Block matrices and symmetric perturbations

Alicja Smoktunowicz

Faculty of Mathematics and Information Science, Warsaw University of Technology, Pl. Politechniki 1, Warsaw 00-661, Poland

Received 11 July 2007; accepted 13 March 2008
Available online 28 April 2008
Submitted by L. Cvetković

Dedicated to Richard Varga on the occasion of his 80th birthday.

Abstract

We prove that if $A = [A_{ij}] \in \mathbb{R}^{N,N}$ is a block symmetric matrix and $y$ is a solution of a nearby linear system $(A + E)y = b$, then there exists $F = F^T$ such that $y$ solves a nearby symmetric system $(A + F)y = b$, if $A$ is symmetric positive definite or the matricial norm $\mu(A) = (\|A_{ij}\|_2)$ is diagonally dominant. Our blockwise analysis extends existing normwise and componentwise results on preserving symmetric perturbations (cf. [J.R. Bunch, J.W. Demmel, Ch. F. Van Loan, The strong stability of algorithms for solving symmetric linear systems, SIAM J. Matrix Anal. Appl. 10 (4) (1989) 494–499; D. Herceg, N. Krejić, On the strong componentwise stability and $H$-matrices, Demonstratio Mathematica 30 (2) (1997) 373–378; A. Smoktunowicz, A note on the strong componentwise stability of algorithms for solving symmetric linear systems, Demonstratio Mathematica 28 (2) (1995) 443–448]).

© 2008 Elsevier Inc. All rights reserved.

Keywords: Matrical norm; Block matrix; Symmetric perturbations

1. Introduction

In many practical applications we need to solve a linear system of equations $Ax = b$, where $A \in \mathbb{R}^{N,N}$ is nonsingular and has special block structure. We assume that the matrix $A$ is partitioned into $s \times s$ blocks

E-mail address: smok@mini.pw.edu.pl

0024-3795/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2008.03.013
where \( A_{i,j} \in \mathbb{R}^{n_i \times n_j} \) is referred to as the \((i, j)\) block of \( A \), \( \{n_1, \ldots, n_s\} \) is a given set of positive integers, \( n_1 + \cdots + n_s = N \).

Very often, the block matrices \( A_{i,j} \) are sparse and many of them are zero. Numerical algorithms should exploit the structure of the matrix \( A \). Without loss of generality we restrict our attention to the spectral matrix norm (2-norm) and the second vector norm (length of \( x \)). It is well-known that \( \| A \|_2^2 = \rho(A^T A) \), where \( \rho(A) = \max\{ |\lambda| : \lambda \in \text{spect}(A) \} \) denotes the spectral radius of \( A \).

The aim of this paper is to answer the following question:

Assume that \( A \in \mathbb{R}^{N \times N} \) is symmetric and \( y \) is a solution of a nearby linear system \((A + E)y = b\), does it follows that \( y \) solves a nearby symmetric system \((A + F)y = b\)?

Bunch, Demmel, and Van Loan (cf. [1]) show that there is \( F = F^T \) such that \( \| F \|_2 \leq \| E \|_2 \). Smoktunowicz (cf. [10]) proves that if \( |e_{ij}| \leq \epsilon |a_{ij}| \) for all \( i, j = 1, 2, \ldots, N \) and \( A \) is symmetric and diagonally dominant or symmetric positive definite, then there exists a symmetric matrix \( F \) such that \((A + F)y = b \) with \( |f_{ij}| \leq \epsilon c |a_{ij}| \), where \( c \) is a modest constant depending only on \( N \). Herceg and Krejić (cf. [6]) get similar result for \( H \)-matrices. Elsner (cf. [4]) observed that we can take \( c = 2 \) instead of \( c = 3 \) in [10], if \( A \) is diagonally dominant.

We extend existing results on preserving symmetric perturbations to block matrices. We assume that \( \| E_{i,j} \|_2 \leq \epsilon \| A_{i,j} \|_2 \) for all \( i, j = 1, 2, \ldots, s \). We prove that symmetry is preserved in perturbations if \( A \) is symmetric positive definite or \( A = [A_{i,j}] \) is diagonally dominant in an appropriate block sense. A general matrix \( A = [A_{i,j}] \) partitioned as in (1) is block diagonally dominant with respect to the spectral norm if, for all \( i = 1, 2, \ldots, s \),

\[
\| A_{i,i} \|_2 \geq \sum_{j=1, j \neq i}^{s} \| A_{i,j} \|_2.
\]

Similarly, \( A \) is a block \( H \)-matrix if there is a positive vector \( w = [w_1, w_2, \ldots, w_s]^T \) such that for all \( i = 1, 2, \ldots, s \),

\[
\| A_{i,i} \|_2 w_i \geq \sum_{j=1, j \neq i}^{s} \| A_{i,j} \|_2 w_j.
\]

In componentwise case, i.e. for \( s = n \), we call \( A \) satisfying (2) or (3) pointwise diagonally dominant or a pointwise \( H \)-matrix, respectively.

Notice, that if all matrices \( A_{i,i} \) in (1) are nonsingular, then

\[
\| A_{i,i} \|_2 \geq \frac{1}{\| A_{i,i}^{-1} \|_2}.
\]

We see that if, for all \( i \),

\[
\frac{1}{\| A_{i,i}^{-1} \|_2} \geq \sum_{j=1, j \neq i}^{s} \| A_{i,j} \|_2,
\]

then \( A \) satisfies (2), while the opposite is not true.
Our definitions (2) and (3) are different from those considered by Feingold and Varga (cf. [5]) and Cvetković and Kostić (cf. [2]). We do not assume that the diagonal blocks $A_{ii}$ are nonsingular. Moreover, we use the spectral norm instead of the infinity norm. We show that the properties (2) and (3) are quite natural when we introduce *matricial norms*.

2. Matricial norms

For a general matrix $A = [A_{ij}]$ partitioned as in (1) we define a *matricial norm* $\mu(A)$ as follows (cf. [3,8])

$$
\mu(A) = \begin{bmatrix}
\|A_{11}\|_2 & \|A_{12}\|_2 & \cdots & \|A_{1s}\|_2 \\
\|A_{21}\|_2 & \|A_{22}\|_2 & \cdots & \|A_{2s}\|_2 \\
\vdots & \vdots & \ddots & \vdots \\
\|A_{s1}\|_2 & \|A_{s2}\|_2 & \cdots & \|A_{ss}\|_2 
\end{bmatrix}.
$$

(6)

Some important cases of matricial norms are: $\mu(A) = |A|$ for $s = n$ and $\mu(A) = \|A\|_2$ for $s = 1$. The matrix $|A|$ is the matrix whose elements are $|a_{i,j}|$ and we write $|A| \leq |B|$ to mean that inequalities between matrices hold componentwise. The vector $x \in \mathbb{R}^n$ is partitioned conformally: $x = [x_1^T, \ldots, x_s^T]^T$ where $x_i \in \mathbb{R}^{n_i}$ and $\mu(x) = [\|x_1\|_2, \ldots, \|x_s\|_2]^T$.

We review the main facts on the matricial norms (cf. [3,8,10]).

**Theorem 1.** Let $\mu$ be the matricial norm as defined in (6). For matrices $A = [A_{ij}]$ and $B = [B_{ij}]$ partitioned as in (1) and (6) we have

(a) $\mu(cA) = |c| \mu(A)$ for all $c \in \mathbb{R}$,
(b) $\mu(A + B) \leq \mu(A) + \mu(B)$, $\mu(AB) \leq \mu(A)\mu(B)$,
(c) $\mu(x + y) \leq \mu(x) + \mu(y)$ for all $x, y \in \mathbb{R}^N$,
(d) $\mu(Ax) \leq \mu(A)\mu(x)$ for all $x \in \mathbb{R}^N$,
(e) $\rho(A) \leq \rho(\mu(A))$,
(f) $\|A\|_2 \leq \|\mu(A)\|_2$.

The property (e) is a generalization of the Perron–Frobenius inequality and was first proved by Ostrowski (cf. [8]), see also [3]. It is useful in the study of convergence of iterative methods for solving large linear system of equations and for eigenvalue inclusion sets, see Varga (cf. [13, Section 6]), where the infinity norm was used instead of the spectral norm (2-norm).

Notice that the property (2) means that $\mu(A)$ is pointwise diagonally dominant and (3) reads as $\mu(A)$ is a pointwise $H$-matrix.

3. Symmetric perturbations

In practice it is very important to exploit the block structure of $A$. If $y$ solves a problem that is close to the original one, i.e. $(A + E)y = b$ where $\mu(E) \leq \epsilon \mu(A)$, then $A + E$ has the same block structure as $A$: $A_{ij} = 0$ implies that $E_{ij} = 0$.

The blockwise backward error can be computed exactly, from the formulae of Rigal and Gaches (cf. [9]).
Theorem 2 (Rigal and Gaches). The blockwise backward error

\[ \eta_\mu(y) = \min \{ \epsilon : (A + E)y = b, \mu(E) \leq \epsilon \mu(A) \} \]  

is given by

\[ \eta_\mu(y) = \max_i \frac{\|r_i\|_2}{c_i} \]  

where \( r = [r_1^T, \ldots, r_s^T]^T = b - Ay, y = [y_1^T, \ldots, y_s^T]^T, r_i, y_i \in \mathbb{R}^{n_i}, \) and

\[ c_i = [\mu(A)\mu(y)]_i = \sum_{j=1}^s \|A_{ij}\|_2 \|y_j\|_2, \quad i = 1, \ldots, s. \]

Here \( \xi/0 \) is interpreted as zero if \( \xi = 0 \) and infinity otherwise.

Proof. It is obvious that the right-hand side of (8) is a lower bound for \( \eta_\mu(y) \). It remains to prove that there exists \( E = [E_{ij}] \) such that this lower bound is attained.

Without loss of generality we can assume that \( y_j \neq 0 \) for all \( j = 1, \ldots, s \).

Define \( E = [E_{ij}] \) as follows

\[ E_{ij} = \frac{\|A_{ij}\|_2}{c_i} \frac{r_i y_j^T}{\|y_j\|_2}. \]

It is easy to verify that \( Ey = r \) and

\[ \|E_{ij}\|_2 = \frac{\|A_{ij}\|_2 \|r_i\|_2}{c_i}, \]

so \( E = [E_{ij}] \) is the desired matrix. □

We can compute the blockwise relative error to evaluate numerical algorithms. We recommend using special techniques of iterative refinement for improving a computed solution \( \tilde{y} \) to a linear system \( Ax = b \) (cf. [11]). For a deeper discussion of practical error bounds and condition numbers for a linear system with respect to a special class of perturbations we refer the reader to [7, pp. 131–150].

If \( A = [A_{ij}] \) is symmetric then it is reasonable to have a numerical solution \( y \) being a solution of slightly perturbed symmetric system \((A + F)y = b\). We partly resolve this problem using blockwise approach. In our construction of a symmetric matrix \( F \) we use the following lemma [12].

Lemma 1. Let \( a, b \in \mathbb{R}^n \) with \( a \neq 0 \). Then the matrix

\[ S_n = S_n(a, b) = \frac{ab^T + ba^T - (b^T a)I_n}{a^T a} \]  

is the smallest symmetric matrix in the spectral norm such that \( b = S_n a \).

The spectral norm of \( S_n \) equals

\[ \|S_n\|_2 = \frac{\|b\|_2}{\|a\|_2}. \]  

Proof. We omit the proof. It is similar to the proof of Theorem 2.2 in [12].
Lemma 2. Assume $A = [A_{ij}] \in \mathbb{R}^{N,N}$ is symmetric and $(A + E)y = b$ with $\mu(E) \leq \epsilon \mu(A)$. Let $y = [y_1^T, \ldots, y_s^T]^T$, $y_i \in \mathbb{R}^{n_i}$ and
\[
\|y_1\|_2 \leq \|y_2\|_2 \leq \cdots \leq \|y_s\|_2.
\]
Then there exists $F = F^T$ such that $(A + F)y = b$ with $\|Fij\|_2 \leq \epsilon \|Aij\|_2$ for $i \neq j$ and
\[
\|Fii\|_2 \leq \epsilon \left( \|Aii\|_2 + 2 \sum_{j=1}^{i-1} \|A_{ij}\|_2 \right).
\]

Proof. We construct a symmetric $F$ such that $Fy = Ey$. Let $F_{1,1} = 0$ for $y_1 = 0$. If $y_1 \neq 0$ then we apply Lemma 1 for $a = y_1, b = E_{11}y_1$ and $n = n_1$. We take $F_{1,1} = S_{n_1}(y_1, E_{11}y_1)$. Then $F_{11}$ is symmetric, $F_{11}y_1 = E_{11}y_1$ and
\[
\|F_{11}\|_2 = \|E_{11}y_1\|_2 \leq \|E_1\|_2 \leq \epsilon \|A_{11}\|_2.
\]
Define the block off-diagonal of $F$ by $F_{ij} = F_{ji}^T = E_{ij}$ for $j > i$. We find $F_{ii} = F_{ii}^T$ such that
\[
F_{ii}y_i = E_{ii}y_i + \sum_{j=1}^{i-1} (E_{ij} - E_{ji}^T)y_j.
\]
If $y_i = 0$ set $F_{ii} = 0$. Then (14) holds trivially. Otherwise, using a formula
\[
y_j = \left( \frac{y_j y_i^T}{y_i^T y_i} \right) y_i,
\]
define $G_{ii} \in \mathbb{R}^{n_i, n_i}$ as follows
\[
G_{i,i} = E_{ii} + \sum_{j=1}^{i-1} (E_{ij} - E_{ji}^T) \left( \frac{y_j y_i^T}{y_i^T y_i} \right).
\]
Take $F_{i,i} = S_{n_i}(y_i, G_{ii}y_i)$. Then $F_{ii}$ is symmetric and $F_{ii}y_i = G_{ii}y_i$, so (15) holds. It is easy to check that
\[
\|F_{ii}\|_2 \leq \|G_{ii}\|_2 \leq \|E_{ii}\|_2 + \sum_{j=1}^{i-1} (\|E_{ij}\|_2 + \|E_{ji}^T\|_2) \frac{\|y_j\|_2}{\|y_i\|_2},
\]
which together with (12) yields (14). This completes the proof. □

We can now formulate our main results.

Theorem 3. Assume $A = [A_{ij}] \in \mathbb{R}^{N,N}$ is symmetric and $\mu(A)$ is diagonally dominant and $(A + E)y = b$ with $\mu(E) \leq \epsilon \mu(A)$. Then there exists $F = F^T$ such that $(A + F)y = b$ with $\mu(F) \leq 3\epsilon \mu(A)$. 

Proof. First, assume that (12) holds. Then, by Lemma 2 and diagonal dominance
\[
\| F_{ii} \|_2 \leq \epsilon \left( \| A_{ii} \|_2 + 2 \sum_{j=1}^{i-1} \| A_{ij} \|_2 \right) \leq 3\epsilon \| A_{ii} \|_2, \tag{17}
\]
which together with (13) gives the desired bound for \( \mu(F) \).

Now assume that \( y = [y_1^T, \ldots, y_s^T]^T \) satisfies
\[
\| y_{p_1} \|_2 \leq \| y_{p_2} \|_2 \leq \cdots \leq \| y_{p_s} \|_2 \tag{18}
\]
for some permutation \( \{p_1, p_2, \ldots, p_s\} \) of \( \{1, 2, \ldots, s\} \).

Let \( \tilde{y} = [y_{p_1}^T, \ldots, y_{p_s}^T]^T \). Clearly, there is a permutation matrix \( P = [P_{ij}] \) partitioned as in (1) such that \( \tilde{y} = Py \).

Let \( \tilde{A} = PAP^T, \tilde{E} = PEP^T, \tilde{b} = Pb \). Then \( \tilde{A} = [\tilde{A}_{ij}] \) with \( \tilde{A}_{ij} = A_{p_i p_j} \) for \( i, j = 1, \ldots, s \).

Clearly, \( \tilde{A} \) is symmetric diagonally dominant. Thus, we can apply Lemma 2 to \( (\tilde{A} + \tilde{E}) \tilde{y} = \tilde{b} \) with \( \mu(\tilde{E}) \leq 3\epsilon \mu(\tilde{A}) \). With \( F = P^T \tilde{F} P \), we find that \( F \) is symmetric and \( (A + F) y = b \) with \( \mu(F) \leq 3\epsilon \mu(A) \). This finishes the proof. \( \square \)

Theorem 4. Assume \( A = [A_{ij}] \in \mathbb{R}^{N,N} \) is symmetric and \( \mu(A) \) is an \( H \)-matrix and \( (A + E)y = b \) with \( \mu(E) \leq \epsilon \mu(A) \). Then there exists \( F = F^T \) such that \( (A + F)y = b \) with \( \mu(F) \leq 3\epsilon \mu(A) \).

Proof. Notice that if \( \mu(A) \) is an \( H \)-matrix, then, similarly to Cvetković and Kostić (cf. [2]), we can introduce a block diagonal matrix \( W \in \mathbb{R}^{N,N} \) such that
\[
W = \text{diag}(w_1 I_{n_1}, w_2 I_{n_2}, \ldots, w_s I_{n_s}), \tag{19}
\]
where \( I_{n_i} \) denotes the identity matrix of size \( n_i \), i.e. the size of \( A_{i,i} \) in (1). Then \( \mu(A) \) is an \( H \)-matrix if and only if \( AW \) is a block diagonal matrix in the sense of (2). We observe that \( WAW \) is symmetric and block diagonally dominant. We get
\[
\mu(WAW) = \mu(W) \mu(A) \mu(W), \mu(W) = \text{diag}(w_1, w_2, \ldots, w_s) \tag{20}
\]
and just apply Theorem 3 to \( WAW \). This completes the proof. \( \square \)

Now we focus our attention to symmetric positive definite matrices. The definiteness implies certain relations among the submatrices \( A_{ij} \).

Lemma 3. Let \( A = [A_{ij}] \in \mathbb{R}^{N,N} \) be a symmetric positive definite matrix, partitioned as in (1). Then
\[
\| A_{ij} \|_2^2 < \| A_{ii} \|_2 \| A_{jj} \|_2 \quad \text{for } i \neq j \tag{21}
\]
and
\[
\max_{i,j} \| A_{ij} \|_2 = \max_i \| A_{ii} \|_2. \tag{22}
\]

Proof. Let \( A \) be symmetric positive definite. Then \( x^T Ax > 0 \) for every nonzero vector \( x \in \mathbb{R}^N \).

Without loss of generality we can assume that \( i < j \).

It is known (cf. [7, p. 127]) that for every \( B \in \mathbb{R}^{m,n} \)
\[
\| B \|_2 = \max_{\| u \|_2 = \| v \|_2 = 1} | u^T B v |.
\]
Let \( u \in \mathbb{R}^n \) and \( v \in \mathbb{R}^n \) be such that \( \| u \|_2 = \| v \|_2 = 1 \) and \( |A_{ij}|_2 = |u^T A_{ij} v| \). For arbitrary \( t \in \mathbb{R} \) we define \( x = [x_1^T, \ldots, x_s^T]^T \) with \( x_k \in \mathbb{R}^n \), as follows
\[
x_k = \begin{cases} 
  tu & \text{for } k = i, \\
  v & \text{for } k = j, \\
  0 & \text{for } k \neq i, j.
\end{cases}
\]

Then \( 0 < x^T A x = t^2(u^T A_{ii} u) + 2t(u^T A_{ij} v) + (v^T A_{jj} v) \) for arbitrary \( t \), hence
\[
A = 4(u^T A_{ij} v)^2 - 4(u^T A_{ii} u)(v^T A_{jj} v) < 0.
\]

Since \( \| u \|_2 = \| v \|_2 = 1 \), we conclude that
\[
\| A_{ij} \|_2^2 = (u^T A_{ij} v)^2 < (u^T A_{ii} u)(v^T A_{jj} v) < \| A_{ii} \|_2 \| A_{jj} \|_2,
\]
which is our claim.

**Theorem 5.** If \( A = [A_{ij}] \in \mathbb{R}^{N,N} \) is symmetric positive definite and \( (A + E)y = b \) with \( \mu(E) \leq \epsilon \mu(A) \), then there exists \( F = F^T \) such that \( (A + F)y = b \) with \( \mu(F) \leq (2s - 1)\epsilon \mu(A) \).

**Proof.** Let \( A = [A_{ij}] \) be a symmetric positive definite matrix, partitioned as in (1). Define a positive diagonal matrix \( W \) as in (19) with
\[
w_i = \| A_{ii} \|_2^{-1}, \quad i = 1, \ldots, s.
\]
(23)

Define the scaling \( \tilde{A} = W A W \). Notice that \( \tilde{A} \) is symmetric positive definite. We apply Lemma 3 to \( \tilde{A} \). It is clear that \( \| \tilde{A}_{ii} \|_2 = 1 \). By (21) and (23) we obtain
\[
\| \tilde{A}_{ij} \|_2 = \| A_{ij} \|_2 w_i w_j < 1 \quad \text{for } i \neq j.
\]

We use the same reasoning as in Theorems 3 and 4. From Lemma 2 it follows there exists a symmetric matrix \( \tilde{F} \) such that
\[
\| \tilde{F}_{ij} \|_2 \leq \epsilon \| \tilde{A}_{ij} \|_2 \quad \text{for } i \neq j.
\]
and
\[
\| \tilde{F}_{ii} \|_2 \leq \epsilon \left( \| \tilde{A}_{ii} \|_2 + 2 \sum_{j=1}^{i-1} \| \tilde{A}_{ij} \|_2 \right).
\]

This together with \( \| \tilde{A}_{ij} \|_2 \leq 1 \) gives
\[
\mu(\tilde{F}) \leq (\epsilon + 2\epsilon(s - 1))\mu(\tilde{A}) = (2s - 1)\epsilon \mu(\tilde{A}).
\]
With \( F = W^{-1}\tilde{F} W^{-1} \), we find that \( F \) is symmetric and \( \mu(F) \leq \mu(W^{-1})\mu(\tilde{F})\mu(W^{-1}) \leq (2s - 1)\epsilon \mu(A) \). This gives the desired conclusion. \( \square \)

**References**

