Comm. Korean Math. Soc. 9 (1994), No. 2, pp. 439–448

THE STEEPEST DESCENT METHOD AND THE CONJUGATE GRADIENT METHOD FOR SLIGHTLY NON-SYMMETRIC, POSITIVE DEFINITE MATRICES

DONG HO SHIN, DO HYUN KIM AND MAN SUK SONG

1. Introduction and preliminaries

It is known that the steepest descent(SD) method and the conjugate gradient(CG) method [1, 2, 5, 6] converge when these methods are applied to solve linear systems of the form

$$Ax = b,$$

where A is symmetric and positive definite. For some finite difference discretizations of elliptic problems, one gets positive definite matrices that are almost symmetric. Practically, the SD method and the CG method work for these matrices. However, the convergence of these methods is not guaranteed theoretically. The SD method is also called Orthores(1) in iterative method papers. Elman [4] states that the convergence proof for Orthores(k), with k a positive integer, is not heard. In this paper, we prove that the SD method and the CG method converge when the l^2 matrix norm of the non-symmetric part of a positive definite matrix is less than some value related to the smallest and the largest eigenvalues of the symmetric part of the given matrix.

For non-symmetric matrices, many iterative methods [3, 4] have been developed that come from the CG method by changing the number of terms, the number of iterations, or the inner product and so forth. The convergence of most of these CG-like methods were proven by the use of Krylov space techniques. The convergence proofs that we do in this

Received July 27, 1993.
The first author was supported by KOSEF-GARC.
The second author was supported by a Scientific Research grant from the Ministry of Education, 1993.

paper are done without the use of the Krylov space. Hence our setting and proof are much different from those that are done so far about the CG-like methods for non-symmetric matrices.

We begin with some notations. Let

$$(x, y) := x^t y$$

be the Euclidean inner product of two vectors $x, y \in \mathbf{R}^n$ and let

$$\|x\| := \sqrt{(x, x)}$$

be the induced Euclidean norm. The associated matrix norm is given by

$$\sup_{\|x\|=1} \|Ax\|.$$

Let A be an $n \times n$ matrix. Note that A can be represented as

$$A = A_S + A_N,$$

where

$$A_S = (A + A^t)/2 \qquad \text{and} \qquad A_N = (A - A^t)/2$$

are the symmetric and the non-symmetric parts of A respectively. Let A be positive definite, then the symmetric part A_S is also positive definite. Hence A and A_S are invertible and the eigenvalues of A_S are all positive real numbers.

Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A_S such that

$$0 < \lambda_1 \leq \cdots \leq \lambda_n,$$

then $\lambda_1^{-1}, \dots, \lambda_n^{-1}$ are the eigenvalues of A_S^{-1}. The condition number of A_S is defined to be

$$\kappa := \|A_S\| \|A_S^{-1}\|.$$

Since A_S is symmetric,

$$\kappa = \lambda_n / \lambda_1 \geq 1.$$

The following lemmas will be used frequently in the proof of the main theorems.

LEMMA 1.1. *For any vector* x,

$$\lambda_1 \|x\|^2 \leq (x, A_S x) \leq \lambda_n \|x\|^2.$$

Proof. The Rayleigh quotients of a symmetric matrix are bounded by the smallest and the largest eigenvalues of the matrix. Thus, we have

$$\lambda_1 \leq \frac{(x, A_S x)}{\|x\|^2} \leq \lambda_n$$

for any x. Hence our claim follows.

Similarly, we get the next lemma.

LEMMA 1.2. *For any vector* x,

$$\lambda_n^{-1} \|x\|^2 \leq (x, A_S^{-1} x) \leq \lambda_1^{-1} \|x\|^2.$$

LEMMA 1.3. *For any vector* x,

$$(x, Ax) = (x, A_S x) \qquad \text{and} \qquad (x, A_N x) = 0.$$

Proof. Since $x^t A_N x$ is a real number,

$$x^t A_N x = (x^t A_N x)^t = x^t A_N^t x = -x^t A_N x.$$

Hence $(x, A_N x) = 0$.

The Schwarz inequality for a positive definite matrix is stated in the next lemma.

LEMMA 1.4. *If* P *is a positive definite matrix, then*

$$|(x, Py)| \leq \sqrt{(x, Px)(y, Py)}$$

for any x *and* y.

2. Convergence proofs

The convergence proof for the SD method applied to the linear system $Ax = b$ is shown in the next theorem.

THEOREM 2.1. *If*

$$\|A_N\| < \lambda_1 \sqrt{\kappa^{-1}} \left(-1 + \sqrt{1+\kappa^{-1}}\right),$$

then the SD method, defined by

$$\begin{aligned} r^0 &= b - Au^0 \\ x^{k+1} &= x^k + \alpha_k r^k \\ r^{k+1} &= r^k - \alpha_k A r^k \\ \alpha_k &= \frac{(r^k, r^k)}{(r^k, Ar^k),} \end{aligned}$$

converges.

Proof. By Lemma 1.2 and Lemma 1.3, for any k,

$$\lambda_n^{-1} \le \alpha_k \le \lambda_1^{-1} k. \tag{2.1}$$

We have

$$\begin{aligned} (r^{k+1}, A_S^{-1} r^{k+1}) &= (r^k - \alpha_k A r^k, A_S^{-1} r^k - \alpha_k A_S^{-1} A r^k) \\ &= (r^k, A_S^{-1} r^k) - 2\alpha_k (Ar^k, A_S^{-1} r^k) + \alpha_k^2 (Ar^k, A_S^{-1} A r^k) \\ &= (r^k, A_S^{-1} r^k) - 2\alpha_k \{(r^k, r^k) + (A_N r^k, A_S^{-1} r^k) \\ &\qquad + \alpha_k^2 \{(Ar^k, r^k) + (Ar^k, A_S^{-1} A_N r^k)\} \\ &= (r^k, A_S^{-1} r^k) - \alpha_k (r^k, r^k) - 2\alpha_k (A_N r^k, A_S^{-1} r^k) \\ &\qquad + \alpha_k^2 (Ar^k, A_S^{-1} A_N r^k) \\ &= (r^k, A_S^{-1} r^k) - \alpha_k (r^k, r^k) - 2\alpha_k (A_N r^k, A_S^{-1} r^k) \\ &\qquad + \alpha_k^2 \{(r^k, A_N r^k) + (A_N r^k, A_S^{-1} A_N r^k)\}. \end{aligned} \tag{2.2}$$

Let $\epsilon := \|A_N\|$ and $a^k := (r^k, A_S^{-1} r^k)$ for any k. By Lemma 1.2 and (2.1), we have

$$\alpha_k(r^k, r^k) \geq \lambda_n^{-1}\lambda_1 a^k = \kappa^{-1} a^k. \tag{2.3}$$

Applying Lemma 1.2 twice, we get

$$\begin{aligned} |(A_N r^k, A_S^{-1} A_N r^k)| &\leq \lambda_1^{-1}\|A_N r^k\|^2 \leq \lambda_1^{-1}\epsilon^2\|r^k\|^2 \\ &\leq \lambda_1^{-1}\epsilon^2\lambda_n(r^k, A_S^{-1} r^k) \leq \kappa\epsilon^2 a^k. \end{aligned} \tag{2.4}$$

By Lemma 1.4 and (2.4),

$$|(A_N r^k, A_S^{-1} r^k)| \leq \sqrt{(A_N r^k, A_S^{-1} A_N r^k) a^k} = \sqrt{\kappa}\,\epsilon a^k. \tag{2.5}$$

Using Lemma 1.3 and combining the equations from (2.1) to (2.5),

$$\begin{aligned} a^{k+1} &\leq a^k - \kappa^{-1} a^k + 2\lambda_1^{-1}\sqrt{\kappa}\epsilon a^k + \lambda_1^{-2}\kappa\epsilon^2 a^k \\ &\leq a^k(1 - \kappa^{-1} + 2\lambda_1^{-1}\sqrt{\kappa}\epsilon + \lambda_1^{-2}\kappa\epsilon^2). \end{aligned}$$

For convergence, we require

$$\begin{aligned} &1 - \kappa^{-1} + 2\lambda_1^{-1}\sqrt{\kappa}\epsilon + \lambda_1^{-2}\kappa\epsilon^2 < 1, \\ &\lambda_1^{-2}\kappa\epsilon^2 + 2\lambda_1^{-1}\sqrt{\kappa}\epsilon - \kappa^{-1} < 0, \\ &\epsilon^2 + 2\lambda_1\sqrt{\kappa^{-1}}\epsilon - \lambda_1^2\kappa^{-2} < 0. \end{aligned}$$

This happens when

$$\epsilon < -\lambda_1\sqrt{\kappa^{-1}} + \sqrt{\lambda_1^2\kappa^{-1} + \lambda_1^2\kappa^{-2}} = \lambda_1\sqrt{\kappa^{-1}}\left(-1 + \sqrt{1 + \kappa^{-1}}\right).$$

The main result for the CG method for solving the linear system $Ax = b$ is the following.

THEOREM 2.2. *If*

$$\|A_N\| < \lambda_1 \left(-1 + \sqrt{1 + \kappa^{-1}}\right),$$

then the CG method, defined by

$$\begin{aligned} p^0 &= r^0 = b - Au^0 \\ x^{k+1} &= x^k + \alpha_k p^k \\ r^{k+1} &= r^k - \alpha_k A p^k \\ p^{k+1} &= r^{k+1} + \beta_k p^k \\ \alpha_k &= \frac{(p^k, r^k)}{(p^k, Ap^k)} \\ \beta_k &= -\frac{(r^{k+1}, Ap^k)}{(p^k, Ap^k)}, \end{aligned}$$

converges.

We will use the notations

$$\epsilon := \|A_n\| \qquad \text{and} \qquad a_k := (r_k, A_S^{-1} r_k)$$

as we did in the proof of the SD method.

LEMMA 2.3. *For any x and y,*

$$|(x,y)| \le \sqrt{(x, A_S x)(y, A_S^{-1} y)}.$$

Proof. By Lemma 1.4,

$$\begin{aligned} |(x,y)| &= |(x, A_S(A_S^{-1} y))| \\ &\le \sqrt{(x, A_S x)(A_S^{-1} y, A_S(A_S^{-1} y))} = \sqrt{(x, A_S x)(A_S^{-1} y, y)}. \end{aligned}$$

LEMMA 2.4. *For any positive integer k,*

$$|\alpha_k| \leq \sqrt{\frac{a^k}{\lambda_1 \|p^k\|^2}}.$$

Proof. By Lemma 1.1, Lemma 1.3 and Lemma 2.3,

$$\begin{aligned}|\alpha_k| &= \frac{(p^k, r^k)}{(p^k, Ap^k)} \leq \frac{\sqrt{(p^k, A_S p^k)(r^k, A_S^{-1} r^k)}}{(p^k, Ap^k)} \\ &= \frac{\sqrt{(p^k, A_S p^k) a^k}}{(p^k, A_S p^k)} \leq \sqrt{\frac{a^k}{(p^k, A_S p^k)}} \leq \sqrt{\frac{a^k}{\lambda_1 \|p^k\|^2}}.\end{aligned}$$

LEMMA 2.5. *For any positive integer k,*

$$(p^k, r^k) = (r^k, r^k).$$

Proof. Note that for any k,

$$(p^k, r^{k+1}) = (p^k, r^k) - \alpha_k (p^k, Ap^k) = 0,$$

by the definition of α_k. Thus,

$$(p^k, r^k) = (r^k, r^k) + \beta_{k-1}(p^{k-1}, r^k) = (r^k, r^k).$$

LEMMA 2.6. *For any positive integer k,*

$$(p^k, Ap^k) = (p^k, Ar^k).$$

Proof. By the definition of β_k, we have

$$(p^{k+1}, Ap^k) = (r^{k+1}, Ap^k) + \beta_k (p^k, Ap^k) = 0$$

for any k. Hence,

$$(p^k, Ap^k) = (p^k, Ar^k) + \beta_{k-1}(p^k, Ap^{k-1}) = (p^k, Ar^k).$$

LEMMA 2.7. *For any positive integer k,*

$$\alpha_k(p^k, r^k) \geq \kappa^{-1} a^k.$$

Proof. By Lemma 1.4 and Lemma 2.6,

$$(p^k, Ap^k) = (p^k, Ar^k) \leq \sqrt{(p^k, Ap^k)(r^k, Ar^k)}.$$

Thus, we get

$$(p^k, Ap^k) \leq (r^k, Ar^k).$$

Using Lemma 1.1, Lemma 1.2, Lemma 1.3 and Lemma 2.5,

$$\alpha_k(p^k, r^k) = \frac{(r^k, r^k)^2}{(p^k, Ap^k)} \geq \frac{(r^k, r^k)^2}{(r^k, Ar^k)} \geq \frac{(r^k, r^k)}{\lambda_n} \geq \frac{\lambda_1}{\lambda_n} a^k.$$

With the lemmas above, the convergence of our CG method can be shown similarly to the way that the SD method was proven.

Proof of Theorem 2.2. We have

$$\begin{aligned}(r^{k+1}, A_S^{-1} r^{k+1}) &= (r^k - \alpha_k Ap^k, A_S^{-1} r^k - \alpha_k A_S^{-1} Ap^k) \\ &= (r^k, A_S^{-1} r^k) - 2\alpha_k (Ap^k, A_S^{-1} r^k) + \alpha_k^2 (Ap^k, A_S^{-1} Ap^k) \\ &= (r^k, A_S^{-1} r^k) - 2\alpha_k \{(p^k, r^k) + (A_N p^k, A_S^{-1} r^k)\} \\ &\qquad + \alpha_k^2 \{(Ap^k, p^k) + (Ap^k, A_S^{-1} A_N p^k)\} \\ &= (r^k, A_S^{-1} r^k) - \alpha_k (p^k, r^k) - 2\alpha_k (A_N p^k, A_S^{-1} r^k) \\ &\qquad + \alpha_k^2 (Ap^k, A_S^{-1} A_N p^k) \\ &= (r^k, A_S^{-1} r^k) - \alpha_k (p^k, r^k) - 2\alpha_k (A_N p^k, A_S^{-1} r^k) \\ &\qquad + \alpha_k^2 \{(p^k, A_N p^k) + (A_N p^k, A_S^{-1} A_N p^k)\}.\end{aligned} \tag{2.6}$$

By Lemma 1.2,

$$|(A_N p^k, A_S^{-1} A_N p^k)| \leq \lambda_1^{-1} \|A_N p^k\|^2 \leq \lambda_1^{-1} \epsilon^2 \|p^k\|^2. \tag{2.7}$$

Using (2.7) and Lemma 1.4,

$$(2.8) \qquad |(A_N p^k, A_S^{-1} r^k)| \le \sqrt{(A_N p^k, A_S^{-1} A_N p^k) a^k} \le \sqrt{\lambda_1^{-1} \epsilon^2 \|p^k\|^2 a^k}.$$

By Lemma 1.3, Lemma 2.4, Lemma 2.7 and equations from (2.6) to (2.8), one gets

$$a^{k+1} \le a^k - \kappa^{-1} a^k + 2\sqrt{\frac{a^k}{\lambda_1 \|p^k\|^2}} \sqrt{\lambda_1^{-1} \epsilon^2 \|p^k\|^2 a^k} + $$
$$+ \frac{a^k}{\lambda_1 \|p^k\|^2} \lambda_1^{-1} \epsilon^2 \|p^k\|^2 = a^k (1 - \kappa^{-1} + 2\lambda_1^{-1}\epsilon + \lambda_1^{-2}\epsilon^2).$$

For convergence, we require

$$1 - \kappa^{-1} + 2\lambda_1^{-1}\epsilon + \lambda_1^{-2}\epsilon^2 < 1,$$
$$\lambda_1^{-2}\epsilon^2 + 2\lambda_1^{-1}\epsilon - \kappa^{-1} < 0,$$
$$\epsilon^2 + 2\lambda_1\epsilon - \kappa^{-1}\lambda_1^2 < 0.$$

Hence, the sufficient condition for convergence is

$$\epsilon < -\lambda_1 + \sqrt{\lambda_1^2 + \kappa^{-1}\lambda_1^2} = \lambda_1 \left(-1 + \sqrt{1 + \kappa^{-1}}\right).$$

References

1. Owe Axelsson, *Solution of linear systems of equations : iterative methods*, Sparse Matrix Techniques, Springer-Verlag, New York, 1976, pp. 1–51.
2. Rati Chandra, *Conjugate Gradient Methods for Partial Differential Equations*, Ph.D. Thesis, Department of Computer Science, Yale University, 1978.
3. P. Concus, G. H. Golub and D. P. O'Leary, *A generalized conjugate gradient method for the numerical solution of elliptic partial differential equations*, Sparse Matrix Computations, J. R. Bunch and D. J. Rose, ed., Academic Press, New York, 1976, pp. 309–322.
4. H. C. Elman, *Iterative methods for large sparse nonsymmetric, systems of linear equations*, Ph.D. Thesis, Yale University, 1982.
5. Magnus R. Hestenes and Eduard Stiefel, *Methods of conjugate gradients for solving linear systems*, Journal of Research of the National Bureau of Standards **49** (1952), 409–435.

6. J. K. Reid, *On the method of conjugate gradients for the solution of large sparse systems of linear equations*, Large Sparse Sets of Linear Equations, Academic Press, New York, 1971, pp. 231–254.

Department of Mathematics
Inje University
Kimhae 621-749, Korea

Department of Mathematics Education
Cheju University
Cheju 690-756, Korea

Department of Computer Sciences
Yonsei University
Seoul 120-749, Korea