ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 4, Number 2, Spring 1974

BOUNDS FOR MATRIX MOMENTS

GENE H. GOLUB*

1. **Introduction.** Let A be a real symmetric positive definite $n \times n$ matrix with

$$Au_i = \lambda_i u_i, \quad (i = 1, 2, \cdots, n)$$

${u_i}^T u_j = \delta_{ij}$, and $0 < \lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_n$. Let r_0 be an arbitrary vector and consider the *Krylov sequence*

$$r_{i+1} = Ar_i, \text{ for } i = 0, 1, \cdots, k-1,$$

so that

$$r_i = A^i r_0 \quad (i = 0, 1, \cdots, k).$$

Let

$$\begin{aligned} \mu_{p,q} = {r_p}^T r_q &= (A^p r_0)^T A^q r_0 \\ &= {r_0}^T A^{p+q} r_0 \\ &\equiv \mu_{p+q}. \end{aligned}$$

Thus if $r_0 = \sum_{i=1}^{n} \alpha_i u_i$,

$$\mu_m = \sum_{i=1}^{m} {\alpha_i}^2 {\lambda_i}^m \equiv \int \lambda^m \, d\alpha(\lambda) \quad (m = 0, 1, \cdots, 2k)$$

where $\alpha(\lambda) = 0$ for $\lambda \leqq \lambda_1$,

$$\begin{aligned} &= {\alpha_1}^2 + \cdots + {\alpha_t}^2 \quad \lambda_t < \lambda \leqq \lambda_{t+1}, \\ &= {\alpha_1}^2 + \cdots + {\alpha_n}^2 \quad \lambda_n < \lambda. \end{aligned}$$

Thus, $\{\mu_m\}_{m=1}^{2k}$ are a set of moments associated with the distribution function $\alpha(\lambda)$.

In certain applications (cf. [1]) we are interested in determining bounds for μ_s where s is a positive integer greater than $2k$ or a negative integer. We shall construct algorithms for computing bounds on μ_s where we have an upper bound on the largest eigenvalue and a positive lower bound on the smallest eigenvalue, e.g.,

Received by the editors February 8, 1973.
*This research was in part supported by the AEC and the NSF.
Dedicated to the memory of Professor C. P. Welter (1914-1972).

$$0 < a \leqq \lambda_i \leqq b \quad i = 1, \cdots, n.$$

Note that when the eigenvalues are known precisely linear programming may be used for determining upper and lower bounds on μ_s (cf. [1]).

2. **Construction of bounds.** Suppose we are given $\{\mu_i\}_{i=0}^{2k}$ and a function $\varphi(\lambda)$, $(a \leqq \lambda \leqq b)$ and we wish to determine (L, U) so that

$$L \leqq \int_a^b \varphi(\lambda)\, d\alpha(\lambda) \leqq U.$$

We can determine a Gauss-Radau quadrature rule so that

$$\mu_r = \sum_{i=0}^{k} A_i t_i^r \quad \text{for } r = 0, 1, \cdots, 2k$$

where $\{A_i\}_{i=0}^{k}$ and $\{t_i\}_{i=1}^{k}$ are unknown and t_0 is specified. Then

$$\int_a^b \varphi(\lambda)\, d\alpha(\lambda) = \sum_{i=0}^{k} A_i \varphi(t_i) + R[\varphi]$$

where

$$R[\varphi] = \frac{\varphi^{(2k+1)}(\eta)}{(2k+1)!} \int_a^b (\lambda - t_0) \left[\prod_{i=1}^{k} (\lambda - t_i) \right]^2 d\alpha(\lambda), \; a < \eta < b.$$

Thus if $\varphi(\lambda) = \lambda^s$,

$$R[\lambda^s] = \binom{s}{2k+1} \eta^{s-(2k+1)} \times \int_a^b (\lambda - t_0) \left[\prod_{i=1}^{k} (\lambda - t_i) \right]^2 d\alpha(\lambda), \; a < \eta < b.$$

Hence, if $s > 2k$ and $t_0 = a$, then $R[\lambda^s] \geqq 0$ and hence the Gauss-Radau rule yields an upper bound and if $t_0 = b$, a lower bound. However if $s < 0$ and $t_0 = a$, then $R[\lambda^s] \leqq 0$ and hence the Gauss-Radau rule yields an upper bound and if $t_0 = b$, a lower bound. It can be shown that these bounds are attainable (cf. [2]).

Unfortunately, using the moments for computing the quadrature rules is a very ill-conditioned numerical problem [3]. We can avoid this difficulty by working with orthogonal polynomials which are defined by the distribution function $\alpha(\lambda)$.

3. **The Lanczos algorithm.** Associated with the distribution function $\alpha(\lambda)$, there is a set of orthogonal polynomials $\{p_j(\lambda)\}$ such that

$$\int_a^b p_m(\lambda)p_\ell(\lambda)\,d\alpha(\lambda) = 0 \text{ when } m \neq \ell.$$

It is well known that these polynomials satisfy the relationship

$$p_{j+1}(\lambda) = (\xi_{j+1} - \lambda)p_j(\lambda) - \eta_j^2 p_{j-1}(\lambda) \tag{3.1}$$

with $p_{-1}(\lambda) = 0$, $p_0(\lambda) = 1$. The zeros of $p_k(\lambda)$ are the nodes of the Gauss quadrature rule associated with $\alpha(\lambda)$ and $\{\mu_j\}_{j=0}^{2k-1}$; e.g., $p_k(t_i) = 0$ $(i = 1, 2, \cdots, k)$, $\int_a^b \lambda^r\,d\alpha(\lambda) = \sum_{i=1}^k A_i t_i^r$ $(i = 0, 1, \cdots, 2k-1)$. The coefficients of (3.1) can be calculated directly from the moments but this is also a numerically unstable process.

The coefficients $\{\xi_j\}_{j=1}^k$, $\{\eta_j^2\}_{j=1}^{k-1}$ can be computed directly using the Lanczos algorithm [4]. We generate a sequence of vectors $\{z_j\}_{j=0}^{k+1}$ such that

$$z_i^T z_j = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j. \end{cases}$$

Let $z_0 = r_0 \times (\|r_0\|)^{-1}$ $(\|r_0\| = (r_0^T r_0)^{1/2})$. Then for $j = 0, 1, \cdots, k$,

$$\begin{aligned} \xi_{j+1} &= z_j^T A z_j, \\ w_{j+1} &= Az_j - \xi_{j+1} z_j - \eta_j^2 z_{j-1}, \quad (\eta_0 = 0), \\ \eta_{j+1} &= \|w_{j+1}\| \\ z_{j+1} &= \eta_{j+1} \times w_{j+1}. \end{aligned}$$

For numerical stability, one must reorthogonalize z_{j+1} with respect to all the previous z_j's (cf. [5]). It is well known (cf. [6]) that the eigenvalues of the symmetric tridiagonal matrix $J_k = \{\eta_{j-1}, \xi_j, \eta_j\}$ are the roots of the polynomial $p_k(\lambda)$ and that the square of the first component of the orthonormalized eigenvectors is the associated weight of the quadrature rule when $\mu_0 = 1$. The eigenvalues of J_k and the first component of the eigenvectors can be efficiently and accurately computed by the QR method of Francis ([7]).

Let

$$\bar{J}_{k+1} = \begin{bmatrix} \xi_1 & \eta_1 & & & & & \\ \eta_1 & \xi_2 & \eta_2 & & & & \\ & \eta_2 & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \cdot & \cdot & \eta_k \\ & & & & & \eta_k & \bar{\xi}_{k+1} \end{bmatrix}$$

We wish to compute the element $\bar{\xi}_{k+1}$ so that $p_{k+1}(t_0) = 0$, and thus the eigenvalues and eigenvectors of $\bar{J}_{k+1}$ yield the Gauss-Radau rule. Now

$$0 = p_{k+1}(t_0) = (t_0 - \bar{\xi}_{k+1})p_k(t_0) - \eta_k{}^2 p_{k-1}(t_0),$$

and hence,

$$\bar{\xi}_{k+1} = t_0 - \eta_k{}^2 p_{k-1}(t_0)/p_k(t_0).$$

The quantity $\rho_{k-1} \equiv p_k(t_0)/p_{k-1}(t_0)$ can be easily computed. Defining $\rho_j = p_{j+1}(t_0)/p_j(t_0)$ $(j = 0, 1, \cdots, k-1)$, then $\rho_j = (t_0 - \xi_{j+1}) - \eta_j{}^2/\rho_{j-1}$ with $\rho_1 = (t_0 - \xi_1)$.

It is not necessary to compute the eigenvalues and eigenvectors of $\bar{J}_{k+1}$ to compute upper and lower bounds on μ_s. Let

$$\bar{J}_{k+1} = QTQ^T, QQ^T = I_{k+1},$$

where T is the diagonal matrix of eigenvalues of $\bar{J}_{k+1}$ and Q is the matrix of eigenvectors. The vector $w_0 = Q^T e_1$, $(e_1{}^T = (1, 0, \cdots, 0))$ consists of the first element of each eigenvector of $\bar{J}_{k+1}$. Hence

$$\sum_{i=0}^{k} A_i t_i{}^s = e_1{}^T \bar{J}_{k+1}^s e_1.$$

If $s = 2p$ (say), then $\sum_{i=0}^{k} A_i t_i{}^s = (e_1{}^T \bar{J}_{k+1}^p)(\bar{J}_{k+1}^p e_1)$. And if s is negative it is only necessary to repeatedly solve the tridiagonal system of equations $\bar{J}_{k+1}^p f = e_1$.

4. **A numerical example.** Let A be the tridiagonal matrix $A = \{-1, 2, -1\}$ so that $a = 2 - 2\cos(\pi/(n+1))$, $b = 2 + 2\cos(\pi/(n+1))$. The vector r_0 is a random vector with $\|r_0\| = 1.07188$. We wish to construct upper and lower bounds for $s = -2$ when $n = 25$. The exact

solution is $\mu_{-2} = 1.0$. Using the Lanczos algorithm we have the following bounds.

k	*lower bound*	*upper bound*
1	.36834	18.99795
2	.43465	9.62961
3	.46049	4.51138
4	.47367	2.76067
5	.49212	2.11169
23	.999820503726357	1.00000054223912
24	.999999999914178	1.00000000000003

References

1. G. Dahlquist, S. Eisenstat and G. H. Golub, *Bounds for the error of linear systems of equations using the theory of moments,* J. Math. Anal. Appl. 37 (1972), 151-166.

2. S. Karlin and W. J. Studden, *Tchebysheff Systems: With Applications in Analysis and Statistics,* Interscience Publishers, New York, 1966.

3. W. Gautschi, *Construction of Gauss-Christoffel quadrature formulas,* Math. Comp. 22 (1968), 251-270.

4. G. H. Golub, *Some uses of the Lanszos algorithm in numerical linear algebra,* Technical Report No. CS 302, Computer Science Dept., Stanford University, 1972.

5. G. H. Golub, R. Underwood and J. H. Wilkinson, *The Lanczos algorithm for the symmetric $Ax = \lambda Bx$ problem,* Technical Report No. CS 270, Computer Science Dept., Stanford University, 1972.

6. G. H. Golub and J. H. Welsch, *Calculation of Gauss quadrature rules,* Math. Comp. 23 (1969), 221-230.

7. J. G. F. Francis, *The QR transformation: a unitary analogue to the LR transformation.* I, II, Computer J. 4 (1961/62), 265-271; 332-345.

Stanford University, Stanford, California 94503