THE LANCZOS ALGORITHM WITH SELECTIVE ORTHOGONALIZATION*

B.N. Parlett
Mathematics Department, University of California, Berkeley

D.S. Scott
Union Carbide

ABSTRACT

A new stable and efficient implementation of the Lanczos algorithm is presented.

The Lanczos algorithm is a powerful method for finding a few eigenvalues and eigenvectors at one or both ends of the spectrum of a symmetric matrix A. The algorithm is particularly effective if A is large and sparse in that the only way in which A enters the calculation is through a subroutine which computes Av for any vector v. Thus the user is free to take advantage of any sparsity structure in A and A need not even be represented as a matrix at all.

The simple Lanczos algorithm procedes as follows. Choose q_1, an arbitrary unit vector, and define $\beta_0 = 0$ and $q_0 = 0$. Then for $j = 1,2,\ldots$ do 1 to 5.

1. $u_j = Aq_j - q_{j-1}\beta_{j-1}$
2. $\alpha_j = q_j^* u_j$
3. $r_j = u_j - q_j\alpha_j$
4. $\beta_j = \|r_j\|$
5. if $\beta_j = 0$ stop, else $q_{j+1} = r_j/\beta_j$

One cycle through 1-5 is a Lanczos step. Note that only q_{j-1} and q_j are needed to compute q_{j+1} which is another attractive feature of the algorithm.

In exact arithmetic, if $Q_j = (q_1, q_2, \ldots, q_j)$ then it can be shown (cf. ref. 1) that Q_j is an orthonormal matrix, i.e. $1 - Q_j^* Q_j = 0$, and in addition $Q_j^* A Q_j = T_j$ where

*Research supported by Office of Naval Research Contract N00014-76-C-0013.

$$T_j = \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \cdot & \cdot & \\ & & \cdot & \cdot & \beta_{j-1} \\ & & & \beta_{j-1} & \alpha_j \end{bmatrix}$$

is tridiagonal. Furthermore if $T_j = S_j \Theta_j S_j^*$ is the spectral decomposition of T_j with $\Theta_j = \text{diag}(\theta_1^{(j)}, \theta_2^{(j)}, \ldots, \theta_j^{(j)})$ and if $Y_j \equiv (y_1^{(j)}, y_2^{(j)}, \ldots, y_j^{(j)}) = Q_j S_j$ then $(\theta_i^{(j)}, y_i^{(j)})$, $i = 1,2,\ldots,j$, are the (optimal) Rayleigh-Ritz approximations to eigenpairs of A derivable from span(Q_j), the subspace spanned by $q_1, q_2, \ldots, q_j$.

Finally and remarkably, the residual norm of (y_i, θ_i) can be computed without computing the vector y_i. Namely

$$\|Ay_i - y_i\theta_i\| = \beta_{ji}$$

where $\beta_{ji} = \beta_j |s_{ji}|$ and s_{ji} is the (j,i) element of S_j. The quantities β_{ji} show how it is possible for <u>some</u> of the Ritz values (θ's) to be accurate without the appearance of a small β_j. If s_{ji} is tiny then $\theta_i^{(j)}$ will be accurate even if β_j is not small at all.

By construction span(Q_j) = span($q_1, Aq_1, \ldots, A^{j-1}q_1$), a Krylov subspace. It can be shown (refs. 2 and 3) that Rayleigh-Ritz approximations converge rapidly as j increases to well separated extreme eigenpairs of A (those near either end of the spectrum).

Unfortunately, as was known to Lanczos when he introduced the algorithm (ref. 4), finite precision causes the computed quantities to diverge completely from their theoretical counterparts. The Lanczos vectors (the q's) inevitably lose their mutual orthogonality and approach linear dependence. This is the infamous "loss of orthogonality" in the Lanczos algorithm.

Lanczos himself recommended that the simple Lanczos algorithm be augmented by a full reorthogonalization of each newly computed q_{j+1}. That is, q_{j+1} is explicitly orthogonalized against all preceding Lanczos vectors (q_i, for $i \leq j$). This not only greatly increases the number of computations required to compute q_{j+1}, it also requires that all the q's be kept in fast store.

This poses a serious dilemma. For large problems it will be too costly to take more than a few steps using full reorthogonalization but linear independence will surely be lost without some sort of corrective procedure. Selective Orthogonalization (hereafter called SO) interpolates between full reorthogonalization and simple Lanczos to obtain the best of both worlds. Robust linear independence is maintained among the columns of Q_j at a cost which is close to that of simple Lanczos.

SO is based on the analysis of the simple Lanczos algorithm in finite arithmetic given by C. Paige in his doctoral thesis (ref. 3). Paige showed that the loss of orthogonality among the columns of Q_j is highly structured when viewed in the basis of Ritz vectors, the columns of $Y_j = Q_j S_j$, rather than in the basis of Q_j itself.

Theorem (Paige). For any step j of the simple Lanczos algorithm and any $i \leq j$, $$\lvert y_i^{(j)*} q_{j+1} \rvert = \epsilon \lVert A \rVert \gamma_{ji} / \beta_{ji}$$ where $\gamma_{ji} \doteq 1$ and ϵ is the working precision.

A proof is given in reference 5.*

It can also be shown (cf. ref. 5 or 6) that β_{ji} is a very good estimate of the residual norm of $y_i^{(j)}$ despite rounding errors. Thus Paige's Theorem shows that q_{j+1} will lose orthogonality only in the direction of Ritz vectors with small β_{ji}, that is those Ritz vectors which are converging to eigenvectors. This can be stated as

loss of orthogonality $\Leftrightarrow$ convergence

To maintain orthogonality among the Lanczos vectors below some threshold value $\tau \leq 1$ it is only necessary to orthogonalize q_{j+1} against those Ritz vectors which satisfy

$$\lvert y_i^{(j)*} q_{j+1} \rvert > \tau \qquad (1)$$

By Paige's Theorem equation (1) holds only if $\beta_{ji} \leq \epsilon \lVert A \rVert \gamma_{ji}/\tau \doteq \epsilon \lVert A \rVert/\tau$. Thus it is possible to determine which Ritz vectors achieve the threshold (1) merely by inspecting the β_{ji} which can be computed via a small $(j \times j)$ eigenproblem. There are strong theoretical arguments in favor of the choice $\tau = \sqrt{\epsilon}$ (ref. 5 or 6).

Thus SO modifies the simple Lanczos algorithm by explicitly orthogonalizing q_{j+1} against all the Ritz vectors which satisfy

$$\beta_{ji} \leq \sqrt{\epsilon} \lVert A \rVert \qquad (2)$$

We call any Ritz vector which satisfies (2) a good Ritz vector. Good Ritz vectors are already rather well converged and few (if any) of the Ritz vectors at step j will be good, which explains the computational efficiency of the scheme.

In practice it is possible to implement SO even more efficiently. It is not necessary to recompute the good Ritz vectors for orthogonalizing q_{j+1} at each step j as Ritz vectors computed at earlier steps can be used instead.

* The detailed rounding error analysis needed to complete the proof of Paige's Theorem is given in reference 3.

Furthermore it is not necessary to orthogonalize against a particular good Ritz vector at every step.

In conclusion, SO is an efficient way to maintain robust linear independence among the columns of Q_j and so allow the Lanczos algorithm to be run almost as originally conceived. SO points the way to a high quality subroutine package which can be used off the shelf for large sparse symmetric eigenvalue problems.

REFERENCES

1. Kahan, W. and Parlett, B.: How Far Should You Go with the Lanczos Algorithm? In Sparse Matrix Computations, J. Bunch and D. Rose, eds. Academic Press, New York, 1976.

2. Kaniel, S.: Estimates for Some Computational Techniques in Linear Algebra. Mathematics of Computation, vol. 20, no. 95, July 1966, pp. 369-378.

3. Paige, C. C.: The Computation of Eigenvalues and Eigenvectors of Very Large Sparse Matrices. Ph.D. Thesis, University of London, 1971.

4. Lanczos, C.: An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators. J. Res. Nat. Bur. Stand., vol. 45, 1950, pp. 255-282.

5. Scott, D. S.: Analysis of the Symmetric Lanczos Process. Ph.D. Thesis, Mathematics Department, University of California, Berkeley, 1978.

6. Parlett, B. N.: The Symmetric Eigenvalue Problem. Prentice-Hall, Englewood Cliffs, N.J., 1979.