

This article was downloaded by:[Rogers, E.]
On: 5 June 2008
Access Details: [subscription number 793777376]
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Control

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t713393989>

Homotopy methods for solving the optimal projection equations for the H_2 reduced order model problem

Dragan Žigić^a, Layne T. Watson^a, Emmanuel G. Collins Jr.^b; Dennis S. Bernstein^b

^a Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, VA, U.S.A.

^b Harris Corporation, Government Aerospace Systems Division, Melbourne, FL, U.S.A.

Online Publication Date: 01 January 1992

To cite this Article: Žigić, Dragan, Watson, Layne T., Collins Jr., Emmanuel G. and Bernstein, Dennis S. (1992) 'Homotopy methods for solving the optimal projection equations for the H_2 reduced order model problem', International Journal

of Control, 56:1, 173 — 191

To link to this article: DOI: 10.1080/00207179208934308
URL: <http://dx.doi.org/10.1080/00207179208934308>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Homotopy methods for solving the optimal projection equations for the H_2 reduced order model problem

DRAGAN ŽIGIĆ†, LAYNE T. WATSON†, EMMANUEL G. COLLINS, JR‡ and DENNIS S. BERNSTEIN‡

The optimal projection approach to solving the H_2 reduced order model problem produces two coupled, highly nonlinear matrix equations with rank conditions as constraints. Owing to the resemblance of these equations to standard matrix Lyapunov equations, they are called modified Lyapunov equations. The algorithms proposed herein utilize probability-one homotopy theory as the main tool. It is shown that there is a family of systems (the homotopy) that makes a continuous transformation from some initial system to the final system. With a carefully chosen initial problem, a theorem guarantees that all the systems along the homotopy path will be asymptotically stable, controllable and observable. One method, which solves the equations in their original form, requires a decomposition of the projection matrix using the Drazin inverse of a matrix. It is shown that the appropriate inverse is a differentiable function. An effective algorithm for computing the derivative of the projection matrix that involves solving a set of Sylvester equations is given. Another class of methods considers the equations in a modified form, using a decomposition of the pseudogramians based on a contragredient transformation. Some freedom is left in making an exact match between the number of equations and the number of unknowns, thus effectively generating a family of methods.

1. Introduction

Hyland and Bernstein (1985) considered the quadratic (H_2) reduced order model problem, which is to find a reduced order model for a given continuous time stationary linear system which minimizes a quadratic model error criterion. The necessary conditions for the optimal reduced order model are given in the form of two modified Lyapunov equations, matrix equations which resemble the (linear) matrix Lyapunov equations, but are highly nonlinear and mutually coupled.

Among many different approaches for finding reduced order models are component cost analysis (Skelton 1980, Skelton and Hughes 1980, Skelton and Yousuff 1983), balancing (Moore 1981, Pernebo and Silverman 1982), Hankel-norm approximation (Kung and Lin 1981 a, b), aggregation (Aoki 1968, Kwong 1982), non-minimal partial realization (Hickin and Sinha 1980), projection methods (De Villemagne and Skelton 1987) and the optimal reduction method of Wilson (1970). Some other applications of the optimal projection approach include the H_2/H_∞ model reduction problem (Haddad and Bernstein 1989), the fixed order dynamic compensation problem (Hyland and Bernstein 1984) and the reduced order state estimation problem (Bernstein and Hyland 1985). A homotopy based algorithm for

Received 20 May 1991. Revised 7 August 1991.

† Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, U.S.A.

‡ Harris Corporation, Government Aerospace Systems Division, Melbourne, FL 32902, U.S.A.

solving the fixed order dynamic compensation problem is given by Richter and Collins (1989). Unlike that algorithm, the algorithm proposed here calculates the jacobian matrix of the homotopy map explicitly, resulting in more numerically robust and efficient homotopy path tracking.

The complete statement of the reduced order model problem is given in § 2; § 3 explains the basics of probability-one homotopy theory. Section 4 gives a way for constructing an initial problem. Section 5 gives an algorithm for computing a contragredient transformation. The method based on the Drazin inverse is presented in § 6. Methods based on decompositions of pseudogramians are given in § 7. Numerical results obtained by solving a number of model reduction problems are given in § 8; § 9 gives a conclusion.

2. Statement of the problem

Given the controllable and observable time invariant continuous time system

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (1)$$

$$y(t) = Cx(t) \quad (2)$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{l \times n}$, the goal is to find, for given $n_m < n$, a reduced order model

$$\dot{x}_m(t) = A_m x_m(t) + B_m u(t)$$

$$y_m(t) = C_m x_m(t)$$

where $A_m \in \mathbb{R}^{n_m \times n_m}$, $B_m \in \mathbb{R}^{n_m \times m}$, $C_m \in \mathbb{R}^{l \times n_m}$, which minimizes the quadratic model-reduction criterion

$$J(A_m, B_m, C_m) \equiv \lim_{t \rightarrow \infty} E[(y - y_m)' R (y - y_m)]$$

where the input $u(t)$ is white noise with positive definite intensity V and R is a positive definite weighting matrix.

It is assumed that A is asymptotically stable and diagonalizable, and a solution (A_m, B_m, C_m) is sought in the set

$$A_+ = \{(A_m, B_m, C_m): A_m \text{ is stable, } (A_m, B_m) \text{ is controllable}$$

and (A_m, C_m) is observable\}.

Definition 1: Given symmetric positive semidefinite matrices $\hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ such that $\text{rank}(\hat{Q}) = \text{rank}(\hat{P}) = \text{rank}(\hat{Q}\hat{P}) = n_m$, matrices $G, \Gamma \in \mathbb{R}^{n_m \times n}$ and positive semisimple $M \in \mathbb{R}^{n_m \times n_m}$ are called a (G, M, Γ) -factorization (projective factorization) of $\hat{Q}\hat{P}$ if

$$\hat{Q}\hat{P} = G^1 M \Gamma$$

$$\Gamma G^1 = I_{n_m}$$

Positive semisimple means similar to a symmetric positive definite matrix. \square

The following theorem from Hyland and Bernstein (1985) gives necessary conditions for the optimal solution to the reduced order model problem.

Theorem 1: Suppose $(A_m, B_m, C_m) \in A_+$ solves the optimal model-reduction problem. Then there exist symmetric positive semidefinite matrices $\hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ such that for some projective factorization of $\hat{Q}\hat{P}$, A_m, B_m and C_m are given by

$$A_m = \Gamma A G^t \tag{3}$$

$$B_m = \Gamma B \tag{4}$$

$$C_m = C G^t \tag{5}$$

and such that, with $\tau \equiv G^t \Gamma$ the following conditions are satisfied:

$$0 = \tau[A\hat{Q} + \hat{Q}A^t + BVB^t] \tag{6}$$

$$0 = [A^t\hat{P} + \hat{P}A + C^tRC]\tau \tag{7}$$

$$\text{rank}(\hat{Q}) = \text{rank}(\hat{P}) = \text{rank}(\hat{Q}\hat{P}) = n_m \tag{8}$$

The equations (6)–(7) can be written in an equivalent form:

$$A\hat{Q} + \hat{Q}A^t + \tau BVB^t + BVB^t\tau^t - \tau BVB^t\tau^t = 0$$

$$A^t\hat{P} + \hat{P}A + \tau^t C^tRC + C^tRC\tau - \tau^t C^tRC\tau = 0$$

The matrices \hat{Q} and \hat{P} are called the *controllability* and *observability pseudo-gramians*, respectively, since they are analogous to the Gramians G_c and G_o which satisfy the dual Lyapunov equations

$$AG_c + G_cA^t + BVB^t = 0$$

$$A^tG_o + G_oA + C^tRC = 0$$

τ is an oblique projection (idempotent) operator since $\tau^2 = \tau$. The projection matrix τ can be expressed as

$$\tau = (\hat{Q}\hat{P})(\hat{Q}\hat{P})^*$$

where $(\hat{Q}\hat{P})^*$ is the Drazin inverse defined in § 6.

3. Probability-one homotopy methods

Homotopies are a traditional part of topology, and have found significant application in nonlinear functional analysis and differential geometry (Watson 1986). Homotopy methods are globally convergent, which distinguishes them from most iterative methods, which are only locally convergent. The general idea of homotopy methods is to make a continuous transformation from an initial problem, which can be solved trivially, to the target problem.

Following Watson (1989), the theoretical foundation of all probability-one globally convergent homotopy methods is given in the following differential theorem.

Definition 2: Let $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^p$ be open sets, and let $\rho : U[0, 1) \times V \rightarrow \mathbb{R}^p$ be a C^2 map. ρ is said to be transversal to zero if the jacobian matrix $D\rho$ has full rank on $\rho^{-1}(0)$.

Theorem 2: If $\rho(a, \lambda, x)$ is transversal to zero, then for almost all $a \in U$ the map

$$\rho_a(\lambda, x) = \rho(a, \lambda, x)$$

is also transversal to zero; i.e. with probability one the jacobian matrix $D\rho_a(\lambda, x)$ has full rank on $\rho_a^{-1}(0)$.

This recipe for constructing a globally convergent homotopy algorithm to solve the nonlinear system of equations

$$f(x) = 0$$

where $f: \mathbb{R}^p \rightarrow \mathbb{R}^p$ is a C^2 map, is as follows: for an open set $U \subset \mathbb{R}^m$ construct a C^2 homotopy map $\rho: [0, 1) \times \mathbb{R}^p \rightarrow \mathbb{R}^p$ such that

- (1) $\rho(a, \lambda, x)$ is transversal to zero;
- (2) $\rho_a(0, x) = \rho(a, 0, x) = 0$ is trivial to solve and has a unique solution x_0 ;
- (3) $\rho_a(1, x) = f(x)$;
- (4) $\rho_a^{-1}(0)$ is bounded.

Then for almost all $a \in U$ there exists a smooth zero curve γ of ρ_a (1-manifold of points where $\rho_a = 0$), along which the jacobian matrix $D\rho_a$ has rank p , emanating from $(0, x_0)$ and reaching a zero \bar{x} of f at $\lambda = 1$. This zero curve γ does not intersect itself, is disjoint from any other zeros of ρ_a , and has finite arc length in every compact subset of $[0, 1) \times \mathbb{R}^p$. Furthermore, if $Df(\bar{x})$ is non-singular, then γ has finite arc length. The general idea of the algorithm is to follow the zero curve γ emanating from $(0, x_0)$ until a zero \bar{x} of $f(x)$ is reached (at $\lambda = 1$).

The zero curve γ is tracked by the normal flow algorithm (Watson *et al.*, 1987), a predictor–corrector scheme. In the predictor phase, the next point is produced using Hermite cubic interpolation. Starting at the predicted point, the corrector iteration involves computing (implicitly) the Moore–Penrose pseudo-inverse of the jacobian matrix at each point. The most complex part of the homotopy algorithm is the computation of the tangent vectors to γ , which involves the computation of the kernel of the $p \times (p + 1)$ jacobian matrix $D\rho_a$. The kernel is found by computing the QR factorization of $D\rho_a$, and then using back substitution. This strategy is implemented in the mathematical software package HOMPACK (Watson *et al.* 1987), which was used for the curve tracking here.

Two different homotopy maps are used for solving the optimal projection equations. When the initial problem, $g(x; a) = 0$, can be solved, then the homotopy map is (Watson 1990)

$$\rho_a(\lambda, x) = F(a, \lambda, x) \equiv \lambda f(x) + (1 - \lambda)g(x; a) \quad (9)$$

where $f(x) = 0$ is the final problem, and a is a parameter vector used in defining the function g . In general ρ_a need not be a simple convex combination of two functions f and g , so it is not always the case that $\rho_a(0, x) = g(x; a)$.

When the initial problem is not solved exactly, i.e. $g(x_0; b) \neq 0$, then the map is a Newton homotopy (Smale 1976)

$$\rho_a(\lambda, x) = F(b, \lambda, x) - (1 - \lambda)F(b, 0, x_0) \quad (10)$$

where the parameter vector $a = (b, x_0)$ is now different from the parameter vector a in (9). For $\lambda = 0$, $\rho_a(0, x_0) = F(b, 0, x_0) = 0$, and for $\lambda = 1$, $\rho_a(1, x) = F(b, 1, x) = f(x) = 0$.

For the homotopies considered here, the theoretical verification of properties (1) and (4) is highly technical and was not attempted. Examples of such proofs for other classes of problems are to be found in Watson (1986), Watson (1989) and Watson (1990); property (1) is generically true, but verifying property (4) for homotopies like (9) and (10) would be rather difficult.

4. Defining an initial system

While with homotopy algorithms in general an initial problem can be chosen practically at random, this problem has some special limitations. The reason is that Theorem 1 provides necessary conditions on a solution only under certain assumptions. In other words, every intermediate problem solution satisfies these equations only if the system is asymptotically stable, controllable and observable. While the absence of these features does not automatically mean that the intermediate problem solution will not satisfy the equations, it is clearly better to define a homotopy path in such a way that each problem along it corresponds to an asymptotically stable, controllable and observable system. Existence of a solution to the H_2 reduced order problem follows from the work of Spanos *et al.* (1990). Theorem 3 defines a class of initial systems such that these conditions are satisfied.

Theorem 3: For the given system (1)–(2), let $A = X\Lambda X^{-1}$, with Λ diagonal. Define $D = X\Omega X^{-1}$ for any diagonal matrix $\Omega = \text{diag}(\omega_1, \dots, \omega_n)$, such that all ω_i , for $i = 1, 2, \dots, n$, are in the open left half-plane. Then for almost all such D every convex combination $(A(\alpha), B, C)$ of the systems (D, B, C) and (A, B, C) will be asymptotically stable, controllable and observable.

Proof:

Stability. Since

$$\begin{aligned} A(\alpha) &= \alpha A + (1 - \alpha)D = \alpha X\Lambda X^{-1} + (1 - \alpha)X\Omega X^{-1} \\ &= X[\alpha\Lambda + (1 - \lambda)\Omega]X^{-1} = X\Lambda(\alpha)X^{-1} \end{aligned}$$

and $\Lambda(\alpha)$ is diagonal with all diagonal elements in the open left half-plane for $\alpha \in [0, 1]$, the matrix $A(\alpha)$ is asymptotically stable for $\alpha \in [0, 1]$.

Controllability. Let $B = X\hat{B}$. Consider the controllability matrices in the coordinate system obtained by the change of coordinates defined by the matrix X . In that coordinate system the controllability matrix is

$$\hat{B}_c(\alpha) = (\hat{B} \quad \Lambda(\alpha)\hat{B} \quad \dots \quad \Lambda^{n-1}(\alpha)\hat{B})$$

for $\alpha \in [0, 1]$. For almost all choices of Ω the eigenvalues of the diagonal matrix $\Lambda(\alpha)$ will be distinct for $\alpha \in [0, 1)$. That $\text{rank } \hat{B}_c(\alpha) = n$ follows from the results in Brogan (1985). A direct argument follows. Let $\lambda_1, \dots, \lambda_n$ be the (distinct) eigenvalues of $\Lambda(\alpha)$ for some α , and

$$\hat{B} = \begin{bmatrix} \hat{b}_{11} & \hat{b}_{12} & \dots & \hat{b}_{1m} \\ \hat{b}_{21} & \hat{b}_{22} & \dots & \hat{b}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{b}_{n1} & \hat{b}_{n2} & \dots & \hat{b}_{nm} \end{bmatrix}$$

Since the system is controllable for $\alpha = 1$, it follows that each row of \hat{B} has at least one non-zero element, because otherwise $\hat{B}_c(1)$ would have a zero row.

Reorder the rows and columns of \hat{B}_c in the following way: for each row of \hat{B} that has a non-zero element in the first column (suppose there are p_1 of them), exchange rows so that the selected p_1 rows are at the top. Next, exchange columns in such a way that columns 1, $m + 1, \dots, (p_1 - 1)m + 1$ become the first p_1

columns. This produces a submatrix in the upper left corner which can be expressed as

$$\hat{B}_1 = \begin{bmatrix} \hat{b}_1 & 0 & \dots & 0 \\ 0 & \hat{b}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{b}_{p_1} \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 & \dots & \lambda_1^{p_1-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{p_1-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_{p_1} & \dots & \lambda_{p_1}^{p_1-1} \end{bmatrix}$$

Since $\hat{b}_1, \dots, \hat{b}_{p_1}$ were chosen to be non-zero, and the λ_i are distinct, the Vandermonde matrix is non-singular, and hence \hat{B}_1 is also non-singular.

Repeating this procedure for the remaining rows gives a block upper triangular transformation of $\hat{B}_c(\alpha)$, with diagonal blocks $\hat{B}_1, \dots, \hat{B}_r$, for some $r \leq n$. That means that $\text{rank } \hat{B}_c(\alpha) = n$ and the system is controllable.

Observability. The analogous construction for the observability matrices proves that the system is observable. \square

Theorem 3 holds if A is not diagonalizable, but that proof is considerably more complicated and provides no more insight than the above proof. While the random construction of the matrix D given in Theorem 3 is theoretically plausible, in practice it may not be wise. The reason is that the matrix X is complex in general, which for many choices of Ω leads to a complex matrix D , which is undesirable. Hence, it is better to construct a matrix D directly such that Ω satisfies the conditions given in Theorem 3.

One simple choice for D is

$$D \equiv -c_1 I + \text{diag} \{ \varepsilon_1, \dots, \varepsilon_n \} \quad (11)$$

where $c_1 > 0$ and ε_i are small random numbers that correspond to the parameter a in the theory. In this case Ω is a small perturbation of $-c_1 I$.

Also, the matrix D can be defined as

$$D \equiv -c_1 I + c_2 A \quad (12)$$

for $c_1, c_2 > 0$. In this case $\Omega = -c_1 I + c_2 A$.

The following strategy can be applied to find a good approximation to a solution of the initial system. Since the matrix D is asymptotically stable, the Lyapunov equation

$$DQ + QD^t + BVB^t = 0 \quad (13)$$

has a unique solution Q . Let $Q = T\Sigma T^t$, where T is orthogonal and

$$\Sigma = \text{diag} \{ \sigma_1, \dots, \sigma_n \}$$

Next, define

$$\Sigma_1 \equiv \text{diag} \{ \sigma_1, \dots, \sigma_{n_m}, 0, \dots, 0 \}, \quad Q_0 \equiv T\Sigma_1 T^t$$

If Q_0 is substituted for Q in (13), the equation will not be satisfied, but in general, if σ_i are sufficiently small, it will not be very different from zero. A similar procedure can be applied to compute P_0 that will 'almost' satisfy the equation

$$D^t P + PD + C^t RC = 0$$

The point $x_0 = (Q_0, P_0)$ chosen in this way may lead to small values of the initial system $g(x_0, a)$. Also, this x_0 can be used as the initial guess for a quasi-Newton algorithm which may find a solution to the initial problem

$$\begin{aligned} \tau[DQ + QD^t + BVB^t] &= 0 \\ [D^tP + PD + C^tRC] \tau &= 0 \end{aligned}$$

5. Contragredient transformation

The following lemma from Hyland and Bernstein (1985), which is a special case of a result in Glover (1984), gives an algorithm for simultaneous reduction of pseudogramians to diagonal forms using a contragredient transformation. The constructive proof given here is different from that in Hyland and Bernstein (1985), and the construction provides both insight and an outline for the numerical computation of the contragredient transformation.

Lemma 1 (Hyland and Bernstein 1985): *Let symmetric positive semidefinite Q , $P \in \mathbb{R}^{n \times n}$ satisfy*

$$\text{rank}(Q) = \text{rank}(P) = \text{rank}(QP) = n_m \tag{14}$$

where $n_m \leq n$. Then, there exists a non-singular $W \in \mathbb{R}^{n \times n}$ (contragredient transformation) and positive definite diagonal $\Sigma, \Omega \in \mathbb{R}^{n_m \times n_m}$ such that

$$Q = W \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} W^t, \quad P = W^{-t} \begin{bmatrix} \Omega & 0 \\ 0 & 0 \end{bmatrix} W^{-1}$$

Proof: Since P is positive semidefinite and symmetric there exists orthogonal $V \in \mathbb{R}^{n \times n}$ such that

$$P = V \begin{bmatrix} D_1 & 0 \\ 0 & 0 \end{bmatrix} V^t$$

where $D_1 \in \mathbb{R}^{n_m \times n_m}$ is diagonal and positive definite. Let

$$T_1 \equiv \begin{bmatrix} D_1^{-1/2} & 0 \\ 0 & I \end{bmatrix}$$

Then

$$T_1^t P T_1 = \begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad T_1^{-1} Q T_1^{-t} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^t & Q_2 \end{bmatrix}$$

where $Q_1 \in \mathbb{R}^{n_m \times n_m}$. Since $\text{rank}(QP) = \text{rank}(P)$ implies that Q is one-to-one on the eigenspace of P corresponding to positive eigenvalues, and the quadratic form $x^t Q_1 x$ corresponds to the quadratic form $y^t Q y$ restricted to this eigenspace, and $y^t Q y = 0$ implies $Q y = 0$ for symmetric positive semidefinite Q , it follows that Q_1 is also symmetric positive definite. Therefore, there exists a positive definite diagonal $D_2 \in \mathbb{R}^{n_m \times n_m}$ and orthogonal $U \in \mathbb{R}^{n_m \times n_m}$ such that $Q_1 = U D_2 U^t$. Let

$$T_2 \equiv \begin{bmatrix} U & 0 \\ Q_{12}^t U D_2^{-1} & -I \end{bmatrix}$$

Then

$$\begin{aligned} T_2^t T_1^t P T_1 T_2 &= \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \\ T_2^{-1} T_1^{-1} Q T_1^{-t} T_2^{-t} &= \begin{bmatrix} D_2 & 0 \\ 0 & Q_2 - Q_{12}^t U D_2^{-1} U^t Q_{12} \end{bmatrix} \\ &= \begin{bmatrix} D_2 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

The last equality is a consequence of the rank conditions (14). If $W \equiv T_1 T_2$ then

$$Q = W \begin{bmatrix} D_2 & 0 \\ 0 & 0 \end{bmatrix} W^t, \quad P = W^{-1} \begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix} W^{-1} \quad (15)$$

which completes the proof. \square

Remark 1 (Hyland and Bernstein 1985): Let Q and P be as in Lemma 1. Then there exists a non-singular $U \in \mathbb{R}^{n \times n}$ and positive definite diagonal $\Lambda \in \mathbb{R}^{n_m \times n_m}$ such that

$$Q = U \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} U^t, \quad P = U^{-t} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} U^{-1}$$

Proof: The statement follows from (15) using

$$U \equiv W \begin{bmatrix} D_2^{1/4} & 0 \\ 0 & I \end{bmatrix} \quad \square$$

The following lemma defines a projective factorization of the product of the pseudogramians and gives an effective way to compute it using a contragredient transformation. The proof here is slightly different from that of Hyland and Bernstein (1985).

Lemma 2 (Hyland and Bernstein 1985): Let symmetric positive semidefinite $\hat{Q}, \hat{P} \in \mathbb{R}^{n \times n}$ satisfy the rank conditions (14). Then, there exist $G, \Gamma \in \mathbb{R}^{n_m \times n}$ and positive semisimple (positive semisimple means similar to a symmetric positive definite matrix) $M \in \mathbb{R}^{n_m \times n_m}$ such that

$$\hat{Q}\hat{P} = G^t M \Gamma \quad (16)$$

$$\Gamma G^t = I_{n_m} \quad (17)$$

Proof: Owing to Remark 1, there exist non-singular $W \in \mathbb{R}^{n \times n}$ and positive definite diagonal $\Sigma \in \mathbb{R}^{n_m \times n_m}$ such that

$$\hat{Q} = W \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} W^t, \quad \hat{P} = W^{-t} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} W^{-1} \quad (18)$$

The equations (18) can be expressed in the equivalent form

$$\hat{Q} = W_1 \Sigma W_1^t, \quad \hat{P} = U_1^t \Sigma U_1 \quad (19)$$

where

$$W = \begin{bmatrix} \overbrace{W_1}^{n_m} & W_2 \end{bmatrix}, \quad W^{-1} = U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \quad (20)$$

From (20) with $G \equiv W^n$ and $\Gamma \equiv U_1$ follow (16)–(17). □

Matrices G , M and Γ from Lemma 2 are a (G, M, Γ) -factorization of (\hat{Q}, \hat{P}) .

6. Homotopy method based on the Drazin inverse

One approach designing a homotopy algorithm for solving the optimal projection equations is to use the decomposition of the projection matrix τ based on the Drazin inverse.

6.1. Theoretical results

Since the homotopy algorithm involves computation of derivatives at each step, it is essential that (6)–(7) be differentiable. The problem of proving the differentiability of the equations (6)–(7) with the rank conditions (8) reduces to the problem of proving the differentiability of the Drazin inverse.

Definition 3: The *index* of $A \in \mathbb{R}^{n \times n}$ is the smallest non-negative integer k such that $\text{im}(A^k) = \text{im}(A^{k+1})$.

If $A \in \mathbb{R}^{n \times n}$ has index k , then $\mathbb{R}^n = \text{im}(A^k) \oplus \text{ker}(A^k)$. Thus any $x \in \mathbb{R}^n$ has the unique decomposition $x = u + v$, $u \in \text{im}(A^k)$, $v \in \text{ker}(A^k)$. Observe further that A is invertible on $\text{im}(A^k)$.

Definition 4: Let $A \in \mathbb{R}^{n \times n}$ have index k , where $x = u + v$, $u \in \text{im}(A^k)$, $v \in \text{ker}(A^k)$. The *Drazin inverse* A^* of A is defined by $A^*x = A_1^{-1}u$, where A_1 is A restricted to the image of A^k .

If $k = 1$, the Drazin inverse A^* is called the *group inverse*, and in fact that is all that is needed here.

Theorem 4 (Campbell and Meyer 1979): *Let $A \in \mathbb{R}^{n \times n}$ have index k . The Drazin inverse A^* of A is the unique matrix A^* such that:*

$$A^*AA^* = A^*$$

$$A^*A = AA^*$$

$$A^{k+1}A^* = A^k$$

Theorem 5 (Campbell and Meyer 1979): *If $A \in \mathbb{R}^{n \times n}$ has index k , there there exist non-singular matrices T and C , and a nilpotent matrix N of index k such that*

$$A = T \begin{bmatrix} C & 0 \\ 0 & N \end{bmatrix} T^{-1} \quad \text{and} \quad A^* = T \begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix} T^{-1}$$

A^* is computed using the Hermite echelon form as described by Campbell and Meyer (1979). Hearon and Evans (1968) give conditions for the differentiability of the Drazin inverse. That theorem is proven here in a simpler way and using different terminology.

Theorem 6: Let $A(t) \in C^s(I)$ for some interval I and

$$\text{rank}(A(t)) = \text{rank}(A^2(t)) = r$$

for each $t \in I$. If, for each $t \in I$, $B(t)$ is the Drazin inverse of $A(t)$, then $B(t) \in C^s(I)$.

Proof: The rank condition implies A has either index 0 or 1. For index 0, $A^* = A^{-1}$ is differentiable and there is nothing to prove. So assume A has index 1. Then for each $t \in I$, it has the Jordan decomposition

$$A = T \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} T^{-1}$$

where A_1 is square and non-singular. Using the characteristic polynomial of A_1 , a divisor of the characteristic polynomial of A , A_1^{-1} can be expressed as a polynomial in A_1 whose coefficients are polynomials in the elements of A :

$$A_1^{-1} = p(A_1)$$

Let

$$B_1 \equiv p(A) = Tp \left(\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} \right) T^{-1} = T \begin{bmatrix} A_1^{-1} & 0 \\ 0 & p(0) \end{bmatrix} T^{-1}$$

Then

$$B \equiv B_1 A B_1 = T \begin{bmatrix} A_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} T^{-1} = A^*$$

is the Drazin inverse of A . Furthermore, since the elements of B are polynomial functions of the elements of A , $B \in C^s(I)$. \square

The derivative τ' of τ is actually computed using the Sylvester equations. Since $\tau = (\hat{Q}\hat{P})(\hat{Q}\hat{P})^*$, it follows that

$$\tau' = (\hat{Q}\hat{P})'(\hat{Q}\hat{P})^* + (\hat{Q}\hat{P})[(\hat{Q}\hat{P})^*]'$$

where everything is directly computable except $[(\hat{Q}\hat{P})^*]'$. From (19),

$$\begin{aligned} \text{rank}(\hat{Q}\hat{P})^2 &= \text{rank}(W_1 \Sigma W_1^1 U_1^1 \Sigma U_1)^2 \\ &= \text{rank}(W_1 \Sigma^4 U_1) = \text{rank}(\Sigma^4) = \text{rank}(\Sigma^2) \\ &= \text{rank}(W_1 \Sigma W_1^1 U_1^1 \Sigma U_1) = \text{rank}(\hat{Q}\hat{P}) = n_m < n \end{aligned}$$

and therefore $\hat{Q}\hat{P}$ has index 1. The following describes a procedure for the computation of the derivative of the Drazin inverse X of a matrix A of index 1.

For given A of index 1, X is the unique matrix that satisfies

$$AXA = A \tag{21}$$

$$XAX = X \tag{22}$$

$$AX = XA \tag{23}$$

Differentiating (21)–(23) yields

$$A'XA + AX'A + AXA' = A' \tag{24}$$

$$X'AX + XA'X + XAX' = X' \quad (25)$$

$$AX' - X'A = XA' - A'X \quad (26)$$

Substituting AX' from (26) into (24) and summing up equations (24)–(26) gives the Sylvester equation

$$\begin{aligned} (A + XA - I)X' + X'(AX - A + A^2) \\ = XA' - A'X - XA'X + A' - XA'A - AXA' \end{aligned} \quad (27)$$

which has a unique solution with probability one due to the randomness in $A(\lambda)$.

Solving (27) for $X' = [(\hat{Q}\hat{P})^*]'$ completes the computation of τ' .

6.2. Description of the algorithm

The following is a description of the algorithm. The algorithm is based on the normal flow algorithm for dense jacobian matrices described by Watson *et al.* (1987), slightly modified here (Steps 13, 14 of the pseudocode below) to handle the rank requirements of the solution (\hat{Q}, \hat{P}) .

The algorithm starts at the point

$$(\lambda, x) = (0, x_0) = (0, \hat{Q}_0, \hat{P}_0)$$

with some $x_0 = (\hat{Q}_0, \hat{P}_0)$ chosen as explained in § 4. Then it follows the zero curve γ of the homotopy map (9) until a point where $\lambda = 1$ is reached.

$F(a, \lambda, x)$ is represented by two equations

$$A(\lambda)\hat{Q} + \hat{Q}A'(\lambda) + \tau BVB' + BVB'\tau' - \tau BVB'\tau' = 0$$

$$A'(\lambda)\hat{P} + \hat{P}A(\lambda) + \tau'C'RC + C'RC\tau - \tau'C'RC\tau = 0$$

where

$$A(\lambda) = \lambda A + (1 - \lambda)D$$

Recall from (11) that D is a function of the parameter vector a .

Since the equations are symmetric, only the upper right triangles are considered, i.e. q_{ij} and p_{ij} are computed only for $j \geq i$. Therefore, the number of variables is $2[n(n+1)/2] = n(n+1)$.

The mathematical software package HOMPACk requires that the user provide routines to evaluate $\rho_a(\lambda, x)$ and the jacobian matrix $D\rho_a$ at each step. While the former is relatively simple, the latter involves considerable computational effort.

The jacobian matrix consists of $n(n+1) + 1$ derivative vectors, which correspond to the partial derivatives with respect to q_{ij} , p_{ij} , and λ . The terms that do not include τ are simple to evaluate analytically. For example,

$$\left[\frac{\partial}{\partial q_{ij}} (\hat{Q}\hat{P}) \right]_{kl} = \frac{\partial}{\partial q_{ij}} \sum_{m=1}^n q_{km} p_{ml} = \delta_{ik} p_{jl} \quad (28)$$

On the other hand, the components of τ' have to be evaluated numerically. Each evaluation involves solving a Sylvester equation (27). Fortunately, since for different q_{ij} and p_{ij} only the right-hand side of the equation changes, the whole process

can be done efficiently. In order to compute $\partial/\partial q_{11}$ all the computations related to (27) are done completely. For all subsequent partial derivative evaluations only the right-hand side of (27) is evaluated and submitted to the procedure that solves a Sylvester equation. The procedure used for that purpose (Bartels and Stewart 1972) supports this approach very efficiently.

The derivatives of $(1 - \lambda)F(a, 0, x_0)$ with respect to q_{ij} and p_{ij} are zero, and the derivative with respect to λ is $-F(a, 0, x_0)$.

When a final solution to (6)–(7) is obtained, the computation of (A_m, B_m, C_m) is completed by applying the formulae (3)–(5), where G and Γ are obtained as explained in Lemma 2.

In summary the whole algorithm is as follows.

- Step 1. Define $D = -cI + \text{diag} \{ \varepsilon_1, \dots, \varepsilon_n \}$, with $c > 0$ and small random ε_i .
- Step 2. Choose Q_0 and P_0 that satisfy (8) as small perturbations of, respectively, $BVB^1/2c$ and $C^1RC/2c$.
- Step 3. Set $\lambda := 0$, $x := x_0$.
- Step 4. Compute Drazin inverse $(\hat{Q}\hat{P})^*$.
- Step 5. Compute $\tau = (\hat{Q}\hat{P})(\hat{Q}\hat{P})^*$.
- Step 6. Evaluate $\rho_a(\lambda, x)$.
- Step 7. (Evaluate $D\rho_a(\lambda, x)$.) For each p_{ij} , q_{ij} such that $j \geq i$, and λ , do Steps 8–11.
 - Step 8. Compute derivatives of terms that do not include τ using analytical formulae similar to (28).
 - Step 9. Compute $[(\hat{Q}\hat{P})^*]'$ using equation (27).
 - Step 10. Complete the computation of τ' as $\tau' := (\hat{Q}\hat{P})'(\hat{Q}\hat{P})^* + (\hat{Q}\hat{P})[(\hat{Q}\hat{P})^*]'$.
 - Step 11. Sum the values obtained in Steps 8 and 10 to the final value of the derivative vector.
- Step 12. Take a step along the curve and obtain $x_1 = (\hat{Q}_1, \hat{P}_1)$.
- Step 13. Compute a contragredient transformation (18) as if x_1 satisfied the rank conditions.
- Step 14. Use formulae (19) to compute $\bar{x}_1 = (\hat{Q}, \hat{P})$.
- Step 15. If $\lambda < 1$, then set $x := \bar{x}_1$, and go to Step 4.
- Step 16. If $\lambda \geq 1$, compute the solution \bar{x}_1 of (6)–(8) at $\lambda = 1$. Obtain G and Γ as explained in Lemma 2.
- Step 17. Compute the reduced order model using equations (3)–(5).

7. Methods based on decompositions of pseudogramians

Homotopy algorithms for solving optimal projection equations can be designed using decompositions of the pseudogramians based on contragredient transformations.

7.1. Descriptions of methods

The equations (6)–(7) can be considered in another, equivalent form. If (6) is multiplied by U_1 from the left, and (7) is multiplied by W_1 from the right, using (19)–(20), the following two equations are obtained:

$$U_1 A W_1 \Sigma W_1^t + \Sigma W_1^t A^t + U_1 B V B^t = 0 \tag{29}$$

$$A^t U_1^t \Sigma + U_1^t \Sigma U_1 A W_1 + C^t R C W_1 = 0 \tag{30}$$

The third equation

$$U_1 W_1 - I = 0 \tag{31}$$

determines the relationship between W_1 and U_1 .

The matrix equations (29)–(31) contain $2nn_m + n_m^2$ scalar equations. On the other side, the only unknowns in (29)–(31), W_1 , U_1 and diagonal Σ , contain $2nn_m + n_m$ variables. Hence, some additional techniques are necessary in order to make an exact match between the number of equations and the number of unknowns.

One approach is to consider Σ to be symmetric and all elements of Σ as unknowns. This is appropriate, since the equations (29)–(31) with a full symmetric Σ can be transformed into equations of the same form with a diagonal Σ by computing

$$\Sigma = T \bar{\Sigma} T^t, \quad \bar{W}_1 = W_1 T, \quad \bar{U}_1 = T^t U_1$$

where $\bar{\Sigma}$ is diagonal and T is orthogonal.

Another approach is to consider the decomposition from the statement of Lemma 1, which leads to the equations

$$U_1 A W_1 \Sigma W_1^t + \Sigma W_1^t A^t + U_1 B V B^t = 0$$

$$A^t U_1^t \Omega + U_1^t \Omega U_1 A W_1 + C^t R C W_1 = 0$$

$$U_1 W_1 - I = 0$$

which also have $2nn_m + n_m^2$ scalar equations. In this case the number of unknowns in W_1 , U_1 and symmetric Σ and Ω is $2nn_m + n_m^2 + n_m$. An additional n_m equations can be obtained, for example, by requiring

$$\sigma_{ii} - \omega_{ii} = 0 \quad \text{for } i = 1, \dots, n_m$$

Alternatively, the number of unknowns can be reduced to $2nn_m + n_m^2$ if the diagonal elements of Ω are actually the diagonal elements of Σ .

7.2. Algorithm

The following is a description of the algorithm for the method determined by the equations (29)–(31). The algorithm is based on the normal flow algorithm for dense jacobian matrices described by Watson *et al.* (1987). Depending on the relative size of $F(a, 0, x_0)$, the algorithm may be modified. If $F(a, 0, x_0)$ is relatively large, computational experience shows that it is desirable (but not theoretically

necessary) to enforce the symmetry of Σ along the homotopy path. This is done by observing that a symmetrized Σ corresponds to *some* homotopy map that *could* have been chosen initially. In effect, x_0 is changed in the homotopy map at each step along the homotopy zero curve γ . Obviously, in that case the homotopy map (10) must be used.

The algorithm uses the homotopy map (9) or (10), where $F(a, \lambda, x)$ is represented by three equations:

$$U_1 A(\lambda) W_1 \Sigma W_1^t + \Sigma W_1^t A^t(\lambda) + U_1 B V B^t = 0 \quad (32)$$

$$A^t(\lambda) U_1^t \Sigma + U_1^t \Sigma U_1 A(\lambda) W_1 + C^t R C W_1 = 0 \quad (33)$$

$$U_1 W_1 - I = 0 \quad (34)$$

A detailed description of the algorithm for evaluating the jacobian matrix $D\rho_a(\lambda, x)$ is given in Appendix B of Žigić (1991). A program that implements this method is given in Appendix C of Žigić (1991).

In summary, the whole algorithm is as follows.

- Step 1. Define D using formula (11) or (12).
- Step 2. Choose a starting point $x_0 = (Q_0, P_0)$ using one of the strategies explained in § 5.2. Compute $(W_1)_0, (U_1)_0$ and Σ_0 using a contragredient transformation.
- Step 3. Set $\lambda := 0, x := x_0 = ((W_1)_0, (U_1)_0, \Sigma_0)$.
- Step 4. Evaluate $\rho_a(\lambda, x)$ given by (9) or (10), and (32)–(34).
- Step 5. Evaluate $D\rho_a(\lambda, x)$.
- Step 6. Take a step along the curve and obtain $x_1 = (W_1, U_1, \Sigma), \bar{\lambda}$.
- Step 7. Compute $\bar{x}_1 = (W_1, U_1, \bar{\Sigma}) = (W_1, U_1, (\Sigma + \Sigma^t)/2)$.
- Step 8. Change the homotopy to

$$F(a, \lambda, x) - (1 - \lambda)v = 0$$

$$\text{where } v = F(a, \bar{\lambda}, \bar{x}_1)/(1 - \bar{\lambda}).$$

- Step 9. If $\bar{\lambda} < 1$, then set $x := \bar{x}_1, \lambda := \bar{\lambda}$, and go to Step 4.
- Step 10. If $\bar{\lambda} \geq 1$, compute the solution \bar{x}_1 at $\bar{\lambda} = 1$. Compute the reduced order model by diagonalizing $\Sigma = T\bar{\Sigma}T^t$.

Note. If $F(a, 0, x_0)$ is small, Steps 7 and 8 can be omitted without a serious loss of efficiency.

8. Numerical results

Three examples are discussed here. The methods were tested on a number of additional examples, reported by Žigić (1991). The results were obtained using the method based on a contragredient transformation that has W_1, U_1 and Σ as unknowns, using the homotopy (9). For all examples $V = R = I$.

Example 1 (Kabamba 1985): The system is given by

$$A = \begin{bmatrix} -0.05 & -0.99 \\ -0.99 & -5000.0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 100 \end{bmatrix}, \quad C = (1 \quad 100)$$

For the starting point

$$y_0 = \begin{bmatrix} 0 \\ 0.0099995 \\ 0.99995 \\ 0.0099995 \\ 0.99995 \\ 0.5 \end{bmatrix}$$

the homotopy algorithm converges to a solution corresponding to the model of order $n_m = 1$ given by

$$A_m = (-4998.078625), \quad B_m = (100.000194), \quad C_m = (100.000194)$$

This model yields the (minimum) cost $J = 96.078058$.

For the starting point

$$y_0 = \begin{bmatrix} 0 \\ 1 \\ 0.25 \\ 1.07 \\ -0.27 \\ 1 \end{bmatrix}$$

the solution found corresponds to the model of order $n_m = 1$ given by

$$A_m = (-0.485152), \quad B_m = (-0.0000011427), \quad C_m = (-0.000000073400)$$

which yields the (maximum) cost $J = 10100$. This example shows that the homotopy method can obtain different solutions. \square

Example 2: This is a model of a synchronous machine connected to an infinite busbar (Hickin and Sinah 1980). The system is given by

$$A = \begin{bmatrix} -6.2036 & 15.054 & -9.8726 & -376.58 & 215.32 & -162.24 & 66.827 \\ 0.53 & -2.0176 & 1.4363 & 0 & 0 & 0 & 0 \\ 16.846 & 25.079 & -43.555 & 0 & 0 & 0 & 0 \\ 337.4 & -89.449 & -162.83 & 57.998 & -65.514 & 68.579 & 157.57 \\ 0 & 0 & 0 & 107.25 & -118.05 & 0 & 0 \\ 0.36992 & -0.14445 & -0.26303 & -0.64719 & 0.49947 & -0.21133 & 0 \\ 0 & 0 & 0 & 0 & 0 & 376.99 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 89.353 & 0 \\ 376.99 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0.21133 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

A model of order $n_m = 3$ is

$$A^m = \begin{bmatrix} -0.0261157 & 9.349756 & -0.0528086 \\ -9.352068 & -0.912444 & 0.506220 \\ -0.0541716 & -0.506226 & -0.198770 \end{bmatrix}$$

$$B_m = \begin{bmatrix} -2.414471 & -0.571953 \\ 14.906052 & 0.0416151 \\ -14.944459 & 0.0512237 \end{bmatrix},$$

$$C_m = \begin{bmatrix} -0.371712 & 0.00240265 & 0.0122915 \\ 2.453290 & 14.906110 & 14.944542 \end{bmatrix}$$

This model yields the cost $J = 0.673079$.

A model of order $n_m = 4$ is

$$A_m = \begin{bmatrix} -37.55440 & -0.0546940 & 0.326197 & -0.0709427 \\ 0.0561170 & -0.0261155 & 9.349755 & -0.0528084 \\ 0.324384 & -9.352067 & -0.912440 & 0.506220 \\ 0.0705453 & -0.054714 & -0.506226 & -0.198769 \end{bmatrix}$$

$$B_m = \begin{bmatrix} -2.666516 & -0.0062470 \\ -2.414464 & -0.571953 \\ 14.906036 & 0.0416151 \\ -14.944458 & 0.0512237 \end{bmatrix}, \quad C_m^1 = \begin{bmatrix} 0.280123 & -2.651769 \\ -0.371711 & 2.453283 \\ 0.00240433 & 14.906094 \\ 0.0122916 & 14.944541 \end{bmatrix}$$

This model yields the cost $J = 3.27495 \times 10^{-6}$.

Example: This is a state space model of the transfer function between a torque activator and an approximately collocated torsional rate sensor for the ACES structure (Collins *et al.* 1991), located at NASA Marshall Space Flight Center, Huntsville, Alabama. The system in this example is of size $n = 17$, $m = 1$, $l = 1$. The non-zero elements of A are

$$\begin{aligned} A(1, 1) = A(2, 2) &= -0.031978272, & A(1, 2) = -A(2, 1) &= -78.54 \\ A(1, 17) &= 0.0097138566, & A(2, 17) &= -0.0060463517 \\ A(3, 3) = A(4, 4) &= -5.152212, & A(3, 4) = -A(4, 3) &= -51.457677 \\ A(3, 17) &= -0.021760771, & A(4, 17) &= 0.0054538246 \\ A(5, 5) = A(6, 6) &= -0.1351159, & A(5, 6) = -A(6, 5) &= -15.417859 \\ A(5, 17) &= -0.02179972, & A(6, 17) &= -0.015063913 \\ A(7, 7) = A(8, 8) &= -0.42811443, & A(7, 8) = -A(8, 7) &= -14.698408 \\ A(7, 17) &= 0.01042631, & A(8, 17) &= 0.0088479697 \\ A(9, 9) = A(10, 10) &= -0.064896745, & A(9, 10) = -A(10, 9) &= -12.077045 \\ A(9, 17) &= -0.030531575, & A(10, 17) &= -0.030260987 \\ A(11, 11) = A(12, 12) &= -0.048520356, & A(11, 12) = -A(12, 11) &= -8.9654448 \\ A(11, 17) &= -0.016843335, & A(12, 17) &= -0.011449591 \\ A(13, 13) = A(14, 14) &= -0.036781718, & A(13, 14) = -A(14, 13) &= -4.9057426 \\ A(13, 17) &= -0.1248007, & A(14, 17) &= 0.0005136047 \\ A(15, 15) = A(16, 16) &= -0.025112482, & A(15, 16) = -A(16, 15) &= -3.8432892 \\ A(15, 17) &= -0.035415526, & A(16, 17) &= -0.028115589 \\ A(17, 17) &= -92.399784. \end{aligned}$$

The matrices B and C are

$$B = \begin{bmatrix} 1.8631111 \\ -1.1413786 \\ -1.2105758 \\ 0.31424169 \\ 0.013307797 \\ -0.211128913 \\ 0.19552894 \\ -0.037391511 \\ -0.01049736 \\ -0.011486242 \\ -0.029376402 \\ 0.0082391613 \\ -0.012609562 \\ -0.0022040505 \\ -0.030853234 \\ 0.0011671662 \\ 0 \end{bmatrix}, \quad C^t = \begin{bmatrix} -0.0097138566 \\ 0.0060463517 \\ 0.021760771 \\ -0.0054538246 \\ 0.02179972 \\ 0.015063913 \\ -0.01042631 \\ -0.0088479697 \\ 0.030531575 \\ 0.030260987 \\ 0.016843335 \\ 0.011449591 \\ 0.1248007 \\ -0.0005136047 \\ 0.035415526 \\ 0.028115589 \\ 184.79957 \end{bmatrix}$$

A model of order $n_m = 8$ is given by

$$A_m = \begin{bmatrix} -70.147 & 21.918 & -2.741 & -2.9917 & -0.3721 & 0.228 & 0.0246 & 0.083 \\ 54.161 & -32.186 & 4.683 & 9.2995 & -0.4958 & 0.180 & 0.0289 & 0.093 \\ 3.5118 & -4.6512 & -0.208 & -51.396 & 0.1211 & -0.013 & -0.0049 & -0.016 \\ -22.253 & 19.045 & 51.85 & -12.043 & 1.0945 & -0.639 & -0.0741 & -0.243 \\ 1.2271 & -1.1976 & -0.200 & 1.1602 & -0.1936 & 15.44 & 0.0243 & 0.0807 \\ 0.5249 & -0.5415 & -0.076 & 0.6934 & -15.450 & -0.014 & -0.0125 & -0.041 \\ -0.0705 & 0.0708 & 0.011 & -0.0770 & 0.0238 & 0.012 & 0.0181 & -78.57 \\ -0.2393 & 0.2397 & 0.036 & -0.2610 & 0.0803 & 0.042 & 78.508 & -0.082 \end{bmatrix}$$

$$B_m = \begin{bmatrix} -0.05753 \\ -0.06445 \\ 0.01043 \\ 0.16983 \\ -0.05959 \\ 0.02622 \\ 0.04591 \\ 0.15167 \end{bmatrix}, \quad C_m^t = \begin{bmatrix} -0.16432 \\ 0.16512 \\ 0.02442 \\ -0.18165 \\ 0.05966 \\ 0.02629 \\ -0.04472 \\ -0.15162 \end{bmatrix}$$

This model yields the cost $J = 3.95223 \times 10^{-5}$.

9. Conclusion

This paper has considered the use of probability-one homotopy methods to solve the optimal projection equations for the model reduction problem. Four different approaches (Drazin inverse, contragredient transformations with full Σ , with full symmetric Σ and Ω , with $\text{diag } \Sigma = \text{diag } \Omega$) have been given for solving the equations. The first approach is based on solving the optimal projection equations in their original form. The three other approaches stem from a decomposition of the pseudogramians based on a contragredient transformation and proved to be more numerically robust than the first approach. The 'best' algorithm was shown to be effective in finding the optimal reduced order models for several examples.

The number of variables associated with the first approach is of order n^2 (n is the dimension of the original model), while the number of variables for the latter approaches is of order nn_m (n_m is the dimension of the reduced order model). Future research will involve the development of homotopy algorithms with fewer variables. It appears that by using a more rudimentary form of the optimal projection equations it is possible to reduce the number of variables to be of order $n(m+l)$ (m and l are, respectively, the number of inputs and outputs). Future research will also consider the H_2/H_∞ reduced order problem (Haddad and Bernstein 1989).

REFERENCES

- AOKI, M., 1968, Control of large-scale dynamic systems by aggregation. *IEEE Transactions on Automatic Control*, **13**, 246–253.
- BARTELS, R. H., and STEWART, G. W., 1972, Solution of the matrix $AX + XB = C$. *Communications of the Association for Computing Machinery*, **15**, 820–826.
- BERNSTEIN, D. S., and HYLAND, D. C., 1985, The optimal projection equations for reduced-order state estimation. *IEEE Transactions on Automatic Control*, **30**, 583–585.
- BROGAN, W. L., 1985, *Modern Control Theory* (Englewood Cliffs, NJ: Prentice Hall).
- CAMPBELL, S. L., and MEYER, C. D. JR, 1979, *Generalized Inverses of Linear Transformations* (London: Pitman).
- COLLINS, E. G., JR, PHILLIPS, D. J., and HYLAND, D. C., 1991, Robust decentralized control laws for the ACES structure. *Control Systems Magazine*, April, 62–70.
- GLOVER, K., 1984, All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ -error bounds. *International Journal of Control*, **39**, 1115–1193.
- HADDAD, W. M., and BERNSTEIN, D. S., 1989, Combined L_2/H_∞ model reduction. *International Journal of Control*, **49**, 1523–1535.
- HEARON, J. Z., and EVANS, J. W., 1968, Differentiable generalized inverses. *Journal of Research of the National Bureau of Standards*, **72B**, 109–113.
- HICKIN, J., and SINHA, N. K., 1980, Model reduction for linear multivariable systems. *IEEE Transactions on Automatic Control*, **25**, 1121–1127.
- HYLAND, D. C., and BERNSTEIN, D. S., 1984, The optimal projection equations for fixed-order dynamic compensation. *IEEE Transactions on Automatic Control*, **29**, 1034–1037; 1985, The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton, and Moore. *Ibid.*, **30**, 1201–1211.
- KABAMBA, P. T., 1985, Balanced gains and their significance for L^2 model reduction. *IEEE Transactions on Automatic Control*, **30**, 690–693.
- KUNG, S., and LIN, D. W., 1981 a, Optimal Hankel-norm model reductions: multivariable systems. *IEEE Transactions on Automatic Control*, **26**, 832–852; 1981 b, A state-space formulation for optimal Hankel-norm approximations. *Ibid.*, **26**, 942–946.
- KWONG, C. P., 1982, Optimal chained aggregation for reduced-order modelling. *International Journal of Control*, **35**, 964–982.
- MOORE, B. C., 1981, Principle component analysis in linear systems: controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, **26**, 17–32.
- PERNEBO, L., and SILVERMAN, L. M., 1982, Model reduction via balanced space representations. *IEEE Transactions on Automatic Control*, **27**, 382–387.
- RICHTER, S., and COLLINS, E. G. JR, 1989, A homotopy algorithm for reduced order compensator design using the optimal projection equations. *Proceedings of the 28th IEEE Conference on Decision and Control*, Tampa, Florida, pp. 506–511.
- SKELTON, R. E., 1980, Cost decomposition of linear systems with application to model reduction. *International Journal of Control*, **32**, 1031–1055.
- SKELTON, R. E., and HUGHES, P. C., 1980, Modal cost analysis for linear matrix-second-order systems. *Journal of Dynamic Systems, Measurement, and Control*, **102**, 151–158.
- SKELTON, R. E., and YOUSUFF, A., 1983, Component cost analysis of large scale systems. *International Journal of Control*, **37**, 285–304.

- SMALE, S., 1976, A convergent process of price adjustment and global Newton methods. *Journal of Mathematical Economics*, **3**, 107–120.
- SPANOS, J. T., MILMAN, M. H., and MINGORI, D. L., 1990, Optimal model reduction and frequency-weighted extension. AIAA paper number AIAA-90-3345-CP, pp. 271–284.
- DE VILLEMAGNE, C., and SKELTON, R. E., 1987, Model reductions using a projection formulation. *International Journal of Control*, **46**, 2141–2169.
- WATSON, L. T., 1986, Numerical linear algebra aspects of globally convergent homotopy methods. *SIAM Review*, **28**, 529–545; 1989, Globally convergent homotopy methods: A tutorial. *Applied Mathematics and Computation*, **31BK**, 369–396; 1990, Globally convergent homotopy algorithms for nonlinear systems of equations. *Nonlinear Dynamics*, **1**, 143–191.
- WATSON, L. T., BILLUPS, S. C., and MORGAN, A. P., 1987, HOMPACk: A suite of codes for globally convergent homotopy algorithms. *ACM Transactions on Mathematical Software*, **13**, 281–310.
- WILSON, D. A., 1970, Optimum solution of model-reduction problem. *Proceedings of the Institution of Electrical Engineers*, **117**, 1161–1165.
- ŽIGIĆ, D., 1991, Homotopy methods for solving the optimal projection equations for the reduced order model problem. M.S. thesis, Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.