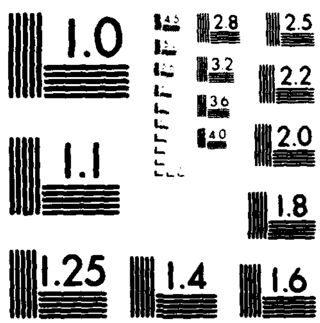
AD-A115 431   YALE UNIV NEW HAVEN CT DEPT OF COMPUTER SCIENCE   F/6 12/1
OPTIMAL BLOCK DIAGONAL SCALING OF BLOCK 2-CYCLIC MATRICES.(U)
1976   S C EISENSTAT, J W LEWIS, M H SCHULTZ   N00014-76-C-0277
UNCLASSIFIED   RR-204   NL

1 OF 1

END
DATE
FILMED
07-82
DTIC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS 1963 A

12

YALE UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE

(12)

Abstract

In this paper, we describe a class of optimal block diagonal scalings (preconditionings) of a symmetric positive definite block 2-cyclic matrix, generalizing a result of Forsythe and Straus [1] for (point) 2-cyclic matrices.

Optimal Block Diagonal Scaling of Block 2-cyclic Matrices

Stanley C. Eisenstat[1], John W. Lewis[2], and Martin H. Schultz[1]

Research Report 204

1 Department of Computer Science and Research Center for Scientific Computation, Yale University, New Haven, CT 06520.

2 General Electric Corporate Research and Development, Schenectady, NY 12345.

This research was supported in part by ONR Grant N00014-76-C-0277.

1. Introduction

Let A be an $n \times n$ symmetric positive definite matrix and let Δ be a class of $n \times n$ nonsingular matrices. In this paper, we consider the problem of finding an $E \in \Delta$ which minimizes the condition number

$$K(E^tAE) = \frac{\alpha_1(E^tAE)}{\alpha_n(E^tAE)}$$

of the scaled (preconditioned) matrix E^tAE (here $\alpha_1(E^tAE)$ and $\alpha_n(E^tAE)$ denote respectively the largest and smallest eigenvalues of E^tAE). In particular, we describe a class of optimal block diagonal scalings (preconditionings) of a block 2-cylic matrix, generalizing a result of Forsythe and Strauss [1] that a (point) 2-cyclic matrix is optimally conditioned with respect to the class of positive diagonal matrices if all the elements of the diagonal are equal.

2. Optimal 2 × 2 Block Diagonal Scaling

Let A be an $n \times n$ symmetric positive definite matrix. Then, for any $0 < m < n$, A can be written as a 2×2 block matrix

$$A = \begin{vmatrix} D_1 & C^t \\ C & D_2 \end{vmatrix}$$

where D_1 is $m \times m$, D_2 is $n\text{-}m \times n\text{-}m$, and C is $n\text{-}m \times m$. Moreover, the 2×2 block diagonal matrix

$$D = \begin{vmatrix} D_1 & 0 \\ 0 & D_2 \end{vmatrix}$$

is also symmetric positive definite and hence nonsingular.

Let $\Delta(m,n-m)$ denote the set of <u>all</u> nonsingular 2×2 block diagonal matrices of the form

$$E = \begin{vmatrix} E_1 & 0 \\ 0 & E_2 \end{vmatrix}$$

where E_1 is $m \times m$ and E_2 is $n-m \times n-m$. The following result describes a class of optimal scalings $\tilde{E} \in \Delta(m,n-m)$ for A.

<u>Theorem 1</u>: If $\tilde{E} \in \Delta(m,n-m)$ and $\tilde{E}\tilde{E}^t = cD^{-1}$ for some positive constant c, then

$$K(\tilde{E}^t A\tilde{E}) = \min \{K(E^tAE) \mid E \in \Delta(m,n-m)\} \quad .$$

<u>Proof</u>:

Let $S = \begin{vmatrix} I_m & 0 \\ 0 & -I_{n-m} \end{vmatrix}$ where I_k denotes the $k \times k$ identity matrix. Then

$$2D-A = \begin{vmatrix} D_1 & -C^t \\ -C & D_2 \end{vmatrix} = \begin{vmatrix} I_m & 0 \\ 0 & -I_{n-m} \end{vmatrix} \begin{vmatrix} D_1 & C^t \\ C & D_2 \end{vmatrix} \begin{vmatrix} I_m & 0 \\ 0 & -I_{n-m} \end{vmatrix} = S^tAS. \tag{1}$$

Let $E \in \Delta(m,n-m)$. Since E is nonsingular, E^tAE is symmetric positive definite so that, by the Rayleigh principle,

$$y^tE^tAEy \le \alpha_1(E^tAE)\ y^ty \qquad \text{for all } y \in R^n \tag{2}$$

and

$$y^tE^tAEy \ge \alpha_n(E^tAE)\ y^ty \qquad \text{for all } y \in R^n. \tag{3}$$

By (1), the fact that SE = ES, and (2),

$$y^tE^t(2D-A)Ey = y^tE^tS^tASEy = y^tS^tE^tAESy \le \alpha_1(E^tAE)\ y^tS^tSy$$

so that, since $S^tS = I_n$,

$$y^tE^t(2D-A)Ey \le \alpha_1(E^tAE)\ y^ty \qquad \text{for all } y \in R^n. \tag{4}$$

Similarly, using (3) instead of (2),

$$y^tE^t(2D-A)Ey \;\geq\; \alpha_n(E^tAE)\; y^ty \qquad \text{for all } y \in R^n. \tag{5}$$

Adding $\alpha_n(E^tAE)$ times (2) to $-\alpha_1(E^tAE)$ times (5),

$$\{\alpha_1(E^tAE) + \alpha_n(E^tAE)\}\; y^tE^tAEy \;-\; 2\alpha_1(E^tAE)\; y^tE^tDEy \;\leq\; 0$$

so that

$$y^tE^tAEy \;\leq\; \frac{2\alpha_1(E^tAE)}{\alpha_1(E^tAE)+\alpha_n(E^tAE)}\; y^tE^tDEy \qquad \text{for all } y \in R^n. \tag{6}$$

Similarly, adding $\alpha_1(E^tAE)$ times (3) to $-\alpha_n(E^tAE)$ times (4),

$$\{\alpha_1(E^tAE) + \alpha_n(E^tAE)\}\; y^tE^tAEy \;-\; 2\alpha_n(E^tAE)\; y^tE^tDEy \;\geq\; 0$$

so that

$$y^tE^tAEy \;\geq\; \frac{2\alpha_n(E^tAE)}{\alpha_1(E^tAE)+\alpha_n(E^tAE)}\; y^tE^tDEy \qquad \text{for all } y \in R^n. \tag{7}$$

Now let $\tilde{E} \in \Delta(m,n-m)$ with $\tilde{E}\tilde{E}^t = cD^{-1}$ for some positive constant c. Since $\tilde{E}$ is nonsingular, taking $y = E^{-1}\tilde{E}x$ in (6),

$$x^t\tilde{E}^tA\tilde{E}x \;\leq\; \frac{2\alpha_1(E^tAE)}{\alpha_1(E^tAE)+\alpha_n(E^tAE)}\; x^t\tilde{E}^tD\tilde{E}x \qquad \text{for all } x \in R^n.$$

But

$$\tilde{E}^tD\tilde{E} \;=\; \tilde{E}^t\,\{c\tilde{E}^{-t}\tilde{E}^{-1}\}\,\tilde{E} \;=\; cI_n$$

so that

$$x^t\tilde{E}^tA\tilde{E}x \;\leq\; \frac{2c\alpha_1(E^tAE)}{\alpha_1(E^tAE)+\alpha_n(E^tAE)}\; x^tx \qquad \text{for all } x \in R^n. \tag{8}$$

Similarly, using (7) instead of (6),

$$x^t\tilde{E}^t A\tilde{E}x \geq \frac{2c\alpha_n(E^tAE)}{\alpha_1(E^tAE)+\alpha_n(E^tAE)}\, x^t x \qquad \text{for all } x \in R^n. \tag{9}$$

Therefore, by the Rayleigh principle,

$$\frac{2c\alpha_n(E^tAE)}{\alpha_1(E^tAE)+\alpha_n(E^tAE)} \leq \alpha_n(\tilde{E}^tA\tilde{E}) \leq \alpha_1(\tilde{E}^tA\tilde{E}) \leq \frac{2c\alpha_1(E^tAE)}{\alpha_1(E^tAE)+\alpha_n(E^tAE)}$$

so that

$$K(\tilde{E}^tA\tilde{E}) = \frac{\alpha_1(\tilde{E}^tA\tilde{E})}{\alpha_n(\tilde{E}^tA\tilde{E})}$$

$$\leq \frac{2c\alpha_1(E^tAE)}{\alpha_1(E^tAE)+\alpha_n(E^tAE)} \Big/ \frac{2c\alpha_n(E^tAE)}{\alpha_1(E^tAE)+\alpha_n(E^tAE)}$$

$$= \frac{\alpha_1(E^tAE)}{\alpha_n(E^tAE)} = K(E^tAE) \quad .$$

QED

Note that if $\tilde{E} \in \Delta(m,n-m)$ satisfies the conditions of Theorem 1, then

$$\tilde{E}^tA\tilde{E} = \begin{vmatrix} cI_m & \tilde{E}^tC^t\tilde{E} \\ \tilde{E}^tC\tilde{E} & cI_{n-m} \end{vmatrix}$$

and the diagonal blocks of $\tilde{E}^tA\tilde{E}$ are common multiples of the identity matrix. Therefore:

<u>Corollary 2</u>: If $D = cI_n$ for some positive constant c, then A is optimally conditioned with respect to $\Delta(m,n-m)$.

However, not all optimal scalings $\tilde{E} \in \Delta(m,n-m)$ of A satisfy Theorem 1. For example, the matrix

$$A = \begin{vmatrix} \rho & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix}$$

is optimally conditioned with respect to $\Delta(2,1)$ for $1 \le \rho \le 3$ (cf. Forsythe and Strauss [1]).

3. Optimal Block Diagonal Scaling of Block 2-Cyclic Matrices

Let A be a symmetric positive definite block 2-cyclic matrix. Then, without loss of generality (by symmetrically permuting the rows and columns), A can be written as a $t \times t$ block matrix of the form

$$A = \left|\begin{array}{cccc|cccc} D_1 & 0 & \cdots & 0 & & & & \\ 0 & D_2 & \cdots & 0 & & & C^t & \\ \vdots & \vdots & & \vdots & & & & \\ 0 & 0 & \cdots & D_r & & & & \\ & & & & D_{r+1} & 0 & \cdots & 0 \\ & & & & 0 & D_{r+2} & \cdots & 0 \\ & C & & & \vdots & \vdots & & \vdots \\ & & & & 0 & 0 & \cdots & D_t \end{array}\right|$$

where D_i is $n_i \times n_i$ and C is $m \times n{-}m$ with $m = n_1 + n_2 + \cdots + n_r$ and $n = n_1 + n_2 + \cdots + n_t$. Moreover, the block diagonal matrix

$$D = \begin{vmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & D_t \end{vmatrix}$$

is again symmetric positive definite and hence nonsingular.

Extending the previous definition, let $\Delta(n_1,n_2,\ldots,n_t)$ denote the set of <u>all</u> nonsingular $t \times t$ block diagonal matrices of the form

$$\begin{vmatrix} E_1 & 0 & \cdots & 0 \\ 0 & E_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & E_t \end{vmatrix}$$

where E_i is $n_i \times n_i$. The following result characterizes a class of optimal scalings $\tilde{E} \in \Delta(n_1,n_2,\ldots,n_t)$ for A.

<u>Corollary 3</u>: If $\tilde{E} \in \Delta(n_1,n_2,\ldots,n_t)$ and $\tilde{E}\tilde{E}^t = cD^{-1}$ for some positive constant c, then

$$\begin{aligned} K(\tilde{E}^t A\tilde{E}) &= \min \{K(E^tAE) \mid E \in \Delta(n_1,n_2,\ldots,n_t)\} \\ &= \min \{K(E^tAE) \mid E \in \Delta(m,n-m)\} \quad . \end{aligned}$$

Therefore, if $D = cI_n$ for some positive constant c, then A is optimally conditioned with respect to both $\Delta(n_1,n_2,\ldots,n_t)$ and $\Delta(m,n-m)$.

<u>Proof</u>:

Since $\tilde{E} \in \Delta(n_1,n_2,\ldots,n_t) \subset \Delta(m,n-m)$, the result follows immediately from Theorem 1 and Corollary 2.

QED

Forsythe and Strauss [1] proved that a (point) 2-cyclic matrix is optimally conditioned with respect to the class of all positive diagonal matrices if the diagonal entries are all equal.[1] This is the special case

---

[1] An alternate proof is given in Young [2].

$n_i = 1$ of Corollary 3. However, as the Corollary also shows, such a matrix is optimally conditioned with respect to the much larger class $\Delta(m,n-m)$ of 2×2 block diagonal matrices.

References

[1] G. E. Forsythe and E. G. Strauss. On best conditioned matrices. Proceedings of the AMS 6:340-345, 1955.

[2] David M. Young. Iterative Solution of Large Linear Systems. Academic Press, 1971.

7-8

DTIC