Estimates of Eigenvalues for
Iterative Methods*

By Gene H. Golub and Mark D. Kent**

Abstract. We describe a procedure for determining estimates of the eigenvalues of operators used in various iterative methods for the solution of linear systems of equations. We also show how to determine upper and lower bounds for the error in the approximate solution of linear equations, using essentially the same information as that needed for the eigenvalue calculations. The methods described depend strongly upon the theory of moments and Gauss quadrature.

1. Introduction. We wish to solve the system of equations

$$(1) \qquad\qquad A\mathbf{x} = \mathbf{b}$$

where A is an $n \times n$, symmetric, positive definite matrix. It is frequently desirable to rewrite (1) as

$$(2) \qquad\qquad M\mathbf{x} = N\mathbf{x} + \mathbf{c},$$

where M and N are symmetric and M is positive definite. We are interested in those situations where it is a much simpler computational task to solve the system $M\mathbf{z} = \mathbf{d}$ than it is to solve (1).

We shall use an *iterative* procedure of the form

$$(3) \quad \mathbf{x}_{k+1} = \mathbf{x}_{k-1} + \omega_{k+1}(\delta_k \mathbf{z}_k + \mathbf{x}_k - \mathbf{x}_{k-1}), \qquad k = 0, 1, 2, \ldots; \ \mathbf{x}_{-1} = \mathbf{0}$$

and a splitting where $A = M - N$ and $\mathbf{c} = \mathbf{b}$ in (2). The "generalized residual" vectors, $\mathbf{z}_k$, are given by $M\mathbf{z}_k = \mathbf{b} - A\mathbf{x}_k \equiv \mathbf{b} - (M - N)\mathbf{x}_k$. Depending on the choice of parameters, (3) describes the conjugate gradient method, the Richardson second-order method, or the Chebyshev semi-iterative (CSI) method [1]. The success of the latter two methods depends on having good estimates for the smallest and largest eigenvalues of $M^{-1}N$. In this paper we show how to obtain the eigenvalue estimates using modified moments calculated from successive iterates.

The algorithm for estimating the optimal parameters is based on the *modified Chebyshev* algorithm given in [4]. See also [5], [11], [12].

In Section 2, we show how modified moments arise naturally in certain iterative methods and how to determine successive elements in a tridiagonal matrix whose eigenvalues approach the eigenvalues of the iteration matrix. Improvements, specifically for the CSI method, are given in Section 3. Determining error bounds is the topic of Section 4 and this is followed by computational results in Section 5.

2. Modified Moments and Iterative Methods. We now demonstrate how modified moments arise in certain iterative methods for the solution of symmetric, positive definite linear systems $A\mathbf{x} = \mathbf{b}$. In the following discussion let $\delta = \delta_k$ be fixed.

The vectors $\mathbf{z}_k$ are related to $\mathbf{z}_0$ by

$$(4) \qquad \mathbf{z}_k = p_k(B)\mathbf{z}_0, \qquad B \equiv I - \delta M^{-1}A,$$

where $p_k(\lambda)$ is a polynomial of degree k satisfying the three-term recurrence

$$(5) \quad p_{k+1}(\lambda) = \omega_{k+1}\lambda p_k(\lambda) + (1 - \omega_{k+1})p_{k-1}(\lambda), \qquad p_{-1}(\lambda) = 0, \ p_0(\lambda) = 1.$$

The matrix B has a complete set of M-orthogonal eigenvectors, $Q = [\mathbf{q}_1, \ldots, \mathbf{q}_n]$, corresponding to real eigenvalues $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ [13]. Clearly, $p_k(B) = Qp_k(\Lambda)Q^{-1}$. Write $\mathbf{z}_0$ as a linear combination of the eigenvectors $\mathbf{q}_i$ of B,

$$\mathbf{z}_0 = \sum_{i=1}^{n} \alpha_i \mathbf{q}_i;$$

then, from (4), we have

$$\mathbf{z}_k = \sum_{i=1}^{n} \alpha_i p_k(\lambda_i)\mathbf{q}_i.$$

Now, form the inner product

$$(6) \qquad \langle \mathbf{z}_k, \mathbf{z}_l \rangle = (\mathbf{z}_k, M\mathbf{z}_l) = \sum_{i=1}^{n} \alpha_i^2 p_k(\lambda_i)p_l(\lambda_i) \equiv \int p_k(\lambda)p_l(\lambda)\,d\alpha(\lambda),$$

where $\alpha(\lambda)$ is a discrete nonnegative distribution with jumps of height α_i^2 at each eigenvalue λ_i.

Associated with the distribution $\alpha(\lambda)$ is a set of discrete orthogonal polynomials $\{\psi_k(\lambda)\}_{k=0}^{n}$ such that

$$\int \psi_i(\lambda)\psi_j(\lambda)\,d\alpha(\lambda) = 0 \quad \text{when } i \neq j.$$

The "final" orthogonal polynomial, $\psi_n(\lambda)$, has a zero at each point of increase of the distribution, i.e., $\psi_n(\lambda_i) = 0$, $i = 1, \ldots, n$.

With each iteration of (3) we can calculate a modified moment of the distribution $\alpha(\lambda)$:

$$(7) \qquad \nu_k = \langle \mathbf{z}_k, \mathbf{z}_0 \rangle = \int p_k(\lambda)\,d\alpha(\lambda),$$

and these quantities can be used in the modified Chebyshev algorithm [4] to generate the orthogonal polynomials.

Let

$$\lambda p_j(\lambda) = b_j p_{j+1}(\lambda) + a_j p_j(\lambda) + c_j p_{j-1}(\lambda)$$

and

$$(8) \qquad \lambda \psi_j(\lambda) = \beta_j \psi_{j+1}(\lambda) + \alpha_j \psi_j(\lambda) + \gamma_j \psi_{j-1}(\lambda).$$

The modified Chebyshev algorithm starts with

$$
\begin{aligned}
(9) \quad & \sigma_{-1,l} = 0, \quad \sigma_{0,l} = \nu_l \quad \text{for } l = 0, 1, \ldots, 2m - 1, \\
& \alpha_0 = a_0 + b_0 \frac{\sigma_{01}}{\sigma_{00}}, \qquad \gamma_0 = 0,
\end{aligned}
$$

and continues with $k = 1, 2, \ldots, m - 1$:

$$\sigma_{kl} = \frac{1}{\beta_{k-1}}[b_l\sigma_{k-1,l+1} + (a_l - \alpha_{k-1})\sigma_{k-1,l} + c_l\sigma_{k-1,l-1} - \gamma_{k-1}\sigma_{k-2,l}]$$

$$\text{for } l = k, k + 1, \ldots, 2m - k - 1,$$

(10)
$$\alpha_k = a_k + b_k\frac{\sigma_{k,k+1}}{\sigma_{kk}} - b_{k-1}\frac{\sigma_{k-1,k}}{\sigma_{k-1,k-1}},$$

$$\gamma_k = b_{k-1}\frac{\sigma_{kk}}{\sigma_{k-1,k-1}}.$$

The coefficients $\{\beta_i\}$ are not specified by the algorithm and can be chosen to scale the generated polynomials. The choice $\beta_i = 1$ will generate monic polynomials. For the polynomials $p_k(\lambda)$ we have $b_j = 1/\omega_{j+1}$, $a_j = 0$, and $c_j = (\omega_{j+1} - 1)/\omega_{j+1}$ and we choose $\beta_j = b_j$.

This computational procedure is equivalent to the Lanczos algorithm [10]. Thus the extreme eigenvalues of the quasi-symmetric tridiagonal matrix formed by the coefficients in (8),

(11)
$$J_k = \begin{bmatrix} \alpha_0 & \beta_0 & & & \\ \gamma_1 & \alpha_1 & \beta_1 & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma_{k-2} & \alpha_{k-2} & \beta_{k-2} \\ & & & \gamma_{k-1} & \alpha_{k-1} \end{bmatrix},$$

will, in general, provide good approximations to the extreme eigenvalues of B as k increases. Once we have sufficiently accurate estimates for the largest and smallest eigenvalue, we can restart the iterative solution of (1) with (*near*) optimal parameters.

3. Improving the Calculation of the Modified Moments.
We now derive improved expressions for the modified moments ν_k in terms of the vectors z_k for the CSI method. After $2m$ iterations of (3) we can determine J_m and have, in principle, the zeros of $\psi_m(\lambda)$.

Obviously, it would be advantageous to have the roots of $\psi_m(\lambda)$ after only m iterations of (3). It is possible to attain this goal if the iteration (3) is the Chebyshev semi-iterative method [9]. For the CSI method we have estimates, a and b, of the smallest and largest eigenvalues of $M^{-1}A$, and the parameters, in terms of $\mu = (b - a)/(b + a)$, are given by

$$\delta = \frac{2}{b + a}, \quad \omega_{k+1} = \frac{1}{1 - \frac{\mu^2}{4}\omega_k} \quad \text{with } \omega_1 = 1, \ \omega_2 = \frac{1}{1 - \frac{\mu^2}{2}}.$$

It is well known [9] that

(12)
$$p_k(\lambda) = \frac{C_k(\lambda/\mu)}{C_k(1/\mu)},$$

where

$$C_k(x) = \begin{cases} \cos(k\cos^{-1}x), & |x| \le 1, k \ge 0; \\ \cosh(k\cosh^{-1}x), & |x| \ge 1, k \ge 0. \end{cases}$$

$C_k(x)$ satisfies the three-term recurrence relation $C_{k+1}(x) = 2xC_k(x) - C_{k-1}(x)$

with initial conditions $C_0(x) = 1$ and $C_1(x) = x$. From the identity $\cos(k + l)\theta = 2\cos k\theta \cos l\theta - \cos(k - l)\theta$, we have $C_{k+l} = 2C_k C_l - C_{|k-l|}$. Hence, for $l = k$ and $l = k + 1$,

$$(13) \qquad C_{2k} = 2C_k^2 - C_0, \qquad C_{2k+1} = 2C_k C_{k+1} - C_1.$$

Working from (12) and making liberal use of these identities, we obtain

$$p_{2k}(\lambda) = \frac{C_{2k}(\lambda/\mu)}{C_{2k}(1/\mu)} = p_k^2(\lambda) + \frac{1}{C_{2k}(1/\mu)}[p_k^2(\lambda) - 1],$$

which leads directly to the modified moment

$$(14) \qquad \nu_{2k} = \langle \mathbf{z}_k, \mathbf{z}_k \rangle + \frac{1}{C_{2k}(1/\mu)}(\langle \mathbf{z}_k, \mathbf{z}_k \rangle - \nu_0).$$

Similarly, we can derive

$$(15) \qquad \nu_{2k+1} = \langle \mathbf{z}_k, \mathbf{z}_{k+1} \rangle + \frac{1}{\mu C_{2k+1}(1/\mu)}(\langle \mathbf{z}_k, \mathbf{z}_{k+1} \rangle - \nu_1).$$

The first two modified moments are given by $\nu_0 = \langle \mathbf{z}_0, \mathbf{z}_0 \rangle$ and $\nu_1 = \langle \mathbf{z}_0, \mathbf{z}_1 \rangle$. Hence, with each generalized residual vector we can compute two more modified moments.

4. Error Bounds. In this section we indicate how to obtain error bounds for an approximate solution to the system $A\mathbf{x} = \mathbf{b}$. In [2] it is shown how moments, and the associated Gauss-Radau quadrature rule, can be used to obtain bounds on $\|\mathbf{x} - \mathbf{x}_0\|_2$ (i.e., the error in the initial approximate solution), as well as bounds for $\|\mathbf{x} - \mathbf{x}_i\|_2$ for the Jacobi method. These results are extended in [3] to the conjugate gradient method (see also [7]).

We consider an iterative method of the form $\mathbf{x}_{k+1} = B\mathbf{x}_k + \mathbf{g}$ and let $B = I - \delta A$. Hence, the errors, $\varepsilon_k = \mathbf{x} - \mathbf{x}_k$, satisfy

$$\varepsilon_k = B\varepsilon_{k-1} = B^k \varepsilon_0.$$

Now since $\mathbf{z}_k = A\varepsilon_k$, and $A = \frac{1}{\delta}(I - B)$, we see that $A\varepsilon_k = p_k(B)\mathbf{z}_0$, and hence

$$(16) \qquad \varepsilon_k^T A\varepsilon_k = \delta \mathbf{z}_0^T p_k(B)(I - B)^{-1} p_k(B)\mathbf{z}_0 = \delta \nu_0 \int \frac{p_k^2(\theta)}{1 - \theta} \, d\alpha(\theta).$$

Define

$$q_k(\theta) \equiv \frac{p_k^2(\theta)}{1 - \theta} = \frac{p_k^2(\theta) - 1}{1 - \theta} + \frac{1}{1 - \theta}.$$

Here, the first term on the right is a polynomial of degree $2k - 1$ (since $p_k(1) = 1$; see (12)) and

$$\frac{d^r}{d\theta^r}(1 - \theta)^{-1} = r!(1 - \theta)^{-(r+1)} \geq 0 \quad \text{for } |\theta| < 1.$$

Thus, if we apply a $(k + 1)$-point Gauss-Radau quadrature rule to $q_k(\theta)$, we have as a remainder term [2], [3]

$$R[q_k(\theta)] = (1 - \eta)^{-(2k+2)} \int_a^b (\theta - t_0) g_k^2(\theta) \, d\alpha(\theta) \quad \text{with } |\eta| < 1.$$

Here $g_k(\theta)$ is a kth degree polynomial. Hence, if $t_0 = a$, the Gauss-Radau rule yields a lower bound for (16), and if $t_0 = b$, the rule yields an upper bound.

The Gauss-Radau rule is computed as follows. Let (cf. (11))

$$\overline{J}_{k+1} = \begin{bmatrix} \alpha_0 & \beta_0 & & & & \\ \gamma_1 & \alpha_1 & \beta_1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \gamma_{k-1} & \alpha_{k-1} & \beta_{k-1} & \\ & & & \gamma_k & \bar{\alpha}_k \end{bmatrix}.$$

We wish to calculate $\bar{\alpha}_k$ so that t_0 is an eigenvalue of $\overline{J}_{k+1}$ and hence the eigenvalues and eigenvectors of $\overline{J}_{k+1}$ yield the Gauss-Radau rule [6]. To compute $\bar{\alpha}_k$ we solve

$$(J_k - t_0 I)\delta = -\beta_{k-1}\mathbf{e}_k,$$

where $\mathbf{e}_k = (0, 0, \ldots, 0, 1)^T$. Then $\bar{\alpha}_k = t_0 - \gamma_k \delta_k$.

5. Computational Results. In view of the equivalence of the modified Chebyshev algorithm and the Lanczos algorithm we would expect similar convergence properties of the eigenvalue calculations in practice. In general, this is indeed the case, but there are circumstances where the behavior of the two methods is quite different. Two points in particular require further investigation: 1) the sensitivity of the modified Chebyshev algorithm to the distribution, $\alpha(\lambda)$, compared to the Lanczos algorithm, and 2) the effect of the initial parameters of the iterative method on the modified Chebyshev algorithm. The first point is dependent on how well the modified moments we calculate codify the distribution, while the second point affects how well we calculate the modified moments.

Both of the points above affect the termination properties of the modified Chebyshev algorithm. Termination occurs when σ_{kk} in (10) is nonpositive, and this can occur either because of numerical stability problems or simply because there are no more orthogonal polynomials to generate (i.e., if the distribution has m points of increase then only $m + 1$ orthogonal polynomials can be generated).

Tables 5.1 and 5.2 show the results of two sets of experiments with the Chebyshev semi-iterative method. For each problem the same randomly generated, unit-length, starting vector and right-hand side were used. Convergence of the eigenvalue calculations occurred when μ, calculated from the extreme eigenvalues of J_k, was within a prescribed tolerance, ε, of the value of μ from the pervious iteration. Two strategies were used to determine when to change the parameters. The first was to continue with the CSI method with the initial parameters until near optimal eigenvalue estimates were available. In this case we would use $\varepsilon = 10^{-6}$ and, when convergence of μ was realized, would restart the CSI method from the current solution vector using the new parameters. No more eigenvalue calculations would be performed.

The second strategy was to select a decreasing sequence of tolerances, $\varepsilon_1, \varepsilon_2, \ldots,$ ε_s and allow up to s parameter changes. At step i, convergence of μ to within ε_i was required. Then, after changing parameters and restarting the CSI method, convergence of μ to within ε_{i+1} was required. In the following tests we used $s = 3$ and $\varepsilon_i = 10^{-2i}$.

The matrix resulting from the discretization of the two-dimensional Laplace equation was used for the results in Table 5.1. The values of δ and μ are the parameters used to start the iterations (calculated from estimates a and b). The

column labeled 'fixed' gives the number of iterations required to reduce the norm of the error to less than 0.5×10^{-4} when no parameter changes were made, i.e., the calculations were continued with the initial choice of parameters. The column labeled 'dynamic' shows the number of iterations required when the parameters were changed at some point during the execution. The column labeled 'change' indicates the iteration at which the parameters were changed. This stage occurs when the eigenvalue calculations either converge or terminate. Finally, the columns $\hat{\delta}$ and $\hat{\mu}$ indicate the parameters used to restart the process at the 'change' point.

TABLE 5.1

2D Laplace, 4096×4096 in 64×64 blocks. $\mu_{opt} \approx 0.998832$, $\delta_{opt} = 0.25$, $a_{opt} \approx 0.00467$, $b_{opt} \approx 7.99533$. † indicates that the optimal parameters were obtained to machine accuracy.

	a	b	δ	μ	fixed	dynamic	change	$\hat{\delta}$	$\hat{\mu}$
1	a_{opt}	b_{opt}	δ_{opt}	μ_{opt}	219	237	30	0.248307	0.97872291
							122	0.250054	0.99861292
							176	†	†
2	0.1	7.9	δ_{opt}	0.975	1018	240	30	0.248307	0.97872291
							112	0.250031	0.99866113
							152	†	†
3	0.0	8.0	δ_{opt}	1.0	∞	234	30	0.248307	0.97872291
							126	0.249988	0.99695640
							202	†	†

For this problem, the optimal Chebyshev semi-iterative method converges in 219 iterations. The same (optimal) parameters are obtained from the eigenvalue calculations at iteration 176, after changing parameters at iterations 30 and 122. The desired solution vector is reached at iteration 237. The parameter change at iteration 30 replaces the optimal initial parameters with the values of $\hat{\delta}$ and $\hat{\mu}$ shown. This, of course, slows down convergence of the CSI method until the parameters are improved at iteration 122 and we finally arrive at the optimal parameters at iteration 176.

Next (entry 2) the initial parameters are chosen so that $\mu = 0.975$ and the parameters are changed at iterations 30, 112 and 152. If the parameters were not changed throughout the iterations, then over one thousand iterations would be required to solve the linear system. The dynamic Chebyshev method requires only 240 iterations.

Next, $a = 0$, so that $\omega_k = 2$ for all $k > 1$, and the CSI method neither converges nor diverges. The eigenvalue calculations again perform well and convergence of the CSI method is reached in only 234 iterations.

The next series of experiments use a tridiagonal matrix associated with a discrete set of orthogonal polynomials called Krawtchouk polynomials. The $(n+1) \times (n+1)$

Krawtchouk matrix has the following form:

$$
K = \begin{bmatrix}
\alpha_1 & p & & & & & \\
q/n & \alpha_2 & (n-1)p/n & & & & \\
& 2q/n & \alpha_3 & (n-2)p/n & & & \\
& & \cdot & \cdot & \cdot & \cdot & \\
& & & \cdot & \cdot & \cdot & \\
& & & & \cdot & \alpha_n & p/n \\
& & & & & q & \alpha_{n+1}
\end{bmatrix},
$$

where α_k is chosen to make the row sums equal one. This matrix has the property that for $p > 0$, $q = 1-p$, the $(n+1)$ eigenvalues of K are $\lambda_j = j/n$, for $j = 0, \ldots, n$.

The tests used the 256×256 shifted matrix $\hat{K} = K + \frac{1}{18}I$ which has $\mu_{\text{opt}} = 0.90$, $\delta_{\text{opt}} = 1.8$. Iterations were continued until the norm of the error was less than 0.5×10^{-8}. The extreme eigenvalues are $a = \frac{1}{18} \approx 0.0555556$ and $b = 1 + a$. The eigenvalue calculations were performed with $s = 1$ and $\varepsilon_1 = 10^{-2}$.

TABLE 5.2

Krawtchouk matrix, 256×256, $\mu_{\text{opt}} = 0.90$,
$\delta_{\text{opt}} = 1.8$, $a_{\text{opt}} = 1/18$, $b_{\text{opt}} = 1 + a_{\text{opt}}$.

	a	b	δ	μ	fixed	dynamic	change	$\hat{\delta}$	$\hat{\mu}$
1	a_{opt}	b_{opt}	δ_{opt}	μ_{opt}	43	48	12	1.79	0.8716981
2	0.01	1.1	1.80180	0.981982	104	58	12	1.79	0.8716981
3	0.06	1.0	1.88679	0.886792	> 256	56	12	1.79	0.8716981

The CSI method with optimal parameters required 43 iterations to meet the requested tolerance. In each of the three tests shown in Table 5.2 the eigenvalue calculations converged in 12 steps. The optimal parameters used in the first row are changed by the method, resulting in a penalty of five iterations. However, the second and third examples clearly show the value of the change where the parameters obtained were sufficiently accurate to provide large savings in the number of iterations. In the third example, the error for the fixed iterations is as large as $O(10^{-3})$ at iteration 256.

6. Summary. We have shown how to apply the method of modified moments to obtain an algorithm for computing the optimal parameters for the preconditioned Chebyshev semi-iterative method. Numerical results indicate that the algorithm performs well in practice. More extensive numerical tests will be reported at a later date.

We have also shown how to use modified moments to obtain sharp lower and upper bounds for certain measures of the error in the approximate solution vectors.

Variations of the material in this paper can be found in [8].

Computer Science Department
Stanford University
Stanford, California 94305-2095
E-mail: golub@na-net.stanford.edu
             kent@na-net.stanford.edu

1. P. CONCUS, G. H. GOLUB & D. P. O'LEARY, "A generalized conjugate gradient method for the numerical solution of elliptic partial differential equations," in *Studies in Numerical Analysis* (G. H. Golub, ed.), MAA Studies in Math., Vol. 24, 1985, pp. 178–198.

2. G. G. DAHLQUIST, S. C. EISENSTAT & G. H. GOLUB, "Bounds for the error of linear systems of equations using the theory of moments," *J. Math. Anal. Appl.*, v. 37, 1972, pp. 151–166.

3. G. G. DAHLQUIST, G. H. GOLUB & S. G. NASH, "Bounds for the error in linear systems," in *Proc. of the Workshop on Semi-Infinite Programming* (R. Hettich, ed.), Springer-Verlag, Berlin and New York, 1978, pp. 154–172.

4. W. GAUTSCHI, "On generating orthogonal polynomials," *SIAM J. Sci. Statist. Comput.*, v. 3, 1982, pp. 289–317.

5. W. GAUTSCHI, "Questions of numerical condition related to polynomials," in *Studies in Numerical Analysis* (G. H. Golub, ed.), MAA Studies in Math., Vol. 24, 1985, pp. 140–177.

6. G. H. GOLUB, "Some modified matrix eigenvalue problem," *SIAM Rev.*, v. 15, 1973, pp. 318–334.

7. G. H. GOLUB, *Error Bounds for Iterative Methods*, NA-85-34, Dept. of Computer Science, Stanford University, 1985.

8. G. H. GOLUB & M. D. KENT, *Estimates of Eigenvalues for Iterative Methods*, NA-87-02, Dept. of Computer Science, Stanford University, 1987.

9. G. H. GOLUB & R. S. VARGA, "Chebyshev semi-iterative methods, successive overrelaxation methods, and second order Richardson iterative methods," *Numer. Math.*, v. 3, 1961, pp. 147–168.

10. C. LANCZOS, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *J. Res. Nat. Bur. Standards*, v. 45, 1950, pp. 255–282.

11. R. A. SACK & A. F. DONOVAN, "An algorithm for Gaussian quadrature given modified moments," *Numer. Math.*, v. 18, 1971–72, pp. 465–478.

12. J. C. WHEELER, "Modified moments and Gaussian quadrature," *Rocky Mountain J. Math.*, v. 4, 1974, pp. 287–296.

13. J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.