A Symplectic Lanczos-Type Algorithm to Compute the Eigenvalues of Positive Definite Hamiltonian Matrices

Pierluigi Amodio*

Dipartimento di Matematica, Università di Bari,
Via E. Orabona 4, I-70125 Bari, Italy,
amodio@dm.uniba.it

Abstract. The Lanczos algorithm is a well known procedure to compute few eigenvalues of large symmetric matrices. We slightly modify this algorithm in order to obtain the eigenvalues of Hamiltonian matrices $H = JS$ with S symmetric and positive definite. These matrices represent a significant subclass of Hamiltonian matrices since their eigenvalues lie on the imaginary axis. An implicitly restarted procedure is also considered in order to speed-up the convergence of the algorithm.

1 Introduction

Many applications require the numerical approximation of the eigenvalues of a $2n \times 2n$ real Hamiltonian matrix $H = JS$ where

$$J = \begin{pmatrix} O & I_n \\ -I_n & 0 \end{pmatrix}$$

and S is a large and sparse symmetric matrix. As an example, we cite the solution of the continuous-time algebraic Riccati equations [9,11] of the form

$$Q + A^T X + XA - XGX = 0,$$

where A, G and Q are known $n \times n$ matrices, G and Q symmetric, and the solution X is also symmetric.

It is well known that the considered eigenvalues are symmetric with respect to the real and imaginary axes. In this paper we are particularly interested in a subclass of Hamiltonian matrices where S is symmetric and positive definite. In this case the eigenvalues lie (two by two symmetrically with respect to the origin) along the imaginary axis. These matrices arise in many application fields that deal with evolutionary problems whose solutions satisfy a certain conservation law, typically the energy of the system (see, for example, [5]). In the rest of the paper we refer to these matrices as positive definite Hamiltonian matrices [2].

All the numerical methods for the eigenvalues computation of Hamiltonian matrices use symplectic transformations in order to maintain the Hamiltonian

* Work supported by GNCS.

P.M.A. Sloot et al. (Eds.): ICCS 2003, LNCS 2658, pp. 139–148, 2003.

structure of the matrix. In fact, if P is symplectic (that is, $P^T JP = J$) and H is Hamiltonian, then also $P^{-1}HP$ is Hamiltonian. For this reason, most of the existing algorithms are symplectic modifications of more general approaches as, for instance, the symplectic QR method in [6,7,13] and the symplectic iterative methods [3,4].

In particular, the symplectic Lanczos algorithm in [4] where the directions generated by the procedure are symplectic rather than orthogonal, is strictly related to the method introduced in this paper. The idea inherited this algorithm is the following: let

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{12}^T & S_{22} \end{pmatrix}$$

with S_{11} and S_{22} symmetric, the matrix H may be represented in the form

$$H = \begin{pmatrix} S_{12}^T & S_{22} \\ -S_{11} & -S_{12} \end{pmatrix} \equiv \begin{pmatrix} V & W \\ Z & -V^T \end{pmatrix}.$$

The approach in [4], as well as those in [1,10] which define canonical forms for the Hamiltonian matrices, transform the above matrix in

$$\begin{pmatrix} V_1 & W_1 \\ & -V_1^T \end{pmatrix}$$

with V_1 triangular and V_1 and W_1 having few nonnull diagonals. This is called a Hamiltonian triangular form for the matrix H.

Conversely, based on a theoretical result in [2], for positive definite Hamiltonian matrices we derive a canonical form (W_2 tridiagonal)

$$\begin{pmatrix} & W_2 \\ -I & \end{pmatrix}$$

by means of a transformation that we prove to be symplectic. We will explain how to derive this form in Section 2. Then in Section 3, following the results presented in [4,8], we derive an implicit restarting technique that can be easily applied to the considered class of matrices.

2 Tridiagonal Canonical Form

In [2] existence results on a diagonal form for Hamiltonian matrices $H = JS$ with S positive definite is given. In the same paper, this result is also rearranged in order to obtain the following symplectic transformation in the simplest canonical form.

Theorem 1. *Given a real Hamiltonian matrix $H = JS$ with S symmetric and positive definite, a real symplectic matrix Z exists such that*

$$HZ = Z \begin{pmatrix} & D_1 \\ -D_1 & \end{pmatrix} \equiv ZJ \begin{pmatrix} D_1 & \\ & D_1 \end{pmatrix} \tag{1}$$

with $D_1 > 0$ diagonal.

Even if the proof of Theorem 1 is presented in [2], it is interesting to sketch its main steps. As a remark, we recall that:

- if A is a symmetric matrix, then all the eigenvalues are real and an orthogonal matrix U exists such that

$$AU = U\Lambda;$$

- if A is a skew-symmetric matrix, then all the eigenvalues are pure imaginary and a unitary matrix $U = [U_1\ \overline{U}_1]$ exists such that

$$AU = iU \begin{pmatrix} \Lambda_1 & \\ & -\Lambda_1 \end{pmatrix}.$$

  Moreover, from $V = \sqrt{2}[Re(U_1)\ Im(U_1)]$ one has

$$AV = VJ \begin{pmatrix} \Lambda_1 & \\ & \Lambda_1 \end{pmatrix}, \quad V^T V = I,$$

  that is, a skew-symmetric matrix is similar to a positive definite Hamiltonian matrix by means of an orthogonal matrix.

Proof. Let us start from the decomposition of the symmetric and positive definite matrix S in diagonal form, that is $S = Q\Lambda Q^T$, with Q orthogonal and Λ positive. Then

$$(JS)JQ = (JQ\Lambda Q^T)JQ = JQ(\Lambda Q^T JQ)$$

shows that JS is similar to $\Lambda Q^T JQ$ by means of the transformation matrix JQ. Now, since the matrix $\Lambda^{1/2} Q^T JQ \Lambda^{1/2}$ is skew-symmetric, it will admit the following similarity transformation

$$\Lambda^{1/2} Q^T JQ \Lambda^{1/2} = VJ \begin{pmatrix} D_1 & \\ & D_1 \end{pmatrix} V^T, \tag{2}$$

and hence

$$\Lambda Q^T JQ = \Lambda^{1/2} \left(\Lambda^{1/2} Q^T JQ \Lambda^{1/2} \right) \Lambda^{-1/2} = (\Lambda^{1/2} V) J \begin{pmatrix} D_1 & \\ & D_1 \end{pmatrix} (\Lambda^{1/2} V)^{-1}.$$

Let $U = JQ\Lambda^{1/2} V$. One has that

$$\begin{aligned} U^T JU &= -(V^T \Lambda^{1/2} Q^T J) J (JQ\Lambda^{1/2} V) = V^T \Lambda^{1/2} Q^T JQ \Lambda^{1/2} V \\ &= V^T VJ \begin{pmatrix} D_1 & \\ & D_1 \end{pmatrix} V^T V = J \begin{pmatrix} D_1 & \\ & D_1 \end{pmatrix} \end{aligned}$$

and therefore the matrix

$$Z = U \begin{pmatrix} D_1^{-1/2} & \\ & D_1^{-1/2} \end{pmatrix} = JQ\Lambda^{1/2} V \begin{pmatrix} D_1^{-1/2} & \\ & D_1^{-1/2} \end{pmatrix}$$

is symplectic and satisfies (1). □

Equation (1) allows us to easily obtain the eigenvalues of JS as $i\lambda_j$ and $-i\lambda_j$, where λ_j is a diagonal element of D_1. On the other hand, as it occurs in the case of symmetric matrices, the above theorem cannot be used to compute the eigenvalues numerically. For this reason, different approaches need to be introduced.

We now observe that $-S^{1/2}JSJS^{1/2}$ is symmetric and positive definite, then it is possible to apply to this matrix the Lanczos algorithm, that is one should compute a matrix $Z_{k+1} = [Z_k \; z_{k+1}]$ which is defined by means of $k+1$ orthogonal columns and a tridiagonal matrix T_k such that

$$-S^{1/2}JSJS^{1/2}Z_k = Z_kT_k + \beta_{k+1}z_{k+1}e_k^T,$$

where e_k is the last unit vector of $\mathbb{R}^k$. Then, from $V_{k+1} = S^{-1/2}Z_{k+1} = [V_k \; v_{k+1}]$, one has

$$-H^2V_k = V_kT_k + \beta_{k+1}v_{k+1}e_k^T \tag{3}$$

with V_k of size $2n \times k$ and such that

$$V_k^TSV_k = I_k. \tag{4}$$

If $-i\lambda_j$, for $j = 1, \ldots, n$, are the eigenvalues of H, then the eigenvalues of H^2 are $-\lambda_1^2, -\lambda_1^2, -\lambda_2^2, -\lambda_2^2, \ldots$, that is, H^2 has double eigenvalues with a subspace of size 2 associated. This means that, if no breakdown has previously occurred, this algorithm theoretically will stop after n steps, providing a tridiagonal symmetric and positive definite matrix T. In fact, from

$$-H^2V = VT, \qquad V^TSV = I_n, \tag{5}$$

one has

$$-V^TSH^2V = V^TSVT = T$$

and $-SH^2$ is symmetric and positive definite.

Hence the eigenvalues of H^2 may be obtained applying an algorithm for symmetric and positive definite matrices to a half sized matrix T. Unfortunately, the decomposition (3)-(4) is not stable and can be only used to compute a limited number of extremal eigenvalues. Otherwise, it needs continue reorthogonalizations which make necessary to store all the columns of V.

Anyway, from a theoretical point of view, we derive the following symplectic transformation which is numerically more reliable than that defined in Theorem 1.

Theorem 2. *Given a real Hamiltonian matrix $H = JS$ with S symmetric and positive definite, a real symplectic matrix U exists such that*

$$HU = U \begin{pmatrix} & T \\ -I & \end{pmatrix} \equiv UJ \begin{pmatrix} I & \\ & T \end{pmatrix}$$

with T symmetric and positive definite tridiagonal matrix.

Proof. Let

$$W = -HV, \tag{6}$$

from (5) one has

$$HW = VT \tag{7}$$

and, consequently,

$$H[V\ W] = [-W\ VT] = [V\ W] \begin{pmatrix} & T \\ -I & \end{pmatrix}. \tag{8}$$

The matrix $U = [V\ W]$ is symplectic, that is

$$U^T JU = \begin{pmatrix} V^T JV & V^T JW \\ W^T JV & W^T JW \end{pmatrix} = J.$$

In fact, from (6) one has $JW = SV$ and

$$V^T JW = V^T SV = I.$$

Moreover, $V^T JV = O$ since each column v_i of V may be expressed in terms of an even polynomial of degree $2(i-1)$ in H, that is $v_i = p_{i-1}(H^2)v_1$, and for any integers i and l

$$v_i^T Jv_l = v_1^T (p_{i-1}(H^2))^T Jp_{l-1}(H^2)v_1 = v_1^T (p_{i+l-1}(H))^T Jp_{i+l-1}(H)v_1 = 0$$

being, for any vector v of appropriate length, $v^T Jv = 0$. Finally, $W^T JW = O$ for a similar reasoning. □

The above theorem states that U transforms the matrix H in a matrix

$$J \begin{pmatrix} I & \\ & T \end{pmatrix}$$

which corresponds to the tridiagonal form of symmetric matrices. This is the simplest canonical transformation which can be computed for Hamiltonian matrices and proves the existence of a canonical form for positive definite Hamiltonian matrices which is slightly different from the one obtained in Theorem 1. In fact, from $T = Q_1 D_1^2 Q_1^T$, with D_1 diagonal and Q_1 orthogonal, it results that $\mathrm{diag}(Q_1, Q_1)$ is symplectic and

$$\begin{pmatrix} Q_1^T & \\ & Q_1^T \end{pmatrix} \begin{pmatrix} & T \\ -I & \end{pmatrix} \begin{pmatrix} Q_1 & \\ & Q_1 \end{pmatrix} = \begin{pmatrix} & D_1^2 \\ -I & \end{pmatrix} \equiv J \begin{pmatrix} I & \\ & D_1^2 \end{pmatrix}.$$

Now we derive an algorithm for the computation of matrices U and T in Theorem 2. Let $V_k = [v_1, \ldots, v_k]$ and

$$T_k = \begin{pmatrix} \alpha_1 & \beta_2 & & \\ \beta_2 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_k \\ & & \beta_k & \alpha_k \end{pmatrix}, \tag{9}$$

the generic row of (3) is

$$-H^2 v_j = \beta_j v_{j-1} + \alpha_j v_j + \beta_{j+1} v_{j+1}$$

and allows us to compute v_{j+1}. The coefficient β_j is computed in order to have $v_j^T S v_j = 1$, while α_j derives from

$$0 = v_j^T S v_{j+1} = \frac{1}{\beta_{j+1}}(v_j^T S H w_j - \alpha_j) = \frac{1}{\beta_{j+1}}(-v_j^T H^T S w_j - \alpha_j),$$

where $w_j = -H v_j$ in analogy with (6).

The complete algorithm to obtain the tridiagonal form is the following:

$\tilde{v}_1$ arbitrary
$\beta_1 = (\tilde{v}_1^T S \tilde{v}_1)^{1/2}$
for $j = 1, 2, \ldots, n$
  $v_j = \tilde{v}_j / \beta_j$
  $w_j = -H v_j$
  $\alpha_j = w_j^T S w_j$
  if $j = 1$
    $\tilde{v}_{j+1} = H w_j - \alpha_j v_j$
  elseif $j < n$
    $\tilde{v}_{j+1} = H w_j - \alpha_j v_j - \beta_j v_{j-1}$
  end
  $\beta_{j+1} = (\tilde{v}_{j+1}^T S \tilde{v}_{j+1})^{1/2}$
  if $\beta_{j+1} = 0$, stop
end

We observe that $\alpha_j > 0$ and $\beta_j \geq 0$. The algorithm stops before n steps if $\beta_j = 0$ (i.e., $\tilde{v}_j = 0$), that is, when $\tilde{v}_1$ belongs to an invariant subspace of H^2. The vectors w_j need not to be stored, while the v_j are required for any reorthogonalization. If the above algorithm is applied to matrices JS with S nondefinite, then $\tilde{v}_j^T S \tilde{v}_j$ should be less than zero and the procedure breaks. A modification of the algorithm in order to overcome this problem will be investigated in future.

3 Implicit Restarting

As observed previously, the Lanczos method is widely used to compute only a small subset of the eigenvalues. In this case, the algorithm is stopped after a fixed

number k of iterations, then it restarts with a different initial vector v_1 obtained by the performed iterations. The idea is just to compute the new starting vector in order to obtain, after k steps, $\beta_{k+1} = 0$, that is an invariant subspace which allows us to compute k eigenvalues.

The implicit restarted Lanczos algorithm shows several advantages since the storage requirement is fixed, there are no spurious eigenvalues, and deflation techniques similar to those applied to the QR algorithm may be applied [8].

As an example, we now analyze how it is possible to apply this technique to the Lanczos algorithm described in the previous section. Let us start from $k+2$ steps of equation (3)

$$-H^2 V_{k+1} = V_{k+1} T_{k+1} + \beta_{k+2} v_{k+2} e_{k+1}^T,$$

where e_{k+1} represents the last unit vector of $\mathbb{R}^{k+1}$. Then, for a given real parameter μ and from $T_{k+1} - \mu I = QR$ we have the following equalities

$$\begin{aligned} (-H^2 - \mu I)V_{k+1} &= V_{k+1}(T_{k+1} - \mu I) + \beta_{k+2} v_{k+2} e_{k+1}^T \\ (-H^2 - \mu I)V_{k+1} &= V_{k+1} QR + \beta_{k+2} v_{k+2} e_{k+1}^T \\ (-H^2 - \mu I)(V_{k+1}Q) &= (V_{k+1}Q)RQ + \beta_{k+2} v_{k+2} e_{k+1}^T Q \\ -H^2(V_{k+1}Q) &= (V_{k+1}Q)(RQ + \mu I) + \beta_{k+2} v_{k+2} e_{k+1}^T Q \end{aligned} \tag{10}$$

where $RQ + \mu I$ is a symmetric and positive definite tridiagonal matrix. The last equation cannot be however considered as obtained by a Lanczos procedure since

$$\beta_{k+2} e_{k+1}^T Q = [\widehat{\beta}_{k+1} e_k^T \quad \widetilde{\beta}_{k+2}], \tag{11}$$

that is, it has two elements different from zero (instead of one, see (9)). Anyway, equation (10) is useful to define the vector $v_1^+ = (V_{k+1}Q)e_1$ (e_1 is the first unit vector of $\mathbb{R}^{k+1}$) as the starting vector of the new Lanczos iterations. Its relation with v_1 is

$$(-H^2 - \mu I)v_1 = \rho_1 v_1^+$$

where ρ_1 is the (1,1) element of R, is obtained by applying the second equation of (10) to e_1 and from $Re_1 = \rho_1 e_1$.

Let us now partition the matrices $V_{k+1}Q = [V_k^+ \ \tilde{v}_{k+1}]$ and

$$RQ + \mu I = Q^T T_{k+1} Q = \begin{pmatrix} T_k^+ & \widetilde{\beta}_{k+1} e_k \\ \widetilde{\beta}_{k+1} e_k^T & \widetilde{\alpha}_{k+1} \end{pmatrix}. \tag{12}$$

Substitution of (11) and (12) in the last equation of (10) gives

$$-H^2 [V_k^+ \ \tilde{v}_{k+1}] = [V_k^+ \ \tilde{v}_{k+1} \ v_{k+2}] \begin{pmatrix} T_k^+ & \widetilde{\beta}_{k+1} e_k \\ \widetilde{\beta}_{k+1} e_k^T & \widetilde{\alpha}_{k+1} \\ \widehat{\beta}_{k+1} e_k^T & \widetilde{\beta}_{k+2} \end{pmatrix}. \tag{13}$$

The first k columns of (13) may be rewritten in the form

$$-H^2 V_k^+ = V_k^+ T_k^+ + \beta_{k+1}^+ v_{k+1}^+ e_k^T \tag{14}$$

where

$$v_{k+1}^+ = \frac{1}{\beta_{k+1}^+}(\widetilde{\beta}_{k+1}\tilde{v}_{k+1} + \widehat{\beta}_{k+1} v_{k+2})$$

and β_{k+1}^+ is such that $(v_{k+1}^+)^T S v_{k+1}^+ = 1$. Hence equation (14) represents the implicit application of k steps of the Lanczos algorithm to the starting vector v_1^+.

This technique may be iterated. Starting from (14), we repeat one additional step of the Lanczos iteration and then again the implicit method until $\beta_{k+1} \approx 0$.

The application of p shifts $\mu_1, \mu_2, \ldots, \mu_p$ is straightforward and is used to speed-up the computation of k eigenvalues simultaneously. We summarize its use in the following algorithm.

compute $v_1, \ldots, v_k, \tilde{v}_{k+1}, \alpha_1, \ldots, \alpha_k, \beta_2, \ldots, \beta_{k+1}$ by
  means of k steps of the Lanczos algorithm
while $|\beta_{k+1}|$ is greater than a fixed tolerance
  compute $v_{k+1}, \ldots, v_{k+p}, \tilde{v}_{k+p+1}, \alpha_{k+1}, \ldots, \alpha_{k+p}$,
    $\beta_{k+2}, \ldots, \beta_{k+p+1}$ by means of p additional steps
    of the Lanczos algorithm
  $v_{k+p+1} = \tilde{v}_{k+p+1}/\beta_{k+p+1}$
  construct the matrix T_{k+p} as in (9)
  choose the parameters $\mu_1, \mu_2, \ldots, \mu_p$
  $Q = I_{k+p}$; $\widetilde{T} = T_{k+p}$
  for $i = 1, \ldots, p$
    $Q_i R_i = \widetilde{T} - \mu_i I$
    $\widetilde{T} = Q_i^T \widetilde{T} Q_i$
    $Q = Q Q_i$
  end
  define $q_{k+p,k}$ as the $(k+p, k)$ element of Q
  define $\widetilde{\beta}_{k+1}$ as the $(k+1, k)$ element of $\widetilde{T}$
  $\widehat{\beta}_{k+p+1} = \beta_{k+p+1} q_{k+p,k}$
  $[V_k^+ \ \widehat{V}_p] = V_{k+p} Q$
  define $\hat{v}_{k+1}$ as the first column of $\widehat{V}_p$
  define $v_1, \ldots, v_k$ as the k columns of V_k^+
  $\tilde{v}_{k+1} = \widetilde{\beta}_{k+1}\hat{v}_{k+1} + \widehat{\beta}_{k+p+1} v_{k+p+1}$
  define $\alpha_1, \ldots, \alpha_k$ and $\beta_2, \ldots, \beta_k$ as the main diagonal
    and the lower diagonal of $\widetilde{T}$
  $\beta_{k+1} = (\tilde{v}_{k+1} S \tilde{v}_{k+1})^{1/2}$
end

The matrix-by-matrix operations are not expensive since k and p are small with respect to n. This means that the computational cost of the overall algorithm depends only on the number of iterates in the Lanczos method.

The choice of the parameters μ_j gives rise to different strategies. For example, if we set μ_j, $j = 1, \ldots, p$, as p of the eigenvalues of T_{k+p}, that is, if we use the exact shift selection strategy [12], then $\widetilde{T}$ has the following structure

$$\begin{pmatrix} T_k^+ & \\ & D_p \end{pmatrix},$$

where D_p is diagonal with $\mu_1, \ldots \mu_p$ as main diagonal entries. This strategy gives good results, especially when used to compute the largest eigenvalues in modulus of the positive definite Hamiltonian matrix.

4 Conclusions

The Lanczos process has been modified in order to compute the eigenvalues of positive definite Hamiltonian matrices. The obtained algorithm is symplectic and requires half of the workspace of the original algorithm. Moreover, since the procedure gives a symmetric and positive definite matrix, known techniques for this class of matrices (for example, the implicit restarting) can be used to improve the computation.

References

1. Ammar, G., Mehrmann, V.: On Hamiltonian and symplectic Hessenberg forms. Linear Algebra Appl. **149** (1991), 55–72
2. Amodio, P., Iavernaro, F., Trigiante, D.: Conservative perturbations of positive definite Hamiltonian matrices, Numer. Linear Algebra Appl., (2003), in press
3. Benner, P.: Symplectic balancing of Hamiltonian matrices, SIAM J. Sci. Comput., **22 (5)** (2001), 1885–1904
4. Benner, P., Faßbender, H.: An implicit restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem, Linear Algebra Appl., **263** (1997), 75–111
5. Brugnano, L., Trigiante, D.: Solving ODEs by Linear Multistep Initial and Boundary Value Methods. Gordon & Breach, Amsterdam, 1998
6. Bunse-Gerstner, A.: Matrix factorization for symplectic QR–like methods, Linear Algebra Appl., **83** (1986), 49–77
7. Byers, R.: A Hamiltonian QR–algorithm, SIAM J. Sci. Stat. Comput., **7** (1986), 212–229
8. Calvetti, D., Reichel, L., Sorensen, D.C.: An implicitly restarted Lanczos method for large symmetric eigenvalue problems, ETNA, Electron. Trans. Numer. Anal., **2** (1994), 1–21
9. Lancaster, P., Rodman, L.: The algebraic Riccati equation. Oxford University Press, Oxford, 1995
10. Lin, W., Mehrmann, V., Xu, H.: Canonical forms for Hamiltonian and symplectic matrices and pencils, Linear Algebra Appl. **302-303** (1999), 469–533

11. Rosen, I., Wang, C.: A multilevel technique for the approximate solution of operator Lyapunov and algebraic Riccati equations. SIAM J. Matrix Anal. Appl. **32** (1992), 514–541
12. Sorensen, D.C.: Implicit application of polynomial filters in a k-step Arnoldi method, SIAM J. Matrix Anal. Appl. **13** (1992), 357–385
13. Van Loan, C.: A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix, Linear Algebra Appl., **16** (1984), 233–251