Universidad Central de Venezuela
Facultad de Ciencias
Escuela de Computación

Lecturas en Ciencias de la Computación
ISSN 1316-6239

A secant method for the matrix sign function

Marlliny Monsalve

RT 2009-03

Centro de Cálculo Científico y Tecnológico (CCCT)
Caracas, Febrero, 2009

A secant method for the matrix sign function

Marlliny Monsalve [*]

February 04, 2009

Abstract

A secant method has been recently proposed for solving nonlinear matrix problems with interesting features, including low computational cost and q-superlinear convergence for special cases. In this work, we analyze this secant method for the special problem of computing the sign of a given matrix. The global and q-superlinear convergence of the proposed secant method are proved from specialized initial guesses, and the numerical stability is also established. We complement our analysis with several numerical experiments that show the advantages of using the secant method over the well-known variant of Newton's method for the same problem.

Keywords: Secant method, quasi-Newton methods, matrix sign function.

1 Introduction

Let $A \in \mathbb{C}^{n\times n}$ be a matrix having no pure imaginary eigenvalues, and let

$$A = P \begin{bmatrix} J_- & 0 \\ 0 & J_+ \end{bmatrix} P^{-1}, \tag{1}$$

be the Jordan decomposition of A, where P is nonsingular. The matrices $J_- \in \mathbb{C}^{q\times q}$, $J_+ \in \mathbb{C}^{r\times r}$ with $q + r = n$, are such that their eigenvalues lie in the open left half-plane and in the open right half-plane, respectively. The sign of A is defined as

$$S = sign(A) = P \begin{bmatrix} -I_q & 0 \\ 0 & I_r \end{bmatrix} P^{-1}, \tag{2}$$

where I_q and I_r are identity matrices of dimension q and r, respectively. If A has any pure imaginary eigenvalues, then $sign(A)$ is not defined. The sign function is a useful tool to solve the Lyapunov and the Riccati equation which arise in problems related to control theory [2, 6, 8, 10, 11, 14]. The sign function can also be used to solve some eigenvalues problems [5] and for computing invariant subspaces [1, 3, 9]. The matrix $S = sign(A)$

[*]Departamento de Computación, Facultad de Ciencias, Universidad Central de Venezuela, Ap. 47002, Caracas 1041-A, Venezuela (marlliny.monsalve@ciens.ucv.ve). Supported by CDCH-UCV project 03.00.6640.2007/2.

commutes with A and satisfies that $S^2 = I$, see [7]. Since $S^2 = I$, the matrix S can be computed solving the nonlinear matrix equation $F(X) = 0$, where

$$F(X) = X^2 - I. \tag{3}$$

The Newton method applied to solve (3) is given by

$$X_{k+1} = \frac{1}{2}(X_k + X_k^{-1}), \quad \text{for } k = 0, 2, \cdots \tag{4}$$

If A has no eigenvalues on the imaginary axis and we used $X_0 = A$ then, the iteration (4) is stable and quadratically convergent to $sign(A)$, see [7, 13, 14] for details. The iteration (4) was proposed by Robert in [14] to solve Riccati equations. Since the secant method can be another option to solve a nonlinear equation, in our next section we analyze a recently proposed secant method for solving (3). In Section 3, we prove the new proposed secant method converges globally and q-superlinearly to the sign of a given matrix, from specialized initial guesses. In Section 4, we establish the numerical stability of the new secant scheme. Numerical examples are present in Section 5. Finally, we give some concluding remarks in Section 6.

2 Specialized secant method

Recently, a general secant method and its inverse version have been proposed in [12] to solve $F(X) = 0$, where F is a nonlinear and Fréchet differentiable matrix function. This secant method is based on the following iteration

$$X_{k+1} = X_k - A_k^{-1} F(X_k), \tag{5}$$

where $X_{-1} \in \mathbb{C}^{n\times n}$ and $X_0 \in \mathbb{C}^{n\times n}$ are given. Additionally, A_{k+1} is a suitable linear operator that satisfies

$$A_{k+1} S_k = Y_k, \tag{6}$$

with $S_k = X_{k+1} - X_k$ and $Y_k = F(X_{k+1}) - F(X_k)$. Equation (6) is known as the *secant equation.*

Notice that an $n \times n$ matrix is enough to satisfy the matrix secant equation (6). Hence, the operator A_k can be forced to be a matrix of the same dimension of the step S_k and the map-difference Y_k. Therefore, once X_{k+1} is obtained, A_{k+1} can be computed at each iteration by solving a linear system of n^2 equations. This attractive feature is in sharp contrast with the standard extension of quasi-Newton methods for general Hilbert spaces, (see e.g. [4, 15]). For instance, a quasi-Newton method in the space of matrices would involve an $n^2 \times n^2$ matrix to approximate the derivative of F at X_k. For problem (3), the direct secant method can be written as follows

Algorithm 1 Specialized secant method

Given $X_{-1} \in \mathbb{C}^{n\times n}$, $X_0 \in \mathbb{C}^{n\times n}$
Set $S_{-1} = X_0 - X_{-1}$; **Set** $Y_{-1} = X_0^2 - X_{-1}^2$
Solve $A_0 S_{-1} = Y_{-1}$ $\triangleright$ for A_0
for $k = 0, 1, \cdots$ until convergence **do**
  Solve $A_k S_k = -F(X_k)$ $\triangleright$ for S_k
  Set $X_{k+1} = X_k + S_k$
  Set $Y_k = X_{k+1}^2 - X_k^2$
  Solve $A_{k+1} S_k = Y_k$ $\triangleright$ for A_{k+1}
end for

From Algorithm 1 we obtain

$$X_{k+1} = X_k - (X_k - X_{k-1})(X_k^2 - X_{k-1}^2)^{-1}(X_k^2 - I). \tag{7}$$

Note that any initial guesses commute with identity matrix. Therefore, it is easy to show by induction that $X_i X_j = X_j X_i$, for all i, j, where X_i and X_j are generated by (7). Using this fact, (7) can be written as

$$\begin{aligned} X_{k+1} &= X_k - (X_k - X_{k-1})(X_k - X_{k-1})^{-1}(X_k + X_{k-1})^{-1}(X_k^2 - I) \\ &= (X_k + X_{k-1})^{-1}((X_k + X_{k-1})X_k - (X_k^2 - I)) \\ &= (X_k + X_{k-1})^{-1}(X_{k-1}X_k + I). \end{aligned} \tag{8}$$

Based on equation (8), we propose the following secant method for computing the sign matrix

Algorithm 2 Secant method for computing the matrix sign

Given $X_{-1} \in \mathbb{C}^{n\times n}$, $X_0 \in \mathbb{C}^{n\times n}$
for $k = 0, 1, \cdots$ until convergence **do**
  Solve $[X_k + X_{k-1}]X_{k+1} = X_{k-1}X_k + I$ $\triangleright$ for X_{k+1}
end for

In the next section, we study convergence and stability of Algorithm 2.

3 Convergence

For our analysis we assume that A is diagonalizable, that is, there exists a nonsingular matrix V such that

$$V^{-1}AV = \Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_n), \tag{9}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of A. From (2), it is easy to establish that $sign(V^{-1}AV) = V^{-1}sign(A)V$, for any nonsingular matrix V. Therefore, using this property and from equation (9), we obtain that

$$sign(\Lambda) = sign(V^{-1}AV) = V^{-1}sign(A)V, \tag{10}$$

for details see [7]. On the other hand, if we define $D_k = V^{-1}X_kV$ then we have from equation (8) that

$$D_{k+1} = (D_k + D_{k-1})^{-1}[D_{k-1}D_k + I]. \tag{11}$$

Notice that if D_{-1} and D_0 are diagonal matrices then all successive D_k are diagonal too. From (11), it is enough to prove that $\{D_k\}$ converges to the sign of Λ, to ensure the convergence of the sequence generated by Algorithm 2 to $sign(A)$. We are now ready to establish our convergence result.

Theorem 3.1 *Let $A \in \mathbb{C}^{n\times n}$ be a diagonalizable matrix, that is, there exists a nonsingular matrix V such that*

$$V^{-1}AV = \Lambda = diag(\lambda_1, \lambda_2 \cdots, \lambda_n),$$

where $\lambda_1, \lambda_2 \cdots, \lambda_n$ are the eigenvalues of A. Let us assume that A has no pure imaginary eigenvalues, and that all iterates X_k are well defined. If $X_{-1} = \alpha A$; $\alpha > 0$ and $X_0 = \beta A$; $\beta > 0$ then the sequence $\{X_k\}$, generated by Algorithm 2, converges q-superlinearly to $sign(A)$.

Proof. The matrices X_{-1} and X_0 are defined such that $D_{-1} = V^{-1}X_{-1}V$ and $D_0 = V^{-1}X_0V$ are diagonal matrices, then all successive D_k are diagonal too, therefore we can write (11) as n uncoupled scalar secant iterations to solve $g(x) = 0$, with $g(x) = x^2 - 1$, given by

$$d^i_{k+1} = \frac{d^i_k d^i_{k-1} + 1}{d^i_{k-1} + d^i_k}, \tag{12}$$

where $d^i_k = (D_k)_{ii}$ and $1 \le i \le n$. On the other hand, $sign(D_k) = sign(\Lambda)$ for all $k \ge -1$. From (11) and (12), it is enough to study the convergence of $\{d^i_k\}$ to the sign of λ_i, for all $1 \le i \le n$. From (12) we have that

$$\begin{aligned} d^i_{k+1} &= \frac{(d^i_k \pm 1)(d^i_{k-1} \pm 1) \mp (d^i_{k-1} + d^i_k)}{d^i_{k-1} + d^i_k} \\ d^i_{k+1} \pm 1 &= \frac{(d^i_k \pm 1)(d^i_{k-1} \pm 1)}{d^i_{k-1} + d^i_k}. \end{aligned} \tag{13}$$

Since the eigenvalues of A are not pure imaginary, we have that $sign(\lambda_i) = s_i = \pm 1$. Let us choose i such that $\lambda_i > 0$, i.e., $sign(\lambda_i) = 1$. Now, from (13) we obtain that

$$\frac{d^i_{k+1} - 1}{d^i_{k+1} + 1} = \left(\frac{d^i_k - 1}{d^i_k + 1}\right)\left(\frac{d^i_{k-1} - 1}{d^i_{k-1} + 1}\right). \tag{14}$$

Then, applying (14) recursively, it follows that

$$\frac{d^i_{k+1}-1}{d^i_{k+1}+1} = \left(\frac{d^i_0-1}{d^i_0+1}\right)^{f_k} \left(\frac{d^i_{-1}-1}{d^i_{-1}+1}\right)^{f_{k-1}}, \tag{15}$$

where $f_{k+1} = f_k + f_{k-1}$, for $k \geq 0$, and $f_{-1} = f_0 = 1$. Notice that $\{f_k\}$ is a Fibonacci sequence which appears quite frequently in secant-methods analysis. On the other hand, since $d^i_{-1} = \alpha\lambda_i > 0$ and $d^i_0 = \beta\lambda_i > 0$, then for each i

$$\frac{d^i_{-1}-1}{d^i_{-1}+1} < 1 \text{ and } \frac{d^i_0-1}{d^i_0+1} < 1.$$

From (15) and using that $sign(\lambda_i) = 1$, we obtain that $\lim_{k\to\infty} d^i_k = 1 = sign(\lambda_i)$. When the value of i is such that $sign(\lambda_i) = -1$, we proceed in a similar way to prove that $\lim_{k\to\infty} d^i_k = -1 = sign(\lambda_i)$. Therefore,

$$\lim_{k\to\infty} d^i_k = s_i = sign(\lambda_i)$$

and

$$\lim_{k\to\infty} D_k = sign(\Lambda). \tag{16}$$

Recalling $D_k = V^{-1}X_kV$ and using (16) we have that

$$\lim_{k\to\infty} X_k = V(\lim_{k\to\infty} D_k)V^{-1} = V sign(\Lambda)V^{-1}.$$

Finally using (10), we have that $\lim_{k\to\infty} X_k = sign(A)$ and the convergence is established. Now, to prove the local q-superlinear convergence, we consider that (13) can be written as

$$e^i_{k+1} = c^i_k e^i_k e^i_{k-1}, \tag{17}$$

where $e^i_k = d^i_k - s_i$ and $c^i_k = 1/(d^i_{k-1} + d^i_k)$. Notice that c^i_k tends to $1/2s_i$ when k goes to infinity, and so it is bounded for k sufficiently large. From (17), we conclude that each scalar secant iteration (12) converges locally and q-superlinearly to s_i. Therefore, equivalently, there exists a sequence $\{\widetilde{c}^i_k\}$, for each $1 \leq i \leq n$, such that $\widetilde{c}^i_k > 0$ for all k, $\lim_{k\to\infty} \widetilde{c}^i_k = 0$, and

$$|e^i_{k+1}| \leq \widetilde{c}^i_k |e^i_k|. \tag{18}$$

Using (18) we now obtain in the Frobenius norm

$$\begin{aligned} \|D_{k+1} - sign(\Lambda)\|^2_F &= \sum_{i=1}^n (e^i_{k+1})^2 \leq \sum_{i=1}^n (\widetilde{c}^i_k)^2 (e^i_k)^2 \\ &\leq n\widehat{c}^2_k \sum_{i=1}^n (e^i_k)^2 \leq n\widehat{c}^2_k \|D_k - sign(\Lambda)\|^2_F, \end{aligned} \tag{19}$$

where $\widehat{c}_k = \max_{1\le i\le n}\{\widetilde{c}_k^i\}$. Finally, we have that

$$\begin{aligned}
\|X_{k+1} - sign(A)\|_F &= \|VV^{-1}(X_{k+1} - sign(A))VV^{-1}\|_F \\
&= \|V(D_{k+1} - sign(\Lambda))V^{-1}\|_F \\
&\le \kappa_F(V)\|D_{k+1} - sign(\Lambda)\|_F \\
&\le \kappa_F(V)\sqrt{n}\widehat{c}_k\|D_k - sign(\Lambda)\|_F \\
&= \kappa_F(V)\sqrt{n}\widehat{c}_k\|V^{-1}V(D_k - sign(\Lambda))V^{-1}V\|_F \\
&\le \kappa_F(V)^2\sqrt{n}\widehat{c}_k\|X_k - sign(A)\|_F,
\end{aligned}$$

where $\kappa_F(V)$ is the Frobenius condition number of V. Hence, the sequence $\{X_k\}$ converges globally and q-superlinearly to $sign(A)$. □

4 Stability

We now discuss the stability of our specialized secant method for the matrix sign function in a neighborhood of solution of equation (3). We will analyze how a small perturbation at a given iteration is propagated over the forthcoming iterations. In our analysis we assume exact arithmetic, and we also assume that the second order terms of perturbations are insignificant, so they are not taken into account.

Let Δ_k be the numerical perturbation introduced at the k-th iteration of Algorithm 2, and let be

$$\widehat{X}_k = X_k + \Delta_k.$$

From (8), we have that

$$\begin{aligned}
\widehat{X}_{k+1} &= (\widehat{X}_k + \widehat{X}_{k-1})^{-1}(\widehat{X}_{k-1}\widehat{X}_k + I) \\
&\approx (\widehat{X}_k + \widehat{X}_{k-1})^{-1}[(X_{k-1}X_k + I) + X_{k-1}\Delta_k + \Delta_{k-1}X_k]. \qquad (20)
\end{aligned}$$

On the other hand, using the well-known fact that for any nonsingular matrix B and any matrix C, $(B + C)^{-1} \approx B^{-1} - B^{-1}CB^{-1}$, up to second order terms, we obtain that

$$\begin{aligned}
(\widehat{X}_k + \widehat{X}_{k-1})^{-1} &= [(X_k + X_{k-1}) + (\Delta_k + \Delta_{k-1})]^{-1} \\
&\approx (X_k + X_{k-1})^{-1} - (X_k + X_{k-1})^{-1}(\Delta_k + \Delta_{k-1})(X_k + X_{k-1})^{-1}. \qquad (21)
\end{aligned}$$

Substituting (21) in (20) and after some algebraic manipulations, we have that

$$\begin{aligned}
\widehat{X}_{k+1} &\approx X_{k+1} + (X_k + X_{k+1})^{-1}[X_{k-1}\Delta_k + \Delta_{k-1}X_k] \\
&- (X_k + X_{k+1})^{-1}(\Delta_k + \Delta_{k-1})(X_k + X_{k+1})^{-1}(X_{k-1}X_k + I). \qquad (22)
\end{aligned}$$

Now, recalling $\Delta_k = \widehat{X}_k - X_k$, and using that for k large enough X_k is close to $S = sign(A)$, from (22) we have that

$$\Delta_{k+1} \approx (2S)^{-1}[S\Delta_k + \Delta_{k-1}S] - (2S)^{-1}(\Delta_k + \Delta_{k-1})(2S)^{-1}(S^2 + I). \qquad (23)$$

Applaying that $S^2 = I$ and $S = S^{-1}$ to equation (23) we obtain that

$$\begin{aligned}
\Delta_{k+1} &\approx \frac{1}{2}[\Delta_k + S^{-1}\Delta_{k-1}S] - \frac{1}{2}S^{-1}(\Delta_k + \Delta_{k-1})S^{-1} \\
&= \frac{1}{2}[\Delta_k - S^{-1}\Delta_k S^{-1}] + \frac{1}{2}(S^{-1}\Delta_{k-1}S - S^{-1}\Delta_{k-1}S^{-1}) \\
&= \frac{1}{2}[\Delta_k - S\Delta_k S]. \qquad (24)
\end{aligned}$$

Applaying (24) recursively, and after some algebraic manipulations, we have that

$$\Delta_{k+1} \approx \frac{1}{2}[\Delta_0 - S\Delta_0 S]. \qquad (25)$$

From (25), we can conclude that the perturbation at iteration $k+1$, Δ_{k+1}, is bounded. Therefore, the sequence $\{X_k\}$ generated by Algorithmic 2 is numerically stable.

□

5 Numerical results

All our numerical experiments were run on a Pentium IV, 3.4GHz, using Matlab 7. We compared our specialized secant method (Algorithm 2) with the Newton method (Iteration 4). We used Y_0 to denote the initial guess for Newton method and X_{-1}, X_0 were used to denote the initial guesses for secant method. We stop all considered algorithms when the Frobenius norm of the residual is less than 0.5×10^{-15}. We report the number of required iterations (Iter), and the norm of the residual ($\|F(X_k)\|_F$) when the process is stopped. We consider that an algorithm fails when the number of iterations exceeds 50 and the symbol (**) is used to indicate it. When an algorithm fails we reported the minimal residual reached. All matrices uses in the following tests were taken from Matlab gallery.

Experiment 1: We consider the orthogonal matrix `orthog` with $n = 150$ and $k = 4$. The results are reported in Table 1. We can observe in Figure 1 that secant method required 10 iterations to achieve convergence whereas the sequence generated by Newton method showed stagnation. Fortunately, Newton method produced a good approximation of the matrix sign before showing stagnation.

Method	Iter	$\|F(X_k)\|_F$
Secant	10	4.52e-15
Newton	**	6.45e-15

Table 1: Performance of Secant method and Newton method for finding the sign of $A =$ gallery('orthog',n,4) when $n = 150$, $X_{-1} = 0.5 * A$, $X_0 = 0.5 * A$ and $Y_0 = A$.

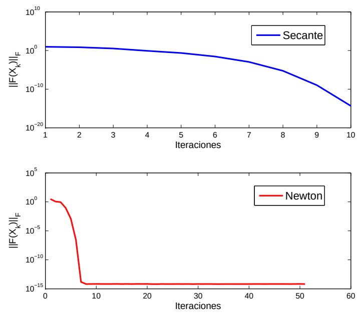

Figure 1: Semilog of the relative error of Secant method and Newton method for finding the sign of $A =$ gallery('orthog',n,4) when $n = 150$, $X_{-1} = 0.5 * A$, $X_0 = 0.5 * A$ and $Y_0 = A$.

Experiment 2: We use parter with $n = 150$. The eigenvalues of this matrix are neither real nor pure imaginary. The results are reported in Table 2. In this case secant method requires more iterations than Newton method to achieve convergence, however the secant method achieves better accuracy than Newton method.

Method	Iter	$\|F(X_k)\|_F$
Secant	16	8.96e-19
Newton	12	2.28e-15

Table 2: Performance of Secant method and Newton method for finding the sign of $A =$ gallery('parter',n) when $n = 150$, $X_{-1} = 0.5 * A$, $X_0 = 0.5 * A$ and $Y_0 = A$.

Experiment 3: For this example we use $n = 5$. The matrix A is diagonal and its diagonal entries are the imaginary number $a_{jj} = \epsilon(-1)^j + ji$, where $\epsilon > 0$. For small values of ϵ the eigenvalues of the matrix A tend to be imaginary pures, and therefore the matrix $sign(A)$ is not defined. In this experiment, we take different small values of ϵ and we report the iteration when the minimal value of $\|F(X_k)\|_F$ was reached and its value. We can see in Table 3 that for all values of ϵ the Newton method as the secant method reduce considerably the residual norm in spite of the ill-conditioned of $sign(A)$. In all cases, the secant method requires more iterations that the Newton method, this can be explained by the velocity of convergence of these methods: the Newton method converges q-quadratically whereas the secant method converges q-superlinearly.

ϵ	Method	Iter	$\|F(X_k)\|_F$
10^{-5}	Secant	37	1.32e-23
	Newton	27	5.09e-23
10^{-8}	Secant	50	3.14e-20
	Newton	37	6.39e-24
10^{-12}	Secant	70	1.38e19
	Newton	50	1.4e-19
10^{-18}	Secant	101	2.22e-16
	Newton	70	2.48e-20

Table 3: Performance of secant method and Newton method for finding the sign of A for different values of ϵ, $n = 5$, $X_{-1} = 0.5 * A$, $X_0 = 0.5 * A$ and $Y_0 = A$. The eigenvalues of A are $\lambda_j = (-1)^j \epsilon + ji$ with $j = 1, 2, \cdots, n$.

6 Concluding Remark

In this work we proposed a new method to find the sign of a given matrix. This method could be obtained adapting a recently proposed secant method for solving nonlinear matrix problems. For the new proposed algorithm we established specialized guesses that guarantee global and q-superlinear convergence. Additionally, the numerical stability of the sequence generated by the new algorithm was proved. In numerical experiment section we compared our new secant algorithm with the well-known Newton's method, and the results show that the secant method can be a suitable option for computing the sign of a given matrix.

References

[1] Z. Bai and J. Demmel. Using the matrix sign function to compute invariant subspaces, *SIAM Journal on Matrix Analysis and Applications* 19:205–225, 1998.

[2] P. Benner and E. S. Quintana-Ortí. Solving stable generalized Lyapunov equations with the matrix sign function. *Numerical Algorithms*, 20(1):75–100, 1999.

[3] R. Byers, C. He and V. Mehrmann. The Matrix Sign Function Method and the Computation of Invariant Subspaces, *SIAM Journal on Matrix Analysis and Applications* 18(3):615–632, 1997.

[4] M. A. Gomes-Ruggiero and J. M. Martínez. The column-updating method for solving nonlinear equations in Hilbert space. *RAIRO Mathematical Modelling and Numerical Analysis* 26: 309–330, 1992.

[5] J. L. Howland. The sign matrix and the separation of matrix eigenvalues. *Linear Algebra and its Appl.*, 49:221–232, 1983.

[6] C. S. Kenney and A. J. Laub. The matrix sign function. *IEEE Trans. Automat. Control*, 40(8):1330–1348, 1995.

[7] C. S. Kenney and A. J. Laub. Rational iteration methods for the matrix sign function. *SIAM J. Math. Anal. Appl.*, 12(2):273–291, 1991.

[8] C. S. Kenney, A. J. Laub and P. M. Papadopoulos. Matrix-Sign Algorithms for Riccati Equations. *IMA J. Math. Control Info.*, 9:331–344, 1992.

[9] C. S. Kenney, A. J. Laub and P. M. Papadopoulos. A Newton-squaring algorithm for computing the negative invariant subspace of a matrix. *IEEE Trans. Automat. Contr.*, 38:1284–1289, 1993.

[10] P. Lancaster and L. Rodman. *The Algebraic Riccati Equation*. Oxford University Press, Oxford, 1995.

[11] V. B. Larin and F. A. Aliev. Generalized Lyapunov equation and factorization of matrix polynomials. *Systems Control Lett.*, 21(6):485–491, 1993.

[12] M. Monsalve and M. Raydan. A secant method for nonlinear matrix problems, To appears in special volume (dedicated to Biswa Datta) in *Numerical Linear Algebra in Signals, Systems and Control. Springer-Verlag*, 2008

[13] E.S. Quintana-Ortí and X. Sun The generalized Newton iteration for the matrix sign function, *SIAM J. Sci. Comput.* 24(2):669-683, 2002.

[14] J.D. Roberts. Linear model reduction and solution of the algebraic Riccati equation by use of the sign function. *Internat. J. Control*, 32(4):677–687, 1980. First issued as report CUED/B-Control/TR13, Department of Engineering, University of Cambridge (1971).

[15] E. Sachs. Broyden's method in Hilbert space, *Math. Programming* 35:71–81, 1986.