Research Contribution

A New Reorthogonalization in the Lanczos Algorithm

HITOSHI IMAI*, MAKOTO NATORI* and EIJI KAWAMURA*

A new reorthogonalization in the Lanczos algorithm is proposed. In this method, the loss of orthogonality among Lanczos vectors is monitored by a recurrence formula. When it is detected, Ritz vectors, which correspond to converging Ritz values, are employed for the orthogonalization. The main feature of this method is that eigenvalues are obtained in any quantity and precision that the user requires. Therefore, the Lanczos algorithm can be stopped as soon as the desired eigenvalues have been obtained.

Numerical computations are carried out to evaluate the method. Here, matrices obtained by discretization of the two-dimensional Laplace operator are used. The features mentioned above are confirmed numerically, and an improvement in the computational time is confirmed.

1. Introduction

The Lanczos algorithm is an interesting method of solving large symmetric eigenvalue problems, for it has the following two attractive features. The first is that it can save machine storage when the matrix is sparse. This is because, unlike other methods (such as the Householder transformation), does not change the given matrix in the course of computation. The second is that theoretically computations can be stopped halfway as soon as the desired eigenvalues have been obtained. This is because accurate eigenvalues are obtained iteratively from outside, and their accuracy is improved if any computational error can be neglected [3]. However, since the Lanczos algorithm is sensitive to roundoff, some computed eigenvalues are not true. This is caused by the loss of orthogonality among Lanczos vectors as the Lanczos algorithm proceeds. Hence some reorthogonalizations have been introduced [3-6]. However, even when these reorthogonalizations are used it is still difficult to know the accuracy of computed eigenvalues or to distinguish true and false eigenvalues, so the second feature has not been realized. In this paper a new reorthogonalization utilizing this feature is proposed.

The Lanczos algorithm was first introduced in 1950 to solve the eigenvalue problem [1]

$$Ax=\lambda x, \tag{1}$$

where A is a symmetric $n\times n$ matrix. This is a simple iterative method, and may be considered as a way of constructing a tridiagonal matrix through orthogonal transformation. The original Lanczos algorithm is given as follows [2].

(I) Choose a starting vector q_1 ($\|q_1\|=1$), and set

$$u_1=Aq_1. \tag{2}$$

(II) Compute α_j and β_j ($j=1\sim n$) iteratively as follows:

$$\alpha_j=u_j^Tq_j, \tag{3}$$

$$r_j=u_j-\alpha_jq_j, \tag{4}$$

$$\beta_j=\|r_j\|, \tag{5}$$

$$q_{j+1}=r_j/\beta_j, \tag{6}$$

$$u_{j+1}=Aq_{j+1}-\beta_jq_j. \tag{7}$$

Here $\{q_j\}_{j=1,n}$ are called Lanczos vectors, which become mutually orthogonal. The procedure (3)-(7) is called the j-th Lanczos step. At the j-th Lanczos step, Ritz values $\{\theta_i^j\}_{i=1,j}$ and Ritz vectors $\{y_i^j\}_{i=1,j}$ can be defined as the following eigensystem:

$$T_js_i^j=\theta_i^js_i^j, \tag{8}$$

$$y_i^j=V_js_i^j, \tag{9}$$

where

$$T_j=\begin{bmatrix}\alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \cdot & & \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & \beta_{j-1} \\ & & & \beta_{j-1} & \alpha_j\end{bmatrix}, \quad V_j=[q_1\cdots q_j]. \tag{10}$$

From the definition, if $j=n$, then the Ritz values and Ritz vectors coincide with the eigenvalues and eigenvectors of the original problem (1), respectively. Note that the Lanczos algorithm is regarded as a natural way of implementing the Rayleigh-Ritz procedure on the sequence of Krylov subspace [3]. Thus Ritz values are re-

*Institute of Information Sciences and Electronics, University of Tsukuba, Ibaraki 305, Japan.

Table 1 Reorthogonalizations SO and PRO.

Method	Monitor	Vectors for reorthogonalization
SO	$\beta_{j,i}$	Ritz vectors
PRO	$\omega_{j+1,i}$	Lanczos vectors

Table 2 Matrices used in numerical computations.

Dimension	60	200	400	800
ID	6	20	25	25
JD	10	10	16	32

garded as approximations to the eigenvalues in (1).

This simple Lanczos algorithm is sensitive to roundoff, so the orthogonality among Lanczos vectors is lost as the Lanczos steps proceed. This loss of orthogonality causes redundant copies of Ritz pairs, so some computed eigenvalues are not true. To relax the sensitivity to roundoff, the reorthogonalization of Lanczos vectors is carried out. However, since the matrices are usually large, full reorthogonalization (FRO) is not available. Some partial reorthogonalizations have therefore been introduced. In particular, the selective orthogonalization (SO) method developed by Parlett and Scott [4] and the partial reorthogonalization (PRO) method developed by Simon [5, 6] are famous. In SO the loss of orthogonality at the j-th step is monitored by

$$\beta_{j,i}=\beta_j\sigma_{j,i}, \quad 1\leqq i\leqq j \tag{11}$$

where $\sigma_{j,i}$ is the j-th component of s^j_i. If $|\beta_{j,i}|<\varepsilon^{1/2}$, then the i-th Ritz vector is chosen for the reorthogonalization. From now on, ε stands for the machine epsilon. The Ritz values are usually computed before the Ritz vectors. Therefore, SO needs to solve eigenvalue problem (8) at every step and to compute adequate Ritz vectors. On the other hand, in PRO the loss of orthogonality at the j-th step is monitored by $\omega_{j+1,i}$ where $\omega_{j+1,i}$ satisfies the following recurrence formula: [6, 7].

$$\omega_{k,0}=0, \quad \omega_{k,k}=1, \quad 1\leqq k\leqq j+1, \tag{12}$$

$$\omega_{k,k-1}=\psi_k, \quad 2\leqq k\leqq j, \tag{13}$$

$$\begin{aligned}\omega_{k+1,i}=\frac{1}{\beta}\{&\beta_i\omega_{k,i+1}+(\alpha_i-\alpha_k)\omega_{k,i}+\beta_{i-1}\omega_{k,i-1}\\&-\beta_{k-1}\omega_{k-1,i}\}+\phi_{k,i},\\&\qquad 2\leqq k\leqq j, \quad 1\leqq i\leqq k-1\end{aligned} \tag{14}$$

where

$$\psi_k=\varepsilon n\frac{\beta_1}{\beta_k}\Psi, \quad \Psi\in N(0, 0.6), \tag{15}$$

$$\phi_{k,i}=\varepsilon(\beta_k+\beta_i)\Phi, \quad \Phi\in N(0, 0.3). \tag{16}$$

If $|\omega_{j+1,i}|>\varepsilon^{1/2}$, then the i-th Lanczos vector is chosen for the reorthogonalization. After the reorthogonalization has been performed, $\omega_{j+1,i}$ is set as $\omega_{j+1,i}\in N(0, 1.5)\varepsilon$. The monitors and the vectors in reorthogonalizations SO and PRO are compared in Table 1.

$$\begin{array}{ccc}\theta_1^{j-1} & \longrightarrow & \theta_1^j\\ \theta_2^{j-1} & \longrightarrow & \theta_2^j\\ \vdots & & \vdots\\ \theta_{j-2}^{j-1} & \longrightarrow & \theta_{j-1}^j\\ \theta_{j-1}^{j-1} & \longrightarrow & \theta_j^j\end{array}$$

Fig. 1 Correspondence of Ritz values for the convergence checking in RIC.

Although these reorthogonalizations give a larger number of true eigenvalues, it is still difficult to discriminate true eigenvalues and to know the accuracy of computed eigenvalues. In the next section, we introduce a new reorthogonalization that can indicate true eigenvalues and their accuracy. This method is also effective in terms of computational time.

2. A New Reorthogonalization

Our new reorthogonalization is based on the following considerations.

To estimate the computed eigenvalues it is better to investigate the behavior of Ritz values among the Lanczos steps. This investigation is only necessary when the behavior changes. From Paige's theorem, this change in the behavior occurs when the orthogonality among Lanczos vectors is lost [2, 3]. This loss of orthogonality is simply monitored by $\omega_{j+1,i}$ as in PRO. Therefore

(i) The behavior of Ritz values is checked only when $\omega_{j+1,i}$ becomes greater than $\varepsilon^{1/2}$.

At the same time, this investigation may also be used for the reorthogonalization. The problem lies in the choice of the vectors to which the Lanczos vectors are orthogonalized. PRO applies the value $\omega_{j+1,i}$ not only to the monitor of the orthogonality but also to the choice of vectors. However, numerical results show that, while SO is successful, PRO is not (see Table 3). Hence, Ritz vectors seem to be suitable for the reorthogonalization. In SO they are chosen by estimating $\beta_{j,i}$. However, it takes a long time to compute $\beta_{j,i}$ since all the eigenvectors in Eq. (8) are necessary. The following choice of Ritz vectors is therefore introduced:

(ii) Choose the tolerance τ for the convergence check. (For example, choose τ in $10^{-10}\sim 10^{-12}$ for double-precision calculations.)

(iii) Then perform a convergence check of Ritz values from outside sequentially by comparing the Ritz values at the j-th and $(j-1)$-st steps (see Fig. 1).

(iv) Ritz vectors that correspond to converging Ritz values are chosen for the reorthogonalization.

This choice is partially supported both by Paige's theorem, which shows the inclination of the Lanczos vector, and by the fact that Ritz values are Rayleigh-

Ritz approximations to eigenvalues. Here note that if the loss of the orthogonality is detected sequentially, then computations of Ritz values at the last step can be omitted. Moreover, there may be no need to recheck Ritz values that have already been determined to be convergent. There may also be no need to recompute Ritz vectors that correspond to these Ritz values if formerly computed Ritz vectors are retained in the memory.

Hereafter, this new reorthogonalization (i)-(iv) is called reorthogonalization with improved convergence-checking (RIC). In RIC, converging Ritz values are regarded as eigenvalues, and at the same time their error is roughly estimated from the tolerance τ. Therefore RIC can be expected to offer some eigenvalues for the given accuracy at any time. Moreover, it can be expected to stop the Lanczos algorithm before it begins to fail. More precisely, if the monitored outermost Ritz values tend to diverge at some step, then the Lanczos algorithm will fail from this step. However, even in this case RIC can be expected to offer some accurate eigenvalues, because converging Ritz values obtained before this step are regarded as good approximations.

Table 3 Averaged relative errors of computed eigenvalues.

Dimension	60	200	400	800
PRO	7.6×10^{-11}	2.7×10^{-1}	3.9×10^{-1}	4.6×10^{-1}
SO	2.9×10^{-16}	5.1×10^{-16}	8.9×10^{-16}	—
RIC*)	3.5×10^{-16}	1.2×10^{-15}	4.5×10^{-12}	1.7×10^{-2}

*) $\tau=10^{-12}$.

Table 4 Number of computed eigenvalues whose relative errors are less than 10^{-8}.

Dimension	60	200	400	800
PRO	59	0	0	0
SO	60	200	400	—
RIC	60	200	400	436

Table 5 Number of converging Ritz values and averaged relative errors in RIC for $\tau=10^{-12}$.

Dimension	Number of converging Ritz values*)	Averaged relative errors
60	17	4.0×10^{-14}
	17	4.6×10^{-14}
200	78	1.6×10^{-15}
	79	4.5×10^{-16}
400	178	1.4×10^{-14}
	179	5.8×10^{-15}
800	219	1.9×10^{-10}
	219	2.8×10^{-11}

*) The number is determined by RIC; the number from the minimum eigenvalue is in the upper row, and the number from the maximum eigenvalue is in the lower row.

3. Numerical Results

In this section we present the numerical results obtained by RIC, SO, and PRO. The matrices used here are obtained by discretization of the two-dimensional Laplace operator that is defined in the rectangular domain and subject to the homogeneous Dirichlet boundary conditions. Exact eigenvalues for these matrices are given as follows:

$$\lambda_{ij}=4\left\{\sin^2\frac{\pi i}{2(ID+1)}+\sin^2\frac{\pi j}{2(JD+1)}\right\},$$
$$i=1\sim ID,\ j=1\sim JD \tag{17}$$

where $(ID+1)$ and $(JD+1)$ are divisions of the rectangular domain along the x-axis and the y-axis, respectively.

Numerical computations are carried out for the matrices in Table 2. Matrices over 800×800 are not used because of the limitation on computational time (see Table 6).

In Table 3, the averaged relative errors of eigenvalues are presented for various matrices with simple eigenvalues. Although SO is not carried out for the 800×800 matrix because of the limitation on computational time, SO would give as accurate results as RIC in this case.

Table 4 exhibits the number of computed eigenvalues with a relative error of less than 10^{-8}. This shows that RIC is superior in precision to PRO.

Table 5 shows the number of Ritz values that are determined by RIC to be convergent. Comparison with Table 4 shows that the approximate eigenvalues can be guaranteed up to the number of converging Ritz values. This observation implies that RIC can stop the Lanczos step as soon as the desired eigenvalues have been obtained.

Table 6 shows the CPU time on a HITAC M682H. Since the CPU time strongly depends on the matching between the program and the machine, these values are not absolute. However, RIC is expected to be superior in computational time to SO.

In Table 7 the averaged relative errors of eigenvalues in RIC for different τ are presented.

Table 6 CPU time (sec., HITAC M682H).

Dimension	60	200	400	800
PRO	0.2	2.2	10.4	56.4
SO	3.0	118.1	1127.8	—
RIC	2.4	89.1	688.5	5402.2

Table 7 τ dependence of RIC for 400×400 matrix.

τ	10^{-10}	10^{-12}
Averaged relative errors of eigenvalues	4.1×10^{-16}	4.5×10^{-12}

When the eigenvalues are degenerate, some Ritz values appear suddenly at a certain step and afterwards converge to degenerate eigenvalues. In this case, neither SO nor PRO is applicable. However, RIC is useful even in this case because it can be programmed to give knowledge about the step at which the convergence is numerically lost. This informs the user that eigenvalues are degenerate or that the Lanczos algorithm fails halfway.

4. Conclusion

A new reorthogonalization of the Lanczos algorithm, called RIC, is proposed. In this method, a recurrence formula is used to monitor loss of orthogonality among Lanczos vectors. When a loss of orthogonality is detected, the adequate Ritz vectors are chosen by investigating the convergence of Ritz values, and are used in the reorthogonalization.

Numerical results show that if the eigenvalues are simple, RIC is as successful as and faster than SO. Moreover, it allows the Lanczos algorithm to be stopped halfway as soon as the desired eigenvalues have been obtained.

5. Acknowledgment

Numerical computations were carried out on a HITAC M682H at the Computer Centre, University of Tokyo.

References

1. Lanczos, C. An iterative method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Standards*, **45** (1950), 255-282.
2. Paige, C. C. Computational Variants of the Lanzcos Method for the Eigenproblem, *J. Inst. Maths. Appl.*, **10** (1972), 373-381.
3. Parlett, B. N. The Symmetric Eigenvalue Problem, Prentice-Hall (1980).
4. Parlett, B. N. and Scott, D. S. The Lanczos Algorithm with Selective Orthogonalization, *Math. Comput.*, **33** (1979), 217-238.
5. Simon, H. D. Analysis of the Symmetric Lanczos Algorithm with Reorthogonalization Methods, *Linear Alg. and Its Applications*, **61** (1984), 101-131.
6. Simon, H. D. The Lanczos Algorithm with Partial Reorthogonalization, *Math. Comput.*, **42** (1984), 115-142.
7. Takahashi, H. and Natori, M. Eigenvalue Problem of Large Sparse Matrices, Rep. Compt. Center University of Tokyo, **4** (1971-1972), 129-148.

(Received August 29, 1989; revised April 16, 1990)