightarrow \beta\, c^j$$

$$u^i \rightarrow \frac{1}{\alpha} u^i \quad , \qquad y^j \rightarrow \frac{1}{\beta} y^j$$

where α and β are non-zero scalar values. Thus, this measure is completely independent of scaling and can be considered as a 'non-dimensional' quantity as used in other scientific disciplines. Also, this quantifier is a true structural measure according to our criterion.

We also observe that the term in the numerator is directly related to the dc gain of the system, showing again the importance of this quantity in input-output studies.

An explanation of the term in the denominator of the measure γ_{ij}, namely $\text{trace}(W_c^{\,i} W_o^{\,j})$, is required. In the analysis of roundoff noise in digital filters, Mullis and Roberts[6] demonstrated that (for discrete-time systems) $\text{trace}(W_c^{\,i} W_o^{\,j})$ can be considered as the 'storage energy capacity' of the system. If we drive the system $S(A,b^i,c^j)$ with unity variance white noise in the input $u^i(t)$ from $t = -\infty$, then

$$\text{as } t \to 0 \quad , \qquad E\left[x^i(t)(x^i(t))^T\right] = W_c^{\,i}$$

From, $t = 0^+$ onwards, if the system is unexcited with $u^i(t) = 0$, then

$$E\left[\int_0^\infty (y^j(t))^2\, dt\right] = \text{trace}(W_c^{\,i} W_o^{\,j})$$

As the term 'storage energy capacity' implies, the response $y^j(t)$ is dictated by the amount of storage energy at $t = 0$. Thus, the quantity $\text{trace}(W_c^{\,i} W_o^{\,j})$ will be determined essentially by dynamic elements (such as capacitors and inductors in electrical networks) rather than static elements (such as resistors). It can be shown (using integration by parts) that the storage energy is also given by

$$\text{trace}(W_c^{\,i} W_o^{\,j}) = \int_0^\infty t(h^{ij}(t))^2\, dt$$

where $h^{ij}(t)$ denotes the impulse response of the system.

Now the coherence measure can be described as the ratio between the 'static effects' of the system to the 'dynamic effects' of the system, since the numerator is related to the dc gain and the denominator to the storage energy capacity.

It should now be obvious that the coherence measure defined through signal processing practice has an uncanny analogue in electrical engineering, namely to the power factor or the inverse of the Q-factor.

In robust control of systems (especially in process control[20] and econometrics[11]), it is usually desired that static (low-frequency) behaviour is dominant and the dynamic (transient or high-frequency) effects are minimal (although, at the level of 'fine tuning', dynamic effects such as slight 'overshoots' are desirable in servomechanisms after compensation). The use of the coherence measure in this respect is self-evident since it is based on the ratio between static and dynamic energies.

For a particular input-output pair, if the coherence measure is very low, then it is best described as a 'tuned-circuit'. Alternatively, if the measure is high, then due to the relatively high dc gain, a robust design of controller is possible[13].

We wish to point out that the Bristol measure, which is essentially based on the dc gain of the system, is a quantifier of the static behaviour of the system while the measure proposed by Denham et al is essentially dynamic. In contrast, the coherence measure defined above depends on both the static and the dynamic behaviour of the system.

6. The coherence measure for internally balanced representations

The system $S(\bar{A},\bar{b}^i,\bar{c}^j)$ is said to be internally balanced if the controllability Gramian matrix $\bar{W}_c^{\,i}$ and the observability Gramian matrix $\bar{W}_o^{\,j}$ are diagonal and equal[1]. If a system $S(A,b^i,c^j)$ is not internally balanced, this can be achieved by similarity transformations, and

hence the coherence measure will also be invariant. We denote the balanced Gramian matrix as W^{ij}, that is

$$W^{ij} = \bar{W}_c{}^i = \bar{W}_o{}^j$$

Since, $(W_{co}{}^{ij})^2 = W_c{}^i W_o{}^j$, it is seen[5] that the cross-Gramian matrix $\bar{W}_{co}{}^{ij}$ is diagonal for balanced representations, and we denote this diagonal matrix as V^{ij}. That is, $\bar{W}_{co}{}^{ij} = V^{ij}$. It is easy to verify that the diagonal values of the matrix V^{ij} denoted by $v_k{}^{ij}$, $k = 1,n$ are the eigenvalues of the cross-Gramian matrix $W_{co}{}^{ij}$.

In fact, these diagonal values are the singular values (second-order modes) of the system except for possible sign variations. If we denote the singular values of the system by $w_k{}^{ij}$, $k = 1,n$, which are the diagonal values of the matrix W^{ij}, then

$$w_k{}^{ij} = |v_k{}^{ij}| \qquad , \qquad k = 1,n$$

The following relationships are then obvious,

$$\text{trace } W_{co}{}^{ij} = \sum_{k=1}^{n} v_k{}^{ij}$$

$$\text{trace } W_c{}^i W_o{}^j = \text{trace } \left(W_{co}{}^{ij}\right)^2 = \sum_{k=1}^{n} (v_k{}^{ij})^2$$

The coherence measure is then given by

$$\gamma_{ij} = \frac{(\sum_{k=1}^{n} v_k{}^{ij})^2}{\sum_{k=1}^{n} (v_k{}^{ij})^2} \qquad (1)$$

7. A modified coherence measure

It is well known that it is difficult to control systems which have both positively and negatively decaying exponentials in their impulse responses. This is a major problem in process control systems[15-17,22] where simple controllers are preferred. The intrinsic reason for non-minimum phase behaviour of systems is also due to these mixed exponentials.

Since the signature of the exponentials are determined by the residues at the poles of the system, these signatures are given by the Cauchy index[18] (Chapter 11) (assuming that the poles are distinct). For an nth order system with all positively decaying exponentials, the Cauchy index is given by n. Similarly, if they are negatively decaying, then it is equal to -n. Such systems are always minimal phase. However, if the Cauchy index lies between these extremes, then non-minimal phase behaviour can occur, depending on the parameters of the system, and with increasing possibility if the magnitude of the index is low. Thus, the Cauchy index is an indicator which can be used in identifying 'troublesome' input-output pairs.

The main difficulty with the Cauchy index is that, like the rank of a matrix[4], it is essentially a non-robust measure which can vary under a small perturbation of the system parameters. Thus, to properly quantify the information in the Cauchy index, we have to qualify it by using a condition number. In this respect, the cross-Gramian matrix W_{co}^{ij} is valuable.

Fernando and Nicholson[18] (Chapter 11) have shown that the signature of the matrix W_{co}^{ij} is equal to the Cauchy index of the system. Since the robust part of the system[1] is dictated by the dominant second-order

modes (as given by the unsigned eigenvalues of the matrix $W_{co}{}^{ij}$), the signature of these values is more important than the signature of the non-dominant values. One way of accounting for the magnitudes of the second-order modes in the definition of the 'condition number' is to weight them proportionally. Such a measure can be defined using the eigenvalues of the cross-Gramian matrix $W_{co}{}^{ij}$ as

$$\beta_{ij} = \frac{(\sum_{k=1}^{n} v_k{}^{ij})^2}{(\sum_{k=1}^{n} |v_k{}^{ij}|)^2} \qquad (2)$$

The numerator of this measure, as in the case of γ_{ij}, is based on the dc gain of the system. The denominator depends on the dc gain of an hypothetical system with a cross-Gramian matrix which has positive eigenvalues given by $|v_k{}^{ij}|$. That is,

$$\beta_{ij} = \frac{(\text{dc gain of the original system } S(A,b^i,c^j))^2}{(\text{dc gain of the hypothetical system})^2}$$

Since β_{ij} is similar to the coherence measure γ_{ij}, we call this new measure β_{ij} the modified coherence measure. It is easily verified that this measure always takes values between zero and one,

$$0 \leq \beta_{ij} \leq 1$$

Furthermore, for first-order systems, it is always equal to unity and for non-proper systems it is equal to zero.

If the magnitude of the Cauchy index of the dominant part of the system is high, then the measure will take values near unity. If it is low, then the measure β_{ij} will take low values.

8. The numerical procedure

To compute the coherence measures between the input i and the output j,

(a) Solve the equation $W_{co}^{ij}A + AW_{co}^{ij} = -b^i c^j$ using the algorithm given in reference 12 or an equivalent algorithm.

(b) Compute the eigenvalues of the matrix W_{co}^{ij} which are denoted by v_k^{ij}, k = 1,n.

(c) Compute the coherence measure γ_{ij} using equation (1) and/or the modified measure β_{ij} using equation (2).

Since eigenvector calculations are avoided, the computation of the measures can be accomplished using orthogonal transformations. Thus, the procedure is numerically well-conditioned.

As described in Section 6, the 'obvious' computational scheme would be to compute the balanced realizations for each input and output. However, this obvious approach is undesirable due to the following reasons.

(1) Computation of balanced realizations is numerically expensive and such realizations are unnecessary as far as the computation of the second-order modes and the measures are concerned.

(2) All published numerical procedures for computation of balanced realizations are based on the eigen-structure of the matrix product $W_c W_o$, which is equal to the square of the matrix W_{co}. Since the matrix product is not formed in our approach, it is a 'square-root' method and thus well-conditioned with respect to roundoff errors.

However, if balanced realizations are required, then they can be obtained[5] (Chapter 9) by computing the eigenvectors of the matrix W_{co}.

9. The experimental procedure

The coherence measure γ_{ij} can be computed using experimental results without knowing the model of the process. The procedure is,

(a) Determine the dc gain from input i to output j using step responses.

(b) Determine the impulse response from input i to output j denoted by $h^{ij}(t)$ and compute the integral,

$$\int_0^\infty t(h^{ij}(t))^2 dt = \text{trace } W_{co}^2 = \text{trace } W_c W_o$$

(c) Compute the coherence measure using the formula,

$$\gamma_{ij} = \frac{(0.5 \times \text{dc gain of the system } S(A,b^i,c^j))^2}{\text{trace } W_c W_o}$$

10. Illustrative examples

Example 1: The single-input single-output system as defined by the transfer function,

$$H(s) = \frac{s}{(s+2)(s+1)}$$

was considered by Wang and Davison[13] to illustrate the difficulties in designing a robust controller. Because of the zero at the origin, the first moment of the system is zero and thus the system cannot track a step input.

The coherence measures are obviously equal to zero,

$$\gamma_{ij} = \beta_{ij} = 0$$

due to the zero at the origin. Thus, our measures also indicate that the system is somewhat ill-conditioned. However, further research is required to compare ill-conditioning in robust controller problems with that of balanced systems.

The diagonal V matrix for the system is given by,

$$V^{11} = \text{diag } (1/36, -1/36)$$

and the singular values (second-order modes) are of the form,

$$W^{11} = \text{diag } (1/36, 1/36)$$

One of the reasons for the ill-conditioned nature of the system is due to the non-distinct singular values[5,14], which is reflected directly in the coherence measures.

This example also indicates that if the first and the higher derivatives of the input are dominant in the impulse response, then the coherence measures are low, tending towards zero. Alternatively, if the derivatives are not dominant, then we may expect high coherence values.

Example 2: To illustrate some of the properties of the measures, we consider the transfer function,

$$\frac{s + z}{(s + 1)(s + 3)}$$

where z is a real number. If z lies between 1 and 3, then the residues at the poles are positive (Cauchy index of 2) and thus the exponentials will be positively decaying. Such systems are relatively easy to control. If z lies outside this domain, then the Cauchy index is zero and the exponentials will have opposite signatures. Due to the form of the measures we would expect large values for them if z lies between 1 and 3.

We now consider a transfer-function matrix of the form,

$$\begin{bmatrix} y_1(s) \\ y_2(s) \end{bmatrix} = \frac{1}{(s + 1)(s + 3)} \begin{bmatrix} k_{11}s & k_{12}(s+2) \\ k_{21}(s+6) & k_{22}(s-10) \end{bmatrix} \begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix}$$

where k_{ij}, i,j = 1,2 are real scalar constants.

Denham et al used the storage energy capacity of the system (sum of the singular values) as the criterion for choosing input-output pairs for control. If k_{11} is large enough, then the combination (u_1,y_1) can be obtained as the 'optimal' result which is obviously a bad choice since the subsystem 1,1 is non-proper.

If we use the dc gain as the deciding criterion (assuming all k_{ij} are equal) then the combination (u_2,y_2) is the first choice. This is inappropriate due to the badly positioned numerator zero.

Table 1 gives the coherence measures (which are invariant of k_{ij}) for this problem. It is seen that the best combinations for control are given by

$$u_2 \rightarrow y_1 \quad , \quad u_1 \rightarrow y_2$$

Table 1: The measure for example 2

	y_1	y_2
u_1	0.	1.10345
u_2	0.86487	0.48251

1(a): The coherence measure γ_{ij}

	y_1	y_2
u_1	0.	1.0
u_2	0.76191	0.48251

1(b): The modified coherence measure β_{ij}

Example 3: To illustrate the measures we have defined for physical systems, we have considered the state-space model of an oil-fired boiler[8]. This is a ninth order model which has an asymptotically stable 8th order subsystem which can be obtained by directly decoupling the ninth state. There are three main inputs and three significant states in this subsystem. They are,

u_1 mass flow of the steam at the superheater

u_2 mass flow of fuel

u_3 mass flow of water at the economiser

x_1 steam density

x_2 superheater steam temperature

x_6 steam drum pressure

Tables 2(a) and 2(b) show half the dc gain and the sum of the singular values, respectively, which are required in the calculation of the coherence measures. It is obvious from these tables that, even for a three-input three-output system, the relationships between inputs and outputs are difficult to assess.

However, from tables 2(c) and 2(d) which give the coherence measures, the inter-relationships between inputs and outputs are quite explicit. These tables indicate that the three inputs are directly related to all the three outputs except that u_3 is almost completely unrelated to x_2. This tabular evidence is in agreement with physical reasoning.

If one-to-one input-output pairing is required, we may obtain the following combinations by observing the highest values in table 2(c) or 2(d) and avoiding the unrelated pair (u_3,x_2) in the combination. One possible combination is given by,

$$u_1 \rightarrow x_1 \quad , \quad u_2 \rightarrow x_2 \quad , \quad u_3 \rightarrow x_6$$

which is a reasonable choice. However, this is not the only possible solution.

Table 2: The measures for example 3

	x_1	x_2	x_6
u_1	-1.021×10^{-2}	-4.807	-1.248×10^{3}
u_2	2.098×10^{-1}	8.470×10^{1}	2.700×10^{4}
u_3	-4.764×10^{-3}	-6.335×10^{-6}	-4.645×10^{2}

2(a): Half the dc gain

	x_1	x_2	x_6
u_1	8.483×10^{-3}	4.698	1.179×10^{3}
u_2	2.088×10^{-1}	8.491×10^{1}	2.695×10^{4}
u_3	5.093×10^{-3}	2.794×10^{-2}	4.956×10^{2}

2(b): Sum of the singular values

	x_1	x_2	x_6
u_1	1.450	1.047	1.120
u_2	1.009	0.995	1.010
u_3	0.875	5.1×10^{-8}	0.878

2(c): The coherence measure γ_{ij}

	x_1	x_2	x_6
u_1	0.958	0.976	0.964
u_2	0.891	0.988	0.924
u_3	0.775	2.4×10^{-8}	0.781

2(d): The modified coherence measure β_{ij}

Example 4: As the final example, we have chosen to study the 16th order model[9] of an F100 turbofan jet engine which was also considered by Denham et al[2]. This model also has three main inputs and three main outputs, where

u_1 main burner fuel flow

u_2 nozzle jet area

u_3 inlet guide vane position

y_1 engine net thrust level

y_2 total engine airflow

y_3 turbine inlet temperature

Tables 3a and 3b give half the dc gain and the sum of the singular values, respectively, and we observe that the values in table 3b are marginally different from those of Denham et al[2]. As mentioned earlier, Denham et al used some lateral arguments to choose input-output pairings for this problem. However, if we consider table 3c or 3d, the conclusions are quite explicit and are given by

$$u_1 \rightarrow y_1 \quad , \quad u_2 \rightarrow y_3 \quad , \quad u_3 \rightarrow y_2$$

and these combinations agree with those of Denham et al[2]. However, our measure γ_{ij} is fundamentally different from that of Denham et al since the sum of the singular values used by them appears in the denominators of eqns 1 and 2.

Table 3: The measures for example 4

	y_1	y_2	y_3
u_1	5.000×10^{-1}	2.610×10^{-3}	6.063×10^{-2}
u_2	-4.819×10^{2}	8.627	1.412×10^{2}
u_3	-7.075	-1.973×10^{-1}	-1.072

3(a): Half the dc gain

	y_1	y_2	y_3
u_1	4.339×10^{-1}	2.528×10^{-3}	5.864×10^{-2}
u_2	1.069×10^{3}	9.594	1.508×10^{2}
u_3	1.048×10^{1}	1.938×10^{-1}	1.500

3(b): Sum of the singular values

	y_1	y_2	y_3
u_1	1.328	1.066	1.069
u_2	0.203	0.809	0.877
u_3	0.456	1.036	0.511

3(c): The coherence measure γ_{ij}

	y_1	y_2	y_3
u_1	0.9086	0.7932	0.7732
u_2	0.0962	0.5860	0.5891
u_3	0.2296	0.8953	0.2758

3(d): The modified coherence measure β_{ij}

10. Conclusion

We have defined coherence measures using fundamental data related to controllability and observability which are consistent with signal processing practice and least-squares theory. The first coherence measure may be interpreted as the ratio between static and dynamic energy. The modified measure is directly related to the Cauchy index of the system. We have illustrated the usefulness and the importance of these measures using non-trivial examples.

The coherence measures can be evaluated for very large-scale systems because of the availability of efficient numerical algorithms required for computation. Alternatively, they may be determined experimentally.

We have also highlighted the relationships between our approach and that of input-output Leontief models. In such economic and other social systems, dc gain is paramount and the transient behaviour is secondary. Since, the second-order modes (singular values) are available in the intermediate calculations, the controllability and observability properties of the system are also available which is a definite advantage in these studies.

We have also established the importance of the coherence measures in determining possible non-minimum phase behaviour and the occurrence of exponentials of mixed signatures in impulse responses. Such information is especially valuable in process control.

We do not claim that the measures defined in this paper are the best or that there are no other alternatives in large-scale system analysis. However, our approach seems to be reasonable and applicable. Furthermore, they are consistent with the requirements in a variety of disciplines.

We have not discussed the possible application of these measures in the control of large or complex systems. However, the coherence measures together with other measures which characterise dynamic and static behaviour can have applications in 'structure free' modelling and control of systems as envisaged by Bristol[20] and others. These measures can also have application in the design of fuzzy controllers.

11. References

1. Moore, B.C.: 'Principal component analysis in linear systems: Controllability, observability and model reduction', IEEE Trans. Automatic Control, 1981, AC-26, (1), pp.17-32.
2. Denham, M.J., and Mahil, S.S.: 'Determination of large scale system structures using principal component analysis', Proc. Int. Conf. on Control and Applications, Univ. of Warwick, UK, March 1981.
3. Fernando, K.V., and Nicholson, H.: 'Degree of controllability due to individual inputs', Electronics Letters, 1981, 17, (9), pp.330-331.
4. Friedland, B.: 'Controllability index based on conditioning number', ASME J. Dyn. Syst. Meas. & Control, 1975, pp.444-445.
5. Fernando, K.V., and Nicholson, H.: 'On the structure of balanced and other principal representations of SISO systems', IEEE Trans. Automatic Control, 1983, AC-28, (1),
6. Mullis, C.T., and Roberts, R.A.: 'Roundoff noise in digital filters: Frequency transformations and invariants', IEEE Trans. Acoustics, speech, and signal processing, 1976, ASSP-24, (6), pp.538-550.
7. Kreindler, E., and Sarachik, P.E.: 'On the concept of controllability and observability of linear systems', IEEE Trans. Automatic Control, 1964, AC-9, (2), pp.129-136.
8. Nicholson, H.: 'Dynamic optimisation of a boiler', Proc.IEE, 1964, 111, pp.1479-1499.
9. Sain, M.K., Peczkowski, J.L., and Melsa, J.L.: 'Alternatives for linear multivariable control', National Engineering Consortium, Chicago, 1978.
10. Bacharach, M.: 'Biproportional matrices and input-output change', The University Press, Cambridge, 1970.

11. Leontief, W.: 'Input-output economics', Oxford Univ. Press, New York, 1966.

12. Bartels, R.H., and Stewart, G.W.: 'Solution of the matrix equation AX + XB = C', Communications ACM, 1972, 15, (9), pp.820-826.

13. Wang, S.H., and Davison, E.J.: 'Penetrating transmission zeros in the design of robust servomechanism systems', IEEE Trans. Automatic Control, 1981, AC-26, (3), pp.784-787.

14. Pernebo, L., and Silverman, L.: 'Balanced systems and model reduction', Proc. 18th IEEE Conf. Decision Control, Fort Lauderdale, Florida, December 1979.

15. Tung, L.S., and Edgar, T.F.: 'Analysis of control-output interactions in dynamic systems', AIChE Journal, 1981, 27, (4), pp.690-693.

16. Bristol, E.H.: 'Recent results on interactions in multivariable process control', AIChE 71st Annual Meeting, Miami, FL, November 1978.

17. Bristol, E.H.: 'The right half plane'll get you if you don't watch out', Proc. JACC, 1981, Charlottesville, Va., USA.

18. Fernando, K.V., and Nicholson, H.: 'On the Cauchy index of linear systems', IEEE Trans. Automatic Control, 1983, AC-28, (3), (tentative).

19. Fernando, K.V., and Nicholson, H.: 'On the minimality of SISO linear systems', Proc. IEEE, to be published.

20. Bristol, E.H.: 'Pattern recognition: An alternative to parameter identification in adaptive control', Automatica, 1977, 13, pp.197-202.

21. Brockett, R.W., and Krishnaprasad, P.S.: 'A scaling theory for linear systems', IEEE Trans. Automatic Control, 1980, AC-25, (2), pp.197-207.

22. Shinskey, C.J.: 'Process-Control Systems', McGraw-Hill, New York, 1979.

CHAPTER 14

On Discrimination of Inputs in Multi-input Systems

Abstract: A metric information measure known as the Mahalanobis distance is used to quantify the dissimilarity between controllable subspaces due to any two inputs in multi-input linear systems.

1. Introduction

Recently, a metric information measure known as the Mahalanobis distance[2-4] was used to quantify the effectiveness of inputs in multi-input linear systems[1]. However, for proper understanding of a control system, this measure of effectiveness alone is not sufficient and a measure for similarity or dissimilarity of controllable subspaces due to individual inputs is required. The object of this Chapter is to define such a measure for linear systems using the Mahalanobis distance.

Apart from the theoretical importance, such measures can be used to analyse complex systems. With some control problems, the designer will have discretion in choosing the system inputs based on physical reasoning and engineering judgement. However, with complex problems such qualitative reasoning might be difficult or absent and more quantitative measures are required.

Quantitative measures for discrimination of inputs are important for system operation, and under emergency operating conditions an alternative control strategy with loss of an input (eg. actuator

failure) can be devised if the similarity or dissimilarity of controllable subspaces due to individual inputs is known. Thus, quantifiers for the discrimination of inputs together with an effectiveness measure[1] (Chapter 12) are crucial in the design and operation of complex systems.

2. The Mahalanobis distance for discrimination of vectors[2-4]

For two classes of random vectors of dimension n, denoted by z^i and z^j which belong to the classes S^i and S^j, respectively, and which are subsets of the general class S,

$$z^i \in S^i \quad , \quad z \in S \quad , \quad S^i \subseteq S \quad , \quad i = 1,m$$

the Mahalanobis distance between the vectors is defined by

$$M^2(z^i,z^j) = \|z^i - z^j\|_{\phi^{-1}} = (z^i - z^j)^T \phi^{-1} (z^i - z^j)$$

The square n,n matrix ϕ is the covariance of vectors z,

$$\phi = E\left[(z-\bar{z})(z-\bar{z})^T\right] \quad , \quad E[z] = \bar{z}$$

and $E[\cdot]$ denotes the expectation operator.

The expected value of the measure is given by,

$$E\left[M^2(z^i,z^j)\right] = \text{trace } \phi^{-1}\{(\phi^{ii}+\phi^{jj}) - (\phi^{ij}+\phi^{ji})$$

where $\phi^{ij} = E\left[(z^i)(z^j)^T\right]$

It is a measure of dissimilarity between the classes S^i and S^j, and gives high values if they are orthogonal and low values when they are similar.

3. Discrimination of inputs

We assume that the linear system (A,B) is asymptotically stable and fully controllable. The controllability Gramian for the system,

$$\dot{x} = Ax + Bu$$

is given by

$$W(\tau) = \int_0^\tau x(t)x^T(t)dt$$

for deterministic unit impulses at the inputs. For stochastic inputs of the form

$$E\left[u(t)\right] = 0 \quad , \qquad E\left[u(t)u^T(\tau)\right] = I\delta(t-\tau)$$

we can take the W-matrix as the covariance,

$$W = \lim_{t\to\infty} W(t) = \lim_{t\to\infty} E\left[x(t)x^T(t)\right]$$

If there is an impulse at the ith input, with all other inputs to the system being held zero, the response of the system is given by

$$x^i(t) = e^{At}b^i \quad , \qquad x^i(t) \ , \ b^i \in M_{n,1}$$

where the vector b^i is the ith column of the matrix B.

We assume that the columns of the matrix B are normal which can be achieved by scaling the inputs. That is, $(b^i)^T b^i = 1$ for all i. This normalisation is required to avoid discrimination of two inputs which are identical except for their amplitudes.

The degree of dissimilarity between the controllable subspaces due to ith and jth inputs can be measured using the Mahalanobis distance measure for discrimination of vectors. We define this measure as,

$$d^2_{ij} = \text{trace } W^{-1}\{(W^{ii} + W^{jj}) - (W^{ij} + W^{ji})\}$$

where $$W_{ij} = \int_0^\infty e^{At} b^i (b^j)^T e^{A^T t} dt$$

$$W = \int_0^\infty e^{At} BB^T e^{A^T t} dt$$

The matrix W^{ij} can be computed by solving the Lyapunov equation,

$$W^{ij} A^T + AW^{ij} = -b^i (b^j)^T$$

Obviously, high magnitudes of d_{ij}^2 indicate dissimilarity of controllable subspaces due to inputs i and j.

4. Properties of the discrimination index

(a) Under similarity transformations of the form,

$$(A,B) \rightarrow (TAT^{-1},B)$$

the scalar d_{ij}^2 is invariant.

(b) The discrimination index is nonnegative

$$d_{ij}^2 \geq 0$$

(c) For input normal systems[5], the Gramian matrix W is equal to unity, and thus the index can be written in the simplified form,

$$d_{ij}^2 = \text{trace}\{(W^{ii} + W^{jj}) - (W^{ij} + W^{ji})\} \qquad (1)$$

5. A modified measure

We assume, without loss of generality, that the system is input normalized[5]. The distance measure for discrimination is then given by eqn 1. If the ith column of matrix B is equal to the jth column,

then the distance between them is zero.

$$d_{ij}^{\,2} = 0 \quad , \quad \text{for } b^i = b^j$$

However, if b^j differs only with respect to sign, then the distance is given by

$$d_{ij}^{\,2} = 4 \text{ trace } W^{ii} \quad , \quad \text{for } b^i = -b^j$$

In most applications, discrimination with respect to sign variations is not required. Thus, we define a modified measure which is insensitive to sign variations as,

$$\tilde{d}_{ij}^{\,2} = \overline{\text{tr}} \{(W^{ii} + W^{jj}) - (W^{ij} + W^{ji})\}$$

where $\overline{\text{tr}}$ is defined as

$$\overline{\text{tr}}\, W = \sum_{i=1}^{n} |w_{ii}|$$

where w_{ii} are the diagonal elements of the matrix W. Thus, for column vectors b^i and b^j which differ with respect to sign only, the modified distant measure is given by,

$$\tilde{d}_{ij}^{\,2} = 0 \quad , \quad \text{for } b^i = -b^j$$

6. Conclusions

A measure for discrimination of controllable subspaces due to any two inputs in a multi input system has been defined using the Mahalanobis distance. A modified measure was also defined which is insensitive to sign variations. These measures, together with that for controllability, provide meaningful measures for quantifying the effectiveness of individual inputs. By using observability

Gramians instead of controllability Gramians, these measures can also be used to analyse outputs in multiple-output systems.

7. References

1. Fernando, K.V., and Nicholson, H.: 'The degree of controllability due to individual inputs', Electronics Letters, 1981, 17, 9, pp.330-331.

2. Tou, J.T., and Gonzalez, R.C.: 'Pattern recognition techniques', Addison-Wesley, Reading, Ma., 1974.

3. Marill, T., and Green, D.M.: 'On the effectiveness of receptors in recognition systems', IEEE Trans. Information Theory, 1963, IT-9, pp.11-17.

4. Mahalanobis, P.C.: 'On the generalized distance in statistics', Proc. National Inst. Sci. India, 1936, 2, pp.49-55.

5. Moore, B.C.: 'Principal component analysis in linear systems: Controllability, observability, and model reduction', IEEE Trans. Automatic Control, 1981, AC-26, (1), pp.17-32.

PART 6

Closure

CHAPTER 15

Closure

1. Some percepts (and heresies)

Importance of the Gramian and covariance matrices and their spectral decompositions in control and systems theory including pattern recognition has been highlighted in this thesis. In all the problems studied, removal of redundant or 'almost' redundant information was the recurrent theme. Such superfluous data (often due to noise, uncontrollable/unobservable modes, unreliable high frequency effects, modelling errors etc.) unnecessarily complicates the understanding of underlying mechanisms which govern the process. Thus, removal of redundant data is prerequisite in the analysis of complex or large-scale systems. In any case, the predicted information explosion due to advances in microelectronics, fibre optics, satellite communication and intelligent computers can be deflated by efficient methods of information contraction.

In Part 2 of the thesis, where the Karhunen-Loève expansion and its extensions are studied, the models assumed for data reduction and extrapolation were rather subtle or non-existent. However, in Parts 3 and 4, formal state-space representations were used in the analysis. It is well known that a large class of systems can be modelled via state-space representations (or equivalently by transfer functions) and thus model-order reduction methods and related problems are paramount in system theory. However it is well known that there are non-trivial problems in parameter identification even with linear representations (for example, see reference 1 for

different results given by different techniques of identification for a second-order scalar system). An alternative for identification is mathematical modelling. However, except for systems which are governed by well-behaved physical laws (such as electromagnetic phenomena) modelling is no easy task and this leaves out most of the social phenomena. Thus, state-space and other formal representations, although elegant from a mathematical point of view cannot be considered as universal. Due to these difficulties in formal representations, the author believes that more effort should be spent on investigating the feasibility of using simple, subtle and implicit models in representing dynamical systems. This is the motivation for Part 5 where non-orthodox ways of characterization of dynamical systems are studied.

The airline data which is prominent in Part 2 has been analysed in the literature using a wide variety of complicated time-series techniques. However as indicated in Chapter 4, a simple model based on the Karhunen-Loeve expansion/singular value decomposition can give superior results to that of time-series methods. This is due to the non-conformity of data to man-made assumptions. For example it is often assumed that the data is stationary, does not have trends, is asymptotically stable, etc, which are at odds with reality. The author believes that by reducing the number of assumptions and by using more simple models some of these problems can be avoided.

What are the alternatives if formal models are too complex or non existent, which is the case for most of the phenomena observed in real life?

This is a difficult question to answer. However, one approach could be to decompose the data into their orthogonal (perpendicular) components so that each component can be studied as independent scalar systems. In this way, it may be possible to understand the underlying mechanics (if any) of the process without problems associated with dimensionality. This is the main reason for diagonalization of covariance/Gramian matrices.

One may argue that orthogonalization procedure is an abstract mathematical technique without any importance in reality. This is not true and there are many situations where orthogonality has been used in practice. There are a non-countable infinite number of colours in the visible spectrum. However, we need only three primary (independent) colours to represent the wide spectrum.

2. References

1. Söderström, T.: 'Identification of stochastic linear systems in presence of input noise', Automatica, 1981, 17, (5), pp.713-725.

PART 7

Appendices

APPENDIX 1

The double-sided least-squares problem

Abstract. The double-sided least-squares problem is formulated under a separability condition, using the properties of the Kronecker product to obtain the overall solution based on two standard sub-problems.

1. Introduction. The standard matrix least-squares problem[1] is concerned with obtaining a solution for the 'n' column elements of the unknown state or parameter matrix X which are related to the 'n' column elements of the observed matrix Y by the equation

$$Y = HX+E \quad , \quad Y \in M_{m,n} \quad , \quad H \in M_{m,p} \quad , \quad X \in M_{p,n} \tag{1}$$

where H is a known matrix and E is a matrix of residual errors.

If H is of maximal rank p, the least-squares estimate, obtained by minimizing the error matrix

$$J = E^T P E \quad , \qquad P \in M_{m,m}$$

is given by

$$\hat{X} = (H^T P H)^{-1} H^T P Y$$

The solution corresponds to a linear transformation of Y of the form

$$\hat{Y} = H\hat{X} = H(H^T P H)^{-1} H^T P Y \triangleq NPY$$

giving the observed error matrix

$$Y-\hat{Y} = (I-NP)Y \triangleq P^{-1}LY \quad , \quad L = P-PNP$$

and $$J = Y^{T}LY$$

In terms of the matrix elements, the relationship of eqn 1 is given by

$$y_{ij} = \sum_{k=1}^{p} h_{ik}x_{kj} + e_{ij}$$

The column j of the matrix Y correlates the column j of the matrix X and no interactions are assumed to exist between the columns.

A similar least-squares problem can also be formed in terms of observed and state row vectors, with the matrix representation

$$Y = XG^{T} + E \quad , \quad Y \in M_{m,n} \quad , \quad X \in M_{m,q} \quad , \quad G \in M_{n,q} \qquad (2)$$

where G is the known coefficient matrix. The least-squares solution obtained by minimizing $J = EQE^{T}$ is then given by

$$\hat{X} = YQG(G^{T}QG)^{-1}$$

The elements of eqn 2 are given by

$$y_{ik} = \sum_{k=1}^{q} x_{ij}g_{jk}$$

In this case, row i of the matrix Y correlates the same row of the matrix X and no interactions exist between rows.

The problems represented by eqns 1 and 2 are equivalent since the ordering of rows and columns are not generic properties. The correlations or the 'flows' are in different directions but are, however, uni-directional. Such relationships have very wide application in classical least-squares estimation, but are inadequate for the representation of two-dimensional processes

involving two-directional flows. These require a more general transformation between two matrices which will correlate both row and column elements.

Such a relationship exists between the m,n dimensional matrix Y and a p,q dimensional matrix Z with the Kronecker product mapping[2]

$$\text{vec}(Y) = F\,\text{vec}(Z) \quad , \quad F \in M_{mn,pq} \quad , \quad y \in M_{mn,1} \qquad (3)$$

$$\text{or} \qquad y = (G \otimes H)z \quad , \quad G \in M_{n,q} \quad , \quad H \in M_{m,p} \quad , \quad z \in M_{pq,1}$$

where vec(•) is the operator which stacks columns of a matrix into a column vector and F is the Kronecker product of the matrices G and H. Cross-correlations then exist between rows and columns with the expansion

$$y_{ij} = \sum_{k=1}^{p} \sum_{\ell=1}^{q} h_{ik}\; g_{j\ell}\; z_{k\ell}$$

Each element y_{ij} of the matrix Y now depends on all the elements of the matrix Z instead of on one particular row or column. The matrix equivalent of eqn 3 is given by

$$Y = HZG^{T}$$

This includes a composition of two linear transformations, and the equivalent vector map is formed using the tensor or Kronecker product of two linear transformations. This form of representation was known to Sylvester[3].

2. The double-sided least-squares problem. We now consider the double-sided least-squares formulation with the equation

$$Y = HZG^T + E \tag{4}$$

relating the measurement or observed matrix Y and the parameter or state matrix Z, where H and G are known maximal rank matrices and E is the residual error matrix. The vector form of eqn 4 is

$$y = Fz + e$$

where $e = \text{vec}(E)$. In a statistical framework, we may also introduce the properties

$$E[e] = 0 \quad , \quad E[ee^T] = R$$

where $E[\cdot]$ is the expectation operator and R is the error covariance matrix of the random noise vector e.

With a minimizing function

$$J = e^T S e \quad , \quad S \in M_{mn,mn} \tag{5}$$

and with F of maximal rank (pq), the least-squares solution is given by

$$\hat{z} = (F^T S F)^{-1} F^T S y \quad , \quad (F^T S F) \in M_{pq,pq}$$

The problem dimension can now be reduced if the error criterion matrix S is assumed to be separable, of the form

$$S = Q \otimes P$$

The least-squares solution vector is then given by

$$\begin{aligned}
\hat{z} &= [(G \otimes H)^T (Q \otimes P)(G \otimes H)]^{-1} (G \otimes H)^T (Q \otimes P) y \\
&= (G^T Q G)^{-1} \otimes (H^T P H)^{-1} [(G^T Q) \otimes (H^T P)] y \\
&= [(G^T Q G)^{-1} G^T Q] \otimes [(H^T P H)^{-1} H^T P] y
\end{aligned}$$

and the corresponding least-squares solution matrix is

$$\hat{Z} = (H^T PH)^{-1} H^T PYQG(G^T QG)^{-1} \qquad (6)$$

$$= Z + (H^T PH)^{-1} H^T PEQG(G^T QG)^{-1}$$

The measurement estimate then introduces a composition of two linear transformations of Y of the form

$$\hat{Y} = H\hat{Z}G^T = H(H^T PH)^{-1} H^T PYQG(G^T QG)^{-1} G^T = NPYQM$$

where $N = H(H^T PH)^{-1} H^T$, $M = G(G^T QG)^{-1} G^T$

The observed error matrix is given by

$$Y - \hat{Y} = Y - NPYQM$$

The error function of eqn 5 can also be written in the form

$$J = \sum_{i,k=1}^{m} \sum_{j,\ell=1}^{n} q_{j\ell} e_{ij} p_{ik} e_{k\ell}$$

or $J = Q * E^T PE = EQE^T * P$

where the operation * represents the bilinear scalar product of two similar dimension matrices defined by

$$A * B = B * A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij} \quad , \quad A,B \in M_{m,n}$$

$$I * B = \text{trace } B$$

Alternatively,

$$J = y^T [S - S(M \otimes N)S] y$$

$$= y^T [S - (QMQ) \otimes (PNP)] y$$

$$= Q * (Y^T PY) - (QMQ) * (Y^T PNPY)$$

If the error covariance matrix R is assumed to be a separable process, of the form

$$R = U \otimes V \quad , \quad U \in M_{n,n} \quad , \quad V \in M_{m,m}$$

and if the weighting matrices P and Q are set equal to the inverses of the error covariance matrices respectively, then the error covariance matrix for the least squares vector estimate z is given by

$$E\left[(z-\hat{z})(z-\hat{z})^T\right] = (G^T U^{-1} G)^{-1} \otimes (H^T V^{-1} H)^{-1}$$

2.1 The equivalent decomposed problem. The double-sided least-squares problem represented by eqn 3 can be decomposed into two standard least-squares sub-problems. These are equivalent to column and row 'scanning' and the estimate of the state matrix Z or the 'image' can be formed from a combined solution of the sub-problems.

The overall problem is then represented by the column problem

$$Y = HX + E \quad , \quad J = E^T P E$$

with the estimate $\hat{X}$ used as an 'observed' matrix in the row problem

$$X = ZG^T + E_1 \quad , \quad J_1 = E_1 Q E_1^T$$

which will give the unknown state matrix $\hat{Z}$, corresponding to eqn 6.

3. Conclusion. A solution has been given for the least-squares estimate of the double-sided composite problem. If the error criterion is separable, then the solution can be decomposed into two sub-problems which can be solved sequentially.

The double-sided problem has application in two-dimensional curve fitting and prediction problems[4,5]. It can also be used for the solution of the inverse output feedback problem[6], which requires solution for the unknown matrix P in the equation

$$A - \bar{A} = BPC$$

where A and $\bar{A}$ are the open- and closed-loop system matrices respectively for the linear dynamical system S(A,B,C).

4. References

1. NICHOLSON, H.: 'Sequential least-squares prediction based on spectral analysis', Int. Jl. Compr. Math., Sect B, 1972, 3, pp.257-270.
2. BELLMAN, R.: 'Introduction to matrix analysis', McGraw Hill, NY, 1960, 1970.
3. MARCUS, M.: 'Finite dimensional multilinear algebra', Marcel Dekker, NY, 1973.
4. FERNANDO, K.V.M., and NICHOLSON, H.: 'Discrete double-sided Karhunen-Loeve expansion', Proc.IEE, 1980, 127D, (4), pp.155-160.
5. FERNANDO, K.V.M., and NICHOLSON, H.: 'Two-dimensional curve-fitting and prediction using spectral analysis', Proc.IEE, 1982, 129D, in press.
6. PARASKEVOPOULOS, P.N., and KING, R.E.: 'A Kronecker product approach to pole assignment by output feedback', Int. Jl. Control, 1976, 24, 3, pp.325-334.

APPENDIX 2

The double-sided least-squares problem with diagonal constraints

Abstract The double-sided least-squares problem with a constrained parameter matrix is formulated and solved using multilinear products.

1. Introduction The double-sided least-squares problem has been defined with the relationship[1]

$$Y = HZG^T + E \tag{1}$$

with $Y \in M_{m,n}$, $H \in M_{m,k}$, $Z \in M_{k,\ell}$, $G \in M_{n,\ell}$, where Y is the observed data matrix, Z is the state or parameter matrix, H and G are known maximal rank matrices and E is the residual error matrix.

The quadratic error function for minimization is taken as

$$J = e^T(Q \otimes P)e = e^T S e$$
$$= Q * E^T P E = P * E Q E^T$$

where $\otimes$ is the Kronecker product[2,3] and $e = \mathrm{vec}(E)$, where $\mathrm{vec}(\cdot)$ is the operator which stacks columns of a matrix into a column vector. The operation * denotes the matrix inner product defined by

$$A * B = B * A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij} \quad , \quad A,B \in M_{m,n}$$

The least-squares formulation is now extended to consider the case with Z constrained to be a diagonal matrix. The importance of such constrained expressions are due to the form of spectral expansions of matrices. For example, any square symmetrical matrix W with distinct eigenvalues can be represented by

$$W = \sum_{i=1}^{m} d_{ii} u_i u_i^T = UDU^T \quad , \quad W,D,U \in M_{m,m}$$

where $U = [u_i]$ and $D = \text{diag}[d_{ii}]$ can be identified as eigenvector and eigenvalue matrices respectively. Alternatively, U could be a triangular matrix with diagonal elements equal to unity. Such decompositions are useful in Gaussian elimination techniques and related problems[5]. If the matrix W is rectangular, then it can be represented by the singular value decomposition[4,5]

$$W = \sum_{i=1}^{r} d_{ij} u_i v_i^T = UDV^T \quad , \quad r = \text{rank } (W)$$

with dimensions $W \in M_{m,n}$, $U \in M_{m,r}$, $D \in M_{r,r}$, $V \in M_{n,r}$, where U and V are eigenvector matrices of WW^T and W^TW, respectively, and D is the square-root of the eigenvalue matrix $\Lambda(WW^T)$. Thus

$$WW^T = U\Lambda U^T \quad , \quad W^TW = V\Lambda V^T$$

$$D = \Lambda^{\frac{1}{2}} \quad , \quad \Lambda,D \in M_{r,r}$$

This spectral decomposition is now proposed for the representation of another matrix Y which is assumed to have approximately the same eigenvector matrices U and V as the matrix W. The matrices W and Y could, for example, contain the observed outputs from a plant and its model or the outputs from a system model and its reduced order form. A diagonal eigenvalue matrix can be found by assuming that the modes or eigenvectors of the original matrix W exist in the matrix Y. The measurement equation is then written in the format

$$Y = UZV^T + E$$

where Z^2 will give the eigenvalues or 'energy' values of the spectrum of (Y^TY) with respect to the spectrum of (W^TW).

2. The diagonal problem If the matrix Z is diagonal, then eqn 1 can be written in the form

$$Y = HZG^T + E \quad , \qquad Z \in M_{k,k}$$

$$= \sum_{i=1}^{k} z_{ii} F_i + E \qquad (2)$$

where $F_i = h_i g_i^T \quad , \qquad F_i \in M_{m,n}$

If the matrix set F_i, i = 1..k, is arbitrary, then the least-squares problem is similar to that posed in reference 7.

The Kronecker product can be used in the expansion of eqn 2 instead of the dyadic product $h_i g_i^T$ - both are tensor products and hence equivalent. Then

$$y = \sum_{i=1}^{k} z_{ii} g_i \otimes h_i + e$$

$$= Fz_d + e$$

where y = vec(Y), e = vec(E), $F \in M_{mn,k}$ and z_d is the column vector formed from diagonal elements of the matrix Z,

$$z_d = (z_{11}, \ldots, z_{ii}, \ldots, z_{kk})^T$$

and $F = [g_1 \otimes h_1 \ \cdots \ g_i \otimes h_i \ \cdots \ g_k \otimes h_k]$

The unbiassed least-squares estimate of the vector z_d is then given by

$$\hat{z}_d = (F^T S F)^{-1} F^T S y$$

and the matrix estimate $\hat{Z}$ can be formed using the above definition.

The symmetrical matrix $F^T S F$ is of the form

$$F^TSF = \begin{pmatrix} (g_1^T \otimes h_1^T)(Q \otimes P)(g_1 \otimes h_1) \ldots\ldots (g_1^T \otimes h_1^T)(Q \otimes P)(g_k \otimes h_k) \\ \ldots\ldots\ldots (g_i^T \otimes h_i^T)(Q \otimes P)(g_j \otimes h_j) \ldots\ldots\ldots \\ (g_k^T \otimes h_k^T)(Q \otimes P)(g_1 \otimes h_1) \ldots\ldots (g_k^T \otimes h_k^T)(Q \otimes P)(g_k \otimes h_k) \end{pmatrix}$$

and the ij element can be reduced to the form

$$(F^TSF)_{ij} = (g_i^T Q g_j)(h_i^T P h_j)$$

The matrix F^TSF can then be decomposed into two symmetrical matrices, with

$$F^TSF = A \circ B = B \circ A$$

where o denotes the Hadamard product[3,4,8] or the Schur product[2], which is formed from element-by-element multiplication of the matrices A and B. Thus

$$(F^TSF)_{ij} = a_{ij}b_{ij} \quad , \quad A = H^TPH \quad , \quad B = G^TQG$$

Since the matrices A and B are positive definite, with H,G,P and Q of maximal rank, then by Schur's lemma[3,4,8], the matrix F^TSF is also positive definite and thus nonsingular. The occurrence of the Hadamard product A o B is not unexpected since it is a principal submatrix of the Kronecker product matrix $A \otimes B$[3].

The vector F^TSy is similarly given by

$$F^TSy = (h_1^T PYQg_1 \ \ldots\ h_i^T PYQg_i \ \ldots\ h_k^T PYQg_k)^T$$

3. Conclusions The diagonally constrained double-sided least-squares problem has been formulated and methods of solution indicated. An application in two-dimensional curve fitting and prediction is being investigated[6].

4. References

1. Fernando, K.V.M., and Nicholson, H.: 'Double-sided least-squares problem', Electronics Letters, 1979, 15, 20, pp.624-625.

2. Bellman, R.: 'Introduction to matrix analysis', McGraw-Hill, New York, 1970.

3. Marcus, M., and Minc, H.: 'A survey of matrix theory and matrix inequalities', Allyn & Bacon, Boston, 1964.

4. Rao, C.R., and Mitra, S.K.: 'Generalized inverse of matrices and its applications', Wiley, New York, 1971.

5. Forsythe, G.E., and Moler, C.B.: 'Computer solution of linear algebraic systems', Prentice-Hall, Englewood Cliffs, NJ, 1967.

6. Fernando, K.V.M., and Nicholson, H.: 'Two-dimensional curve-fitting and prediction using spectral analysis', Proc.IEE, 1982, 129D, in press.

7. Basu, J.P., and Basu, R.: 'Estimations of proportions of given populations when observable units contain several populations', IEEE Trans. Systems, Man and Cybernetics, 1976, SMC-6, 11, pp.775-777.

8. Halmos, P.R.: 'Finite-dimensional vector spaces', Van Nostrand Reinhold, New York, 1958.

APPENDIX 3

Singular Perturbational Model Reduction in the Frequency Domain

Abstract: Singular perturbational approximations for linear continuous-time and discrete-time systems are developed in the frequency domain. It is shown that the familiar singular perturbational result is an approximation at the origin in the complex plane. However, if the system has multiple time-scale effects, other approximations can be obtained at different locations on the negative real axis to emphasize such behaviour. The relationship between singular perturbational approximations and direct subsystem elimination is also investigated.

1. Introduction

The singular perturbational method has become one of the best popular methods for obtaining reduced-order representations of linear systems[2,3]. Some of the advantages of this technique are due to its simplicity, the consistency with mathematical models of some physical systems, and the relationship with other established methods such as aggregation[4], the "dominant mode" methods[7], and the Routh approximation[8]. However, the development of this method has been mostly on an ad hoc basis and all theoretic implications of the method have not been fully explained or understood.

The object of this note is to develop singular perturbational approximations for continuous-time and discrete-time systems in the neighbourhood of the negative real axis in the complex plane. The negative real axis is paramount in singular perturbational studies, since

"fast" and "slow" phenomena depend on the real parts of the poles of the system. It is shown that the usual perturbational result is an approximation at the origin. However, for systems with multiple time-scale effects, the origin need not be the most desired position for approximation. Thus, we may generalize the singular perturbational results by considering approximations in other positions of the negative real axis.

2. The generalized approximation

We consider the linear stable system S(A,B,C) defined by

$$\dot{x}(t) = Ax(t) + Bu(t) \quad , \quad y(t) = Cx(t)$$

which has the transfer function

$$H(s) = C(sI - A)^{-1}B$$

where $s = \sigma+j\omega$ is the complex frequency.

The system can be partitioned in the format,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad , \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

$$C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}$$

where we assume that all submatrices conform to the orders of their subsystems. The transfer function can be written in the form,

$$H(s) = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} sI-A_{11} & -A_{12} \\ -A_{21} & sI-A_{22} \end{bmatrix}^{-1} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

If we use the well known lemma for inverse of partitioned matrices (sometimes known as K-partitioning[1], then the transfer function can be expressed as

$$H(s) = H_1(s) + H_2(s)$$

where
$$H_1(s) = \bar{C}(s)\left[sI-\bar{A}(s)\right]^{-1}\bar{B}(s)$$

$$\bar{A}(s) = A_{11} + A_{12}(sI-A_{22})^{-1}A_{21}$$

$$\bar{B}(s) = B_1 + A_{12}(sI-A_{22})^{-1}B_2$$

$$\bar{C}(s) = C_1 + C_2(sI-A_{22})^{-1}A_{21}$$

$$H_2(s) = C_2(sI-A_{22})^{-1}B_2$$

If the subsystem $S(A_{22},B_2,C_2)$ is stable and is not dominant in the neighbourhood of the frequency $s = \sigma_o$, then the system $S(A,B,C)$ can be approximated by the reduced order representation $S(\bar{A}(\sigma_o),\bar{B}(\sigma_o),\bar{C}(\sigma_o))$ which has the transfer function

$$\bar{H}(s) = \bar{C}(\sigma_o)\left[sI-\bar{A}(\sigma_o)\right]^{-1}\bar{B}(\sigma_o)$$

We call this result the "generalized singular perturbational approximation" at $s = \sigma_o$.

In large-scale system studies, we may approximate the system at different frequencies $s = \sigma_o$ to study the behaviour on different time-scales, provided such multiple time-scale effects are present in the system. However, the subsystem being eliminated (generically denoted by $S(A_{22},B_2,C_2)$ in this study) does not have to be the same subsystem at different frequency approximations.

If the point of approximation is the origin, then we obtain the familiar zeroth order singular perturbational result,

$$\bar{A}(0) = A_{11} - A_{12}A_{22}^{-1}A_{21}$$

$$\bar{B}(0) = B_1 - A_{12}A_{22}^{-1}B_2$$

$$\bar{C}(0) = C_1 - C_2A_{22}^{-1}A_{21}$$

provided that the first moment (dc gain) of the second subsystem which is given by $C_2A_{22}^{-1}B_2$ is "small" or singular compared with that of the approximation. This seems to be more general than the conventional assumption that the second subsystem $S(A_{22},B_2,C_2)$ is "fast".

The non-dominance of the subsystem $S(A_{22},B_2,C_2)$ at a nominal frequency $s = \sigma_o$ can be due to the pole structure as in "fast" subsystems. However, this is not the only possibility and it could be due to the numerator dynamics of the transfer function including any non-minimal phase properties.

3. Direct subsystem elimination

In this section, we study the approximation at the other extreme of the real axis with $\sigma_o \to -\infty$. If the second subsystem $S(A_{22},B_2,C_2)$ is non-dominant at negative infinity, then the generalized approximation is given by

$$\bar{A}(\sigma_o) \to A_{11} \quad , \quad \bar{B}(\sigma_o) \to B_1, \quad C(\sigma_o) \to C_1 \quad , \quad \text{as } \sigma_o \to -\infty$$

Thus, the generalized singular perturbational approximation at negative infinity can be obtained by direct elimination of the second subsystem.

Direct elimination of subsystems [5] and the singular perturbational approximation at the origin [6] have been suggested for model reduction of internally balanced systems. It is interesting to note that both these methods are generalized singular perturbational approximations.

4. Approximations for discrete time systems

For the discrete-time systems $S_d(A,B,C)$ defined by

$$x_{t+1} = Ax_t + Bu_t \quad , \qquad y_t = Cx_t$$

the transfer function can be written in the form,

$$H(z) = C(zI-A)^{-1}B$$

As in the continuous-time case, generalized singular perturbational approximations can be obtained in a similar manner. However, as pointed out by Blankenship[9], the usual approximation at the origin with $z_o = 0$, corresponds to characterization of "fast" behaviour in the discrete-time case rather than "slow" behaviour as in the continuous-time problem.

If the second subsystem $S_d(A_{22},B_2,C_2)$ is stable and is non-dominant around the neighbourhood of $z = 1$, then the system can be approximated by $S_d(\bar{A}(1),\bar{B}(1),\bar{C}(1))$

where
$$\bar{A}(1) = A_{11} + A_{12}(I-A_{22})^{-1}A_{21}$$

$$\bar{B}(1) = B_1 + A_{12}(I-A_{22})^{-1}B_2$$

$$\bar{C}(1) = C_1 + C_2(I-A_{22})^{-1}A_{21}$$

which characterise slow behaviour. However, this is not the only approximation possible and any point on the real axis, but within the unit circle, is a possible candidate frequency for obtaining a reduced-order model.

5. Conclusions

We have shown the feasibility of singular perturbational approximations in the frequency domain and defined generalized singular perturbational approximations valid in the neighbourhood of the real axis in the complex plane. The results developed in this note give an alternative, and a more refined insight into the singular perturbational model reduction problem.

6. References

1. Nicholson, H.: 'Structure of Interconnected Systems', Peregrinus, London, 1978.

2. Kokotovic, P.V., O'Malley, R.E., and Sannuti, P.: 'Singular perturbations and order reduction in control theory - An overview', Automatica, 1976, 12, pp.123-132.

3. Sandell, Jr., N.R., Varaiya, P., Athans, M., Safonov, M.G.: 'Survey of decentralized control methods for large-scale systems', IEEE Trans. Automatic Control, 1978, vol. AC-23, (1), pp.108-128.

4. Aoki, M.: 'Control of large-scale dynamical systems by aggregation', 1968, IEEE Trans. Automatic Control, AC-13, (3), pp.246-253.

5. Moore, B.C.: 'Principal component analysis in linear systems: Controllability, observability, and model reduction', IEEE Trans. Automatic Control, 1981, AC-26, (1), pp.17-32.

6. Fernando, K.V., and Nicholson, H.: 'Singular perturbational model reduction of balanced systems', IEEE Trans. Automatic Control, 1982, AC-27, (2), pp.466-468.

7. Nicholson, H.: 'Dynamic optimization of a boiler', Proc.IEE, 1964, 111, pp.1479-1499.

8. Hutton, M.F., and Friedland, B.: 'Routh approximations for reducing order of linear systems', IEEE Trans. Automatic Control, 1975, AC-20, pp.329-337.

9. Blankenship, G.: 'Singularly perturbed difference equations in optimal control', IEEE Trans. Automatic Control, 1981, AC-26, (4), pp.911-917.

APPENDIX 4

On the Applicability of Routh Approximations and Allied Methods in Model-Order Reduction

Abstract: Routh approximations used in model-order reduction tend to preserve high-frequency behaviour while low-frequency approximations are usually required in control design. This high-frequency bias can be remedied using reciprocal transformations. We demonstrate that if low frequencies are not dominant or unimportant, then reciprocal transformations should be avoided. We also show that some of the disadvantages of the Routh method recently reported in the literature are avoidable.

1. Introduction

The Routh approximation method of Hutton and Friedland[1] and similar techniques known variously as Hurwitz approximations and Routh-Hurwitz approximations[4,5,6] provide a convenient and simple procedure to obtain reduced-order representations of linear systems described by transfer functions. The resultant reduced-order models are always stable provided the original systems are stable. The main peculiarity of this method is the tendency to preserve high-frequency behaviour of the system at the expense of the low frequencies. However, most model-order reduction methods are geared to give low-frequency approximations rather than high-frequency approximations, a tradition initiated in references 7 and 8 following Rosenbrock's original work on modal control. This practice is

due to the importance of slow-time behaviour in control systems design. Furthermore, in most physical systems low-frequency effects dominate and thus low-frequency approximations are paramount in design procedures. In Routh approximation methods, the high-frequency bias is avoided by using a reciprocal transformation, before and after model-order reduction[1]. This reciprocal transformation, essentially changes low-frequency effects into high-frequency effects and vice versa. Reciprocal transformations are easy to implement and are of the form given by

$$H(s) \qquad \hat{H}(s) = 1/s\ H(1/s)$$

where the transfer functions $H(s)$ and $\hat{H}(s)$ denote the original system and the transformed system, respectively. If the transfer function is of the form,

$$H(s) = \frac{b_1 s^{n-1} + \ldots\ldots + b_n}{a_0 s^n + \ldots\ldots + a_n}$$

then the transformed system is given by

$$\hat{H}(s) = \frac{b_n s^{n-1} + \ldots\ldots + b_1}{a_n s^n + \ldots\ldots + a_0}$$

which only involves reordering of the coefficients of the original transfer function. The use of reciprocal transformations has become the standard (or rather the orthodox) practice in Routh approximations, and often the application of this transformation is not explicitly acknowledged.

The main objective of this Appendix is to indicate the frequent misuse of reciprocal transformations in the literature. It is obvious that if the high frequencies dominate in a system, reciprocal transformations should be avoided and the Routh approximation technique should be used directly. Otherwise, the dominant high frequencies will be attenuated and the resultant reduced-order model will represent the non-dominant low-frequency effects. Similarly, if the high-frequency behaviour is important (say, in understanding the transient response) and a high-frequency approximation is desired (perhaps, in addition to a low-frequency reduced-order representation) then reciprocal transformations should not be used.

Recently, Shamash[2,3] used two examples to discredit the Routh approximation and allied techniques. We investigate these two examples in the context of the use of reciprocal transformations and show that the defects pointed out by Shamash can be easily rectified.

To avoid ambiguities, we call the technique the Direct Routh Approximation (DRA) if reciprocal transformations are not used. Otherwise, we call it the Reciprocal Routh Approximation (RRA).

2. Illustrative Examples

Example 1: Shamash[3] used the following problem as a counter-example for the Routh approximation method. The transfer function is of the form,

$$H(s) = \frac{100s^2 + 1100s + 1000}{s^3 + 111s^2 + 1110s + 1000} = \frac{100(s+1)(s+10)}{(s+1)(s+10)(s+100)}$$

where two of the numerator zeros cancel out two of the poles. These additional poles representing low-frequency effects were introduced to 'confuse' those model-order reduction methods which are based on truncation of high-frequency behaviour. Obviously, the preferred first-order reduced model should be of the form,

$$\frac{100}{(s + 100)}$$

which is, in fact, the first-order Pade' approximation of $H(s)$[3].

The first-order RRA is of the form,

$$H_1(s) = \frac{0.9009}{(s + 0.9009)} \qquad \text{(RRA)}$$

which is clearly, a bad approximation as pointed out by Shamash. However, in model-order reduction of large-scale systems, this will not be evident and the validity has to be checked by accounting for the impulse energy of the system. The impulse energy, defined by

$$\|h\|^2 = \int_0^\infty h^2(t)\, dt$$

where h(t) is the impulse response of the system H(s), is given by the α and β parameters of the system[1] in the form,

$$\|h\|^2 = \sum_{i=1}^{n} \frac{\beta_i^2}{2\alpha_i}$$

These parameters occur in Alpha and Beta tables which have to be computed in the Routh approximation method[1].

The value of the impulse energy of the original system H(s) is given by $\|h\|^2 = 50$ while the value for $H_1(s)$ is $\|h_1\|^2 = 50/111$. This conclusively indicates that the reduced-order result $H_1(s)$ is a bad approximation. In this case, energy loss is due to high-frequency 'leakage' if signal processing terminology is used.

To investigate the high-frequency behaviour, we have computed the DRA as,

$$G_1(s) = \frac{100}{s + 111} \qquad \text{(DRA)}$$

Although, this approximation is not as good as the Pade' solution, the result is reasonable. The impulse energy of the approximation is given by $\|g_1\|^2 = 50(100/111)$ which is near the value of the original system. Thus, by comparing energy values, we may conclude that high frequencies dominate in this system and that the DRA gives the best overall approximation. By computing the RRA as well as the DRA and their impulse energies, we have avoided the disadvantages pointed out by Shamash.

Example 2: The following transfer function

$$H(s) = \frac{8169.13s^3 + 50664.97s^2 + 9984.32s + 500}{100s^4 + 10520s^3 + 32101s^2 + 10105s + 500}$$

$$= \frac{81.6913(s + 6.004)(s + 0.1009 + j0.0025)(s + 0.1009 - j0.0025)}{(s + 100)(s + 5)(s + 0.1)^2}$$

was also investigated by Shamash[2,3]. Approximate cancellation of the poles at -0.1 is possible and the preferred second-order approximation should be of the form

$$H(s) \qquad \frac{81.6913(s + 6.004)}{(s + 100)(s + 5)}$$

Clearly, the high-frequency effects are dominant in this system. The RRA is of the form

$$H_2(s) = \frac{0.1936s + 0.009694}{(s^2 + 0.1959s + 0.009694)} \qquad \text{(RRA)}$$

$$= \frac{0.1936(s + 0.05007)}{(s + 0.09796 + j0.009921)(s + 0.09796 - j0.009921)}$$

and obviously, it represents the low-frequency effects. The impulse energy for this representation is given by $\|h_2\|^2 = 0.1204$ which indicates high leakage when compared with $\|h\|^2 = 34.07$ for the original system.

The DRA is given by

$$G_2(s) = \frac{81.69s + 506.6}{(s^2 + 105.2s + 520.0)} \qquad \text{(DRA)}$$

$$= \frac{81.69(s + 6.201)}{(s + 5.201)(s + 100.0)}$$

which is a high-frequency approximation. The validity of the DRA as an overall approximation is evident from the impulse energy $\|g_2\|^2 = 34.06$ which is near the original value.

These results demonstrate that Routh approximations can give reasonable results even when high-frequency effects are dominant.

3. Conclusions

We have demonstrated that reciprocal transformations should not always be used in obtaining reduced-order models if the high-frequency behaviour of the system is dominant or important. This result, in retrospect, is very obvious. By methodically computing the Reciprocal Routh Approximations and the Direct Routh Approximations and their impulse energy values, the pitfalls reported in the literature concerning Routh approximations can be avoided.

4. References

1. Hutton, M.F., and Friedland, B.: 'Routh approximations for reducing time-invariant systems', IEEE Trans. Automatic Control, 1975, AC-20, (3), pp.329-337.
2. Shamash, Y.: 'Viability of methods for generating stable reduced order models', IEEE Trans. Automatic Control, 1981, AC-26, (6), pp.1285-1286.
3. Shamash, Y.: 'Failure of the Routh-Hurwitz method of reduction', IEEE Trans. Automatic Control, 1980, AC-25, (2), pp.313-314.
4. Appiah, R.K.: 'Linear model reduction using Hurwitz polynomial approximation', Int. J. Control, 1978, 28, pp.477-488.
5. Shamash, Y.: 'Note on equivalence of the Routh and Hurwitz methods of approximation', Int. J. Control, 1979, 30, pp.899-900.
6. Langholz, G., and Feinmesser, D.: 'Model reduction by Routh approximations', Int. J. Systems Science, 1978, vol.9, pp.493-496.
7. Davison, E.J.: 'A method for simplifying linear dynamic systems', IEEE Trans. Automatic Control, 1966, AC-11, (1), pp.93-101.
8. Nicholson, H.: 'Dynamic optimisation of a boiler', Proc.IEE, 1964, 111, (8), pp.1479-1499.