Contragredient Transformations Applied to the Optimal Projection Equations

Dragan Zigic, Layne T. Watson,
and Christopher Beattie

TR 92-28

Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

May 20, 1992

Contragredient Transformations Applied to the Optimal Projection Equations

Dragan Žigić and Layne T. Watson
Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

Christopher Beattie
Department of Mathematics
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

ABSTRACT.

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. It is not obvious from their original form how they can be differentiated and how some algorithm for solving nonlinear equations can be applied to them. A contragredient transformation, a transformation which simultaneously diagonalizes two symmetric positive semidefinite matrices, is used to transform the equations into forms suitable for algorithms for solving nonlinear problems. Three different forms of the equations obtained using contragredient transformations are given. An SVD-based algorithm for the contragredient transformation and a homotopy algorithm for the transformed equations are given, together with a numerical example.

---

1. INTRODUCTION.

In [8] Hyland and Bernstein 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. It is shown here how these equations (known as the optimal projection equations) can be transformed into forms suitable for algorithms for solving nonlinear problems. The crucial step is a contragredient transformation, a transformation which simultaneously diagonalizes two symmetric positive semidefinite matrices $\hat{Q}$, $\hat{P}$ satisfying rank $(\hat{Q})$ = rank $(\hat{P})$ = rank $(\hat{Q}\,\hat{P})$.

Some other applications of the optimal projection approach include the H_2/H_∞ model reduction problem [5], the fixed order dynamic compensation problem [7] and the reduced order state estimation problem [1].

The complete statement of the reduced order model problem is given in Section 2. Section 3 gives some theoretical background on contragredient transformations and their relationship to (G, M, Γ)-factorizations. An SVD-based algorithm for the numerical computation of these decompositions is also derived. Section 4 gives three possible ways to transform the optimal projection equations using contragredient transformations into a computationally useful form. Section 5 describes a numerical homotopy algorithm based on contragredient transformations, and Section 6 summarizes.

2. STATEMENT OF THE PROBLEM.

Given the controllable and observable, time invariant, continuous time system

$$\dot{x}(t) = A\,x(t) + B\,u(t),$$
$$y(t) = C\,x(t),$$

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

$$J(A_m, B_m, C_m) \equiv \lim_{t\to\infty} E\left[(y - y_m)^t\,R\,(y - y_m)\right],$$

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

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

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

$$\hat{Q}\,\hat{P} = G^t\,M\,\Gamma,$$
$$\Gamma\,G^t = I_{n_m}.$$

Positive semisimple means similar to a symmetric positive definite matrix.

The following theorem from [8] gives necessary conditions for the optimal solution to the reduced order model problem.

THEOREM 2. *Suppose $(A_m, B_m, C_m) \in A_+$ solves the optimal model-reduction problem. Then there exist symmetric positive semidefinite matrices $\hat{Q}$, $\hat{P} \in R^{n\times n}$ such that for some (G, M, Γ)-factorization of $\hat{Q}\hat{P}$, A_m, B_m and C_m are given by*

$$A_m = \Gamma\,A\,G^t,$$
$$B_m = \Gamma\,B,$$
$$C_m = C\,G^t,$$

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

$$0 = \tau[A\,\hat{Q} + \hat{Q}\,A^t + B\,V\,B^t], \tag{1}$$
$$0 = [A^t\,\hat{P} + \hat{P}\,A + C^t\,R\,C]\,\tau, \tag{2}$$
$$\text{rank }(\hat{Q}) = \text{rank }(\hat{P}) = \text{rank }(\hat{Q}\,\hat{P}) = n_m. \tag{3}$$

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

$$A\,G_c + G_c\,A^t + B\,V\,B^t = 0,$$
$$A^t\,G_o + G_o\,A + C^t\,R\,C = 0.$$

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

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

where $(\hat{Q}\,\hat{P})^{\sharp}$ is the Drazin inverse [2]. Observe that this implies τ is uniquely defined by $\hat{Q}$ and $\hat{P}$.

Note that τ makes equations (1)–(2) highly nonlinear implicit functions of $\hat{Q}$ and $\hat{P}$, and it is not clear how to differentiate τ (the factorizations defining it are not unique). Even supposing that τ could be differentiated and a Newton, quasi-Newton, or homotopy algorithm applied directly to (1)–(2), it is unclear how to enforce (3). $\hat{Q}$ and $\hat{P}$ could be projected to achieve the correct rank, but that won't make $\hat{Q}\,\hat{P}$ have the correct rank.

3. CONTRAGREDIENT TRANSFORMATION BACKGROUND.

The following theorem from [8], which is a special case of a result in [3], gives a sufficient condition for simultaneous reduction of two symmetric positive semidefinite matrices to a diagonal form using a contragredient transformation. The proof given here, which differs from that in [8], is constructive and provides an outline for the numerical computation of the contragredient transformation. The core of the construction is similar to ideas developed for the full rank case in [9].

THEOREM 3. [8, Proposition 2.3] *Let symmetric positive semidefinite Q, $P \in R^{n\times n}$ satisfy*

$$\text{rank }(Q) = \text{rank }(P) = \text{rank }(Q\,P) = n_m, \tag{4}$$

where $n_m \le n$. Then, there exists a nonsingular $W \in R^{n\times n}$ (contragredient transformation) and positive definite diagonal $\Sigma \in R^{n_m\times n_m}$ such that

$$Q = W\begin{pmatrix}\Sigma & 0\\ 0 & 0\end{pmatrix}W^t, \quad P = W^{-t}\begin{pmatrix}\Sigma & 0\\ 0 & 0\end{pmatrix}W^{-1}. \tag{5}$$

Proof. Since Q and P are symmetric positive semidefinite and have rank n_m, there exist X, $Y \in R^{n\times n_m}$ with full column rank such that $Q = XX^t$ and $P = YY^t$. Let $\bar{X}$, $\bar{Y} \in R^{n\times(n-n_m)}$ have columns that span ker (Q) and ker (P), respectively. Since $QP = X(X^tY)Y^t$, (4) implies that X^tY is nonsingular. Likewise the matrices $\left(X\,\bar{Y}\right)$ and $\left(Y\,\bar{X}\right)$ are nonsingular. To see this, suppose that $\left(Y\,\bar{X}\right)\begin{pmatrix}a\\ b\end{pmatrix} = 0$. Thus $Ya = -\bar{X}b$ and premultiplication by X^t yields $X^tYa = 0$, implying $a = 0$ since X^tY is invertible. Then $b = 0$ also since $\bar{X}$ has full column rank. Thus $\left(Y\,\bar{X}\right)$ is nonsingular and the nonsingularity of $\left(X\,\bar{Y}\right)$ follows similarly. Since

$$\left(X\,\bar{Y}\right)^t\left(Y\,\bar{X}\right) = \begin{pmatrix}X^tY & 0\\ 0 & \bar{Y}^t\bar{X}\end{pmatrix},$$

clearly $\bar{Y}^t\bar{X}$ is nonsingular. Define the two singular value decompositions $X^tY = U\Sigma V^t$ and $\bar{X}^t\bar{Y} = \bar{U}\bar{\Sigma}\bar{V}^t$, and let

$$W = (X\,\bar{Y})\begin{pmatrix} U & 0 \\ 0 & \bar{V} \end{pmatrix}\begin{pmatrix} \Sigma & 0 \\ 0 & \bar{\Sigma} \end{pmatrix}^{-1/2}.$$

Since W is the product of nonsingular matrices it must be nonsingular. Straightforward calculations verify that

$$W^{-1} = \begin{pmatrix} \Sigma & 0 \\ 0 & \bar{\Sigma} \end{pmatrix}^{-1/2}\begin{pmatrix} V^t & 0 \\ 0 & \bar{U}^t \end{pmatrix}\begin{pmatrix} Y^t \\ \bar{X}^t \end{pmatrix},$$

and that W is a contragredient transformation simultaneously diagonalizing Q and P. ∎

The following lemma defines the construction of the projective factorization used in the optimal projection approach for solving the reduced order model problem and relates it to the contragredient transformation of Theorem 3.

LEMMA 4. [8, Lemma 2.1] *Let symmetric positive semidefinite $\hat{Q}$, $\hat{P} \in R^{n\times n}$ satisfy the rank conditions* (4). *Then there exists a (G, M, Γ)-factorization of $\hat{Q}\,\hat{P}$, i.e., there exist $G, \Gamma \in R^{n_m\times n}$ and positive semisimple $M \in R^{n_m\times n_m}$ such that*

$$\hat{Q}\,\hat{P} = G^t\,M\,\Gamma, \tag{6}$$

$$\Gamma\,G^t = I_{n_m}. \tag{7}$$

Proof. Due to Theorem 3 there exist nonsingular $W \in R^{n\times n}$ and positive definite diagonal $\Sigma \in R^{n_m\times n_m}$ such that

$$\hat{Q} = W\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}W^t, \quad \hat{P} = W^{-t}\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}W^{-1}. \tag{8}$$

The equations (8) can be expressed in the equivalent form

$$\hat{Q} = W_1\,\Sigma\,W_1^t, \quad \hat{P} = U_1^t\,\Sigma\,U_1, \tag{9}$$

where

$$W = (\overbrace{W_1}^{n_m} \quad W_2), \quad W^{-1} = U = \begin{matrix} n_m\{ \\ \end{matrix}\begin{pmatrix} U_1 \\ U_2 \end{pmatrix}. \tag{10}$$

From (9)–(10) with $G \equiv W_1^t$, $M \equiv \Sigma^2$ and $\Gamma \equiv U_1$ follow (6) and (7). ∎

The algorithm for computing the (G, M, Γ)-factorization of $\hat{Q}\,\hat{P}$ becomes:

1) Form Cholesky-type factorizations of $\hat{Q}$ and $\hat{P}$ with symmetric pivoting:

$$\hat{Q} = \Pi_Q L_Q L_Q^t \Pi_Q^t = XX^t,$$
$$\hat{P} = \Pi_P L_P L_P^t \Pi_P^t = YY^t,$$

where L_P, $L_Q \in R^{n\times n_m}$ and Π_P, $\Pi_Q \in R^{n\times n}$ are permutation matrices. (See [4] §4.2.9, for example.)

2) Form the singular value decomposition of

$$X^tY = (\Pi_Q L_Q)^t(\Pi_P L_P) = U\Sigma V^t.$$

(Consider avoiding the explicit formation of the product as in [6].)

3) Assign

$$W_1 = \Pi_Q L_Q U \Sigma^{-1/2},$$
$$U_1 = \Sigma^{-1/2} V^t L_P^t \Pi_P^t.$$

4. EQUIVALENT FORMS OF THE OPTIMAL PROJECTION EQUATIONS.

Three different ways of applying the contragredient transformation to obtain simpler forms of the optimal projection equations (1)–(3) will now be given. Homotopy methods based on these forms are given in [10].

4.1. FIRST FORM OF THE EQUATIONS.

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

The equations (1)–(2) can be considered in another, equivalent form. If (1) is multiplied by U_1 from the left, and (2) is multiplied by W_1 from the right, using the contragredient transformation

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

the following two equations are obtained:

$$U_1 A W_1 \Sigma W_1^t + \Sigma W_1^t A^t + U_1 B V B^t = 0, \tag{11}$$
$$A^t U_1^t \Sigma + U_1^t \Sigma U_1 A W_1 + C^t R C W_1 = 0. \tag{12}$$

The third equation

$$U_1 W_1 - I = 0 \tag{13}$$

determines the relationship between W_1 and U_1.

The matrix equations (11)–(13) contain $2\,n\,n_m + n_m^2$ scalar equations. On the other hand, the only natural unknowns in (11)–(13), W_1, U_1 and diagonal Σ, contain $2\,n\,n_m + n_m$ variables. Hence, something else is necessary to match 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 (11)–(13) 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.

4.2. SECOND FORM OF THE EQUATIONS.

Another approach in transforming (1)–(2) is to consider the decomposition

$$\hat{Q} = W_1 \Sigma W_1^t, \quad \hat{P} = U_1^t \Omega U_1,$$

which leads to the equations

$$U_1 A W_1 \Sigma W_1^t + \Sigma W_1^t A^t + U_1 B V B^t = 0, \tag{14}$$
$$A^t U_1^t \Omega + U_1^t \Omega U_1 A W_1 + C^t R C W_1 = 0, \tag{15}$$
$$U_1 W_1 - I = 0, \tag{16}$$

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

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

4.3. THIRD FORM OF THE EQUATIONS.

Another way to design a method using equations (14)–(16) is to reduce the number of unknowns. The number of unknowns can be reduced to $2\,n\,n_m + n_m^2$ if the diagonal elements of Ω are taken to be the diagonal elements of Σ.

5. CONTRAGREDIENT BASED HOMOTOPY ALGORITHM.

A homotopy algorithm for the nonlinear equations (11)–(13) can be developed as follows. Replace A in (11) and (12) by $A(\lambda) = (1-\lambda)D + \lambda A$, where D is chosen so that the problem (D, B, C) is easily solved, or at least a good approximation to a solution is easily obtained. The choice of D is a story by itself that is not central to the development here. Suffice it to say that good algorithms for choosing D exist [10]. One could also replace B and C by $(1-\lambda)B_0 + \lambda B$ and $(1-\lambda)C_0 + \lambda C$, but it turns out that is not really necessary. The homotopy map (initially) is

$$\rho_a(\lambda, x) = F(a, \lambda, x) = \begin{pmatrix} U_1\,A(\lambda)\,W_1\Sigma\,W_1^t + \Sigma\,W_1^t\,A^t(\lambda) + U_1\,B\,V\,B^t \\ A^t(\lambda)\,U_1^t\,\Sigma + U_1^t\,\Sigma\,U_1\,A(\lambda)\,W_1 + C^t\,R\,C\,W_1 \\ U_1\,W_1 - I \end{pmatrix} = 0, \tag{17}$$

where x consists of the matrices U_1, W_1, and Σ (assuming the first form of the equations), $0 \le \lambda \le 1$, and a is a parameter vector involved in the definition of D. a is crucial for theoretical discussions of the global convergence properties of homotopy methods, but can be viewed as a constant and ignored for the linear algebra discussion here.

In general terms, the homotopy algorithm is to track the solutions of $\rho_a(\lambda, x) = 0$ as λ varies (not necessarily monotonically!) from 0 to 1, going from the known solution x_0 at $\lambda = 0$ to a solution of the original problem (11)–(13) at $\lambda = 1$. Under reasonably general hypotheses on ρ_a, the solution set of $\rho_a = 0$ contains a smooth 1-manifold γ emanating from $(0, x_0)$ and guaranteed to reach a solution $\bar{x}$ at $\lambda = 1$. γ can be tracked by robust and sophisticated numerical algorithms (see [10] and the references therein for more details).

Recall that $\hat{Q}$ and $\hat{P}$ are symmetric, and thus, mathematically from (9), Σ is also symmetric. However, depending on the choice of D and the initial point $x_0 = \big((U_1)_0, (W_1)_0, \Sigma_0\big)$, and because of roundoff errors and mathematical approximations along γ, the computed points $(\lambda, x) = (\lambda, U_1, W_1, \Sigma)$ along γ will almost certainly not have Σ being symmetric. Technically, this doesn't matter because Σ will become symmetric at $\lambda = 1$ (and computational experience verifies this). However, 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. Precisely, at each point along the solution locus the homotopy map has the form

$$\rho_a(\lambda, x) = F(a, \lambda, x) - (1-\lambda)F(a, 0, x_0), \tag{18}$$

but x_0 keeps adapting to preserve the symmetry of Σ.

In summary, the whole algorithm is:

1) Define D as in [10].
2) Choose a starting point (Q_0, P_0) using one of the strategies explained in [10]. Compute $(W_1)_0$, $(U_1)_0$ and Σ_0 using the algorithm of Section 3.
3) Set $\lambda := 0$, $x := x_0 = ((W_1)_0, (U_1)_0, \Sigma_0)$.

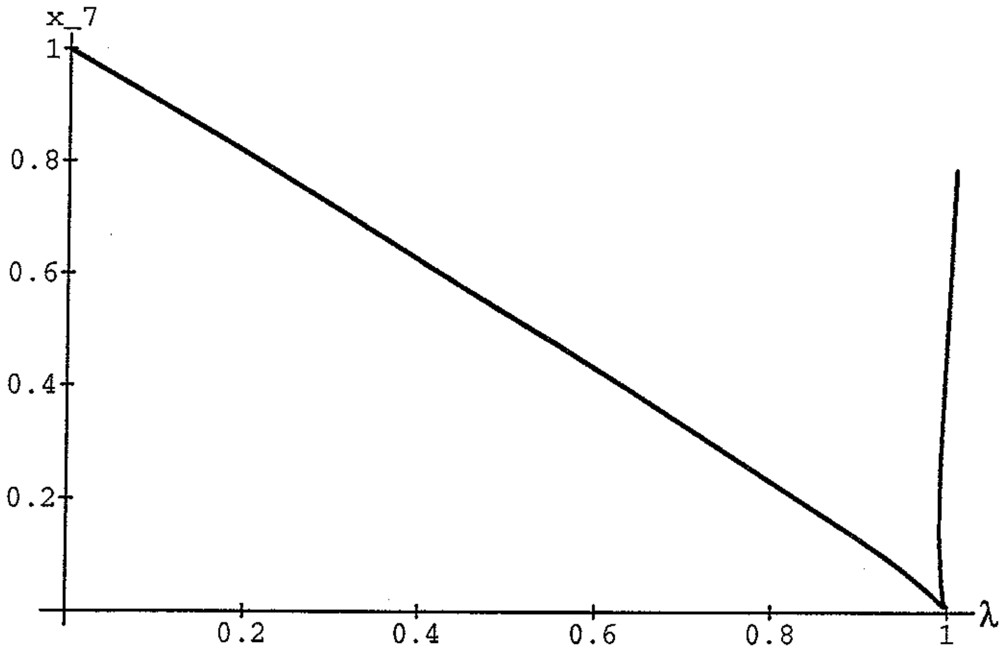

FIGURE 1. *Trace of* Σ_{11}.

4) Evaluate $\rho_a(\lambda, x)$ given by (18).
5) Evaluate the Jacobian matrix $D\rho_a(\lambda, x)$.
6) Take a step along the curve and obtain $x_1 = (W_1, U_1, \Sigma)$, $\bar{\lambda}$.
7) Compute $\bar{x}_1 = (W_1, U_1, \bar{\Sigma}) = (W_1, U_1, (\Sigma + \Sigma^t)/2)$.
8) Change the homotopy $\rho_a(\lambda, x)$ to

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

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

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

5.1. AN EXAMPLE.

Consider the system (System 5 in [10]) defined by

$$A = \begin{pmatrix} -10 & 1 & 0 \\ -5 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad C = (\,1 \quad 0 \quad 0\,).$$

With $V = R = I$ and the initial system $D = -10I$, 21 steps (Jacobian matrix evaluations) were required to find the model of order $n_m = 1$:

$$A_m = (\,-0.157898\,), \quad B_m = (\,0.423088\,), \quad C_m = (\,0.423088\,).$$

This model yields the cost $J = 0.0107792$. Figure 1 shows the trace of Σ_{11}, which is typical of homotopy zero curves γ (note the sharp turn where most of the curve tracking effort was spent).

6. SUMMARY.

This note gives an application of contragredient transformations to the optimal projection equations for solving the H_2 model reduction problem. In their original form the equations are very hard to deal with, since there is no clear way to differentiate them or enforce the rank conditions. When transformed using contragredient transformations the equations become quadratic, and while still very challenging nonlinear equations, at least amenable to quasi-Newton or homotopy algorithms for nonlinear systems. Since the exact match between the number of equations and the number of unknowns can be made in many different ways, the equations can be presented in different equivalent forms, three of which were described here. Computational experience with homotopy algorithms based on contragredient transformations reported in [10] suggests they are robust and practical for small problems ($2\,n\,n_m + n_m^2 < 1000$). Finding practical algorithms for solving the H_2 reduced order model problem for large n and n_m remains an open problem. Finally, it must be pointed out that quasi-Newton or other locally convergent techniques are simply inadequate—homotopy algorithms are the only known reliable way to solve these control problems.

REFERENCES

[1] D. S. Bernstein and D. C. Hyland, The optimal projection equations for reduced-order state estimation, *IEEE Trans. Autom. Contr.* 30:583–585 (1985).

[2] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations,* Pitman, London, England, 1979.

[3] K. Glover, All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds, *Int. J. Contr.* 39:1115–1193 (1984).

[4] G. H. Golub and C. F. Van Loan, *Matrix Computations,* Johns Hopkins University Press, Baltimore, MD, 1989.

[5] W. M. Haddad and D. S. Bernstein, Combined L_2/H_∞ model reduction, *Int. J. Contr.* 49:1523–1535 (1989).

[6] M. T. Heath, A. J. Laub, C. C. Paige, and R. C. Ward, Computing the singular decomposition of a product of two matrices, *SIAM J. Sci. Stat. Comp.* 7:1147–1159 (1986)

[7] D. C. Hyland and D. S. Bernstein, The optimal projection equations for fixed-order dynamic compensation, *IEEE Trans. Autom. Contr.* 29:1034–1037 (1984).

[8] D. C. Hyland and D. S. Bernstein, The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton, and Moore, *IEEE Trans. Autom. Contr.* 30:1201–1211 (1985).

[9] A. J. Laub, M. T. Heath, C. C. Paige, and R. C. Ward, Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms, *IEEE Trans. Autom. Contr.* 32:115–122 (1987).

[10] D. Žigić, L. T. Watson, E. G. Collins, Jr. and D. S. Bernstein, Homotopy methods for solving the optimal projection equations for the H_2 reduced order model problem, *Int. J. Contr.* to appear.