Jitka Křížková; Petr Vaněk

Preconditioning of conjugate gradients by multigrid solver

Applications of Mathematics, Vol. 39 (1994), No. 5, 357–364

Persistent URL: `http://dml.cz/dmlcz/134264`

39 (1994) APPLICATIONS OF MATHEMATICS No. 5, 357–364

PRECONDITIONING OF CONJUGATE GRADIENTS BY MULTIGRID SOLVER

JITKA KŘÍŽKOVÁ, PETR VANĚK, Plzeň

(Received February 5, 1993)

Summary. Solving a system of linear algebraic equations by the preconditioned conjugate gradient method requires to solve an auxiliary system of linear algebraic equations in each step. In this paper instead of solving the auxiliary system one iteration of the two level method for the original system is done.

Keywords: Conjugate gradient method, preconditioning, multigrid method

AMS classification: 65F10

1. INTRODUCTION

Solving a system of linear algebraic equations $Ax = b$ by the preconditioned conjugate gradient method requires to solve an auxiliary system of linear algebraic equations $Cx = r$ in each step. The multigrid method for solving the auxiliary system has been used (e.g. see [4]). In this paper an other usage of the multigrid method for preconditioning is suggested. Instead of solving the system $Cx = r$ one iteration of the two level method for the original system $Ax = b$ is done.

2. PRECONDITIONED CONJUGATE GRADIENT METHOD

We will solve the system of linear algebraic equations

$$Ax = b$$

where A is a positive definite real matrix of order n. We will denote by (x, y) the usual scalar product in $\mathbb{R}_n$, the norm in $\mathbb{R}_n$ being $\|x\| = (x, x)^{\frac{1}{2}}$.

For an arbitrary linear operator L on $\mathbb{R}_n$, $\|L\|$ denotes the operator norm of L defined by the vector norm $\|x\|$ and $\|L\|_A$ denotes the operator norm of L defined by the vector norm $\|x\|_A = (Ax, x)^{\frac{1}{2}}$.

For positive definite operators C, A on $\mathbb{R}_n$ let us define the so called relative condition number

$$\operatorname{cond}(C, A) = \frac{\lambda_{\max}(C^{-1}A)}{\lambda_{\min}(C^{-1}A)}.$$

Usage of the preconditioned conjugate gradient method algorithm follows the scheme:
Step 1. Given $\varepsilon > 0$, $x_0 = 0$, let $k = 0$ and

$$\begin{aligned} r_0 &= b - Ax_0 = b \\ h_0 &= C^{-1}r_0, \\ p_0 &= h_0. \end{aligned}$$

Step 2. Do

$$\begin{aligned} \alpha_k &= \frac{(p_k, r_k)}{(Ap_k, p_k)}, \\ x_{k+1} &= x_k + \alpha_k p_k, \\ r_{k+1} &= r_k - \alpha_k A p_k, \\ h_{k+1} &= C^{-1} r_{k+1}, \\ \beta_k &= \frac{(r_{k+1}, h_{k+1})}{(r_k, h_k)}, \\ p_{k+1} &= h_{k+1} + \beta_k p_k. \end{aligned} \tag{2.1}$$

Step 3. Let c_k be an estimate of $\operatorname{cond}(C, A)$. If $c_k \frac{(r_k, h_k)}{(r_0, h_0)} \leqslant \varepsilon^2$, exit else $k = k + 1$ and go to Step 2—see [1].

Let us define the error $e(x)$ by $e(x) = x - \widehat{x}$, where $\widehat{x} = A^{-1}b$.
Then for the error of the preconditioned conjugate gradient method the following formula can be derived—see [3]:

$$\|e(x_i)\|_A \leqslant 2 \left(\frac{\sqrt{\operatorname{cond}(C, A)} - 1}{\sqrt{\operatorname{cond}(C, A)} + 1} \right)^i \|e(x_0)\|_A. \tag{2.2}$$

Let us remark that $e(x_0) = -\widehat{x}$.

3. Multigrid method

Our aim is to replace the solution of an auxiliary system of linear algebraic equations in (2.1) with one iteration of the two level method constructed for the solution of the system $Ax = b$. Let us consider the space $\mathbb{R}_m$, $m < n$ with the usual scalar product denoted by $(x, x)_2$. Let us suppose that an injective operator $p\colon \mathbb{R}_m \to \mathbb{R}_n$ is given (for technical details see [5]). Now let us denote by r the adjoint operator to p with respect to the both scalar products. Further, a linear iterative method of the form

$$S(x) = Mx + Nb \tag{3.1}$$

is supposed to be given, where M and N are operators on $\mathbb{R}_n$ satisfying the consistence condition

$$I = M + NA. \tag{3.2}$$

Let us note this condition is necessary and sufficient for the fulfilment of

$$S(\widehat{x}) = \widehat{x},$$

where $\widehat{x} = A^{-1}b$ is the exact solution of the given system. The usual two level algorithm follows the scheme:

for given x_i we set

$$\begin{aligned} \tilde{x} &= S(x_i), \\ \overline{x} &= \tilde{x} - p(rAp)^{-1}r(A\tilde{x} - b), \\ x_{i+1} &= S(\overline{x}). \end{aligned} \tag{3.3}$$

It is not difficult to see that the algorithm (3.3) can be written in the form

$$x_{i+1} = Qx_i + \widehat{Q}b, \quad i = 0, 1, \ldots,$$

where

$$Q = M(I - p(rAp)^{-1}rA)M$$

and

$$\widehat{Q} = MN - Mp(rAp)^{-1}r(AN - I) + N. \tag{3.4}$$

These relations yield

$$e(x_{i+1}) = Qe(x_i), \quad i = 0, 1, \ldots. \tag{3.5}$$

If we start with $x_0 = 0$ then $x_1 = \widehat{Q}b$.

R e m a r k 3.1. Let us denote by $P = I - p(rAp)^{-1}rA$. It is easy to show that $P^2 = P$ and

$$(APx, y) = (Ax, Py) \text{ for all } x, y \in \mathbb{R}_n.$$

Therefore P is the A-orthogonal projection.

Now we will use $\widehat{Q}$ instead of C^{-1} in (2.1). For this reason we need symmetry of $\widehat{Q}$.

4. Convergence analysis

Let us consider the preconditioned conjugate gradient method (2.1) and let us set $C^{-1} = \widehat{Q}$, where $\widehat{Q}$ belongs to an iterative method

$$x_{i+1} = Qx_i + \widehat{Q}b, \quad i = 0, 1, \ldots.$$

In this section estimates of rate of convergence will be proved.

Assumptions 4.1. Let us suppose $\|Q\|_A = q < 1$, $\widehat{Q}$ is symmetric, the consistence condition $I = Q + \widehat{Q}A$ is fulfilled and $(AQx, x) \geqslant 0$ for all $x \in \mathbb{R}_n$.

Lemma 4.1. *Let Q, $\widehat{Q}$ fulfil assumptions* 4.1. *Then $\widehat{Q}$ is positive definite.*

P r o o f. Let $f \in \mathbb{R}_n$ be arbitrary nonzero. Let us denote by $\widehat{x} = A^{-1}f, e(x) = x - \widehat{x}$. Let us set $x_0 = 0$. Then $e(x_0) = -\widehat{x} \neq 0$. Assumptions 4.1 imply

$$(AQe(x_0), Qe(x_0)) < (Ae(x_0), e(x_0)).$$

Using the relations $e(x_1) = Qe(x_0)$, where $e(x_1) = x_1 - \widehat{x} = \widehat{Q}f - \widehat{x}$ we get

$$(A(\widehat{Q}f - \widehat{x}), \widehat{Q}f - \widehat{x}) < (A\widehat{x}, \widehat{x}),$$

and using the positive definiteness of A we conclude

$$0 \leqslant \frac{1}{2}(A\widehat{Q}f, \widehat{Q}f) < (\widehat{Q}f, f).$$

This inequality completes the proof. □

The following useful lemma is not difficult to prove.

Lemma 4.2. *Let K, L be two positive definite operators on $\mathbb{R}_n$. Then*
1. $\sigma(KL) = \sigma(K^{\frac{1}{2}}LK^{\frac{1}{2}}) = \sigma(LK) = \sigma(L^{\frac{1}{2}}KL^{\frac{1}{2}})$.

2. $\lambda > 0$ *for all* $\lambda \in \sigma(KL)$ *and* KL *is diagonalizable.*

Lemma 4.3. *For all eigenvalues* λ *of* $A^{\frac{1}{2}}\widehat{Q}A^{\frac{1}{2}}$ *we have*

$$1 - q \leqslant \lambda \leqslant 1.$$

Proof. At first from the assumptions 4.1 we get that $Q = I - \widehat{Q}A$ is $A-$ positive definite and therefore for all eigenvalues λ of $A^{\frac{1}{2}}\widehat{Q}A^{\frac{1}{2}}$ we have (see Lemma 4.2) $\lambda \leqslant 1$. The second part of the above inequality is the consequence of

$$\|I - \widehat{Q}A\|_A = q < 1$$

(see again Assumptions 4.1) and Lemma 4.2. □

Remark 4.1. The relative condition number $\mathrm{cond}(C, A)$ was defined in Section 2 as

$$\mathrm{cond}(C, A) = \frac{\lambda_{\max}(C^{-1}A)}{\lambda_{\min}(C^{-1}A)}.$$

Due to Lemma 4.2 $\mathrm{cond}(C, A)$ is equal to

$$\mathrm{cond}(C, A) = \frac{\lambda_{\max}(AC^{-1})}{\lambda_{\min}(AC^{-1})} = \frac{\lambda_{\max}(A^{\frac{1}{2}}C^{-1}A^{\frac{1}{2}})}{\lambda_{\min}(A^{\frac{1}{2}}C^{-1}A^{\frac{1}{2}})}.$$

Theorem 1. *Let us consider the preconditioned conjugate gradient method* (2.1) *in which* $C^{-1} = \widehat{Q}$, $\widehat{Q}$ *being defined by the iterative method* $x_{i+1} = Qx_i + \widehat{Q}b$. *Let us suppose assumptions* 4.1 *are fulfilled. Then*

1. $\mathrm{cond}(C, A) \leqslant \frac{1}{1-q}$,
2. $\|e(x_i)\|_A \leqslant 2q_G^i\|e(x_0)\|_A$, *where* $q_G = \frac{1-\sqrt{1-q}}{1+\sqrt{1-q}}$

is the estimate for the preconditioned conjugate gradient method.

Proof. The statement 1 is the immediate consequence of Lemma 4.3. The estimate 2 follows from 1, (2.2) and Remark 4.1. □

Theorem 2. *Let us consider the preconditioned conjugate gradient method* (2.1) *in which* $C^{-1} = \widehat{Q}$, $\widehat{Q}$ *being the operator of the two level method (see 3.4). Let us suppose that* $M = I - BA$, $N = B$, *where* B *is symmetric,* $\|Q\|_A = q < 1$ *(for* Q *see 3.4). Then*

1. $\mathrm{cond}(C, A) \leqslant \frac{1}{1-q}$,
2. $\|e(x_i)\|_A \leqslant 2q_G^i\|e(x_0)\|_A$, $q_G = \frac{1-\sqrt{1-q}}{1+\sqrt{1-q}}$.

Proof. We will verify assumptions 4.1. First, the consistence condition $I = Q + \widehat{Q}A$ is easy to prove using (3.4). The operator $\widehat{Q}$ can be written in the form

$$\widehat{Q} = 2B - BAB + (I - BA)p(rAp)^{-1}r(I - BA)^T,$$

its symmetry now being evident. For an arbitrary $x \in \mathbb{R}_n$ we have

$$(AQx, x) = (A(I - BA)P(I - BA)x, x) = (AP(I - BA)x, (I - BA)x) \geqslant 0$$

as $P = I - p(rAp)^{-1}rA$ is the A-orthogonal projection (see Remark 3.1). $\square$

Remark 4.2. Conditions

$$M = I - BA, \qquad N = B, \tag{4.1}$$

B symmetric are fulfilled for the damped Jacobi method

$$S(x_{i+1}) = (I - \omega D^{-1}A)x_i + \omega D^{-1}b, \quad \omega \in (0, 1), \tag{4.2}$$

D being either the diagonal part of A or generally the block diagonal part of A, as well as for the iterative method

$$S(x_{i+1}) = (I - \omega A)x_i + \omega b. \tag{4.3}$$

If we use the method (4.2) or (4.3) several times (even with different ω) it is possible to show that the obtained iterative method is also of the form (4.1) required by Theorem 2.

Remark 4.3. The relation between q and $q_G = \frac{1-\sqrt{1-q}}{1+\sqrt{1-q}}$ is shown in Fig. 1.

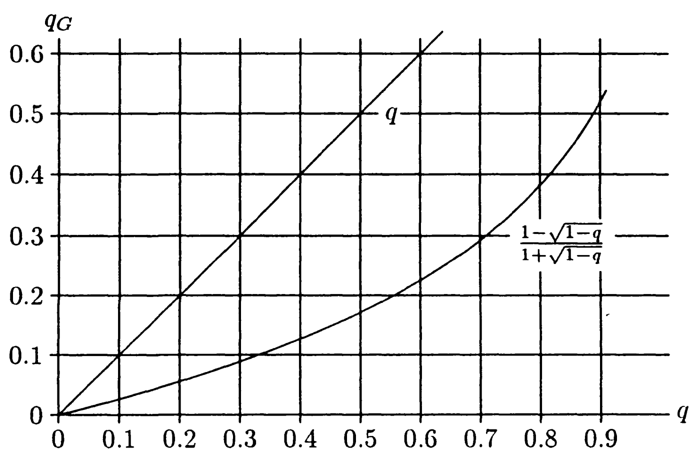

Fig. 1

5. Some remarks on the convergence of two level method

In this section we will present (without proving) some statements about the convergence of the two level method (3.3)—for details see [5]. Let us consider the two level method described in Section 3. Let us denote by $U = \mathrm{Ker}(rA)$, X_U the restriction of an operator X to U. Then

$$\|Q\|_A \leqslant \|M_U\|_A^2$$

(see Lemma 6.5 in [5]).

If the iterative method 4.3 is used the operator norm $\|M_U\|_A$ can be estimated as follows.

Let the constant $c > 0$ be such that for every $x \in \mathbb{R}_n$ there exists $v \in \mathbb{R}_n$,

$$c\|x - pv\| \leqslant \|x\|_A.$$

Then $\|Mu\|_A \leqslant [1 - c\omega(2 - \omega\varrho(A))]\|u\|_A$ for every $u \in U$ (see Lemma 7.2 in [7]).

If the iterative method (4.2) is used the operator norm $\|M_U\|_A$ can be estimated as follows:

Let $D > 0$, let c be a positive constant such that for every $x \in \mathbb{R}_n$ there exists $v \in \mathbb{R}_n$ such that

$$c\|D^{\frac{1}{2}}(x - pv)\| \leqslant \|x\|_A.$$

Then $\|Mu\|_A \leqslant [1 - c\omega(2 - \varrho(D^{-\frac{1}{2}}AD^{\frac{1}{2}}))]\|u\|_A$ for every $u \in U$ (see Lemma 6.8 in [5]).

For model examples the technique for determining the numerical values of c can be found in [4].

References

[1] *J. Mandel*: Adaptive Iterative Solvers in Finite Elements. To appear.

[2] *J. Mandel*: Balancing Domain Decomposition. Communications in numerical methods in engineering, vol. 9 (1993).

[3] *G. Luenberger*: Introduction to Linear and Nonlinear Programming. Addison-Wesley, New York, 1973.

[4] *R. Blaheta*: Iterative Methods for Numerical Solving of the Boundary Value Problems of Elasticity. Thesis, Ostrava, 1989. (In Czech.)

[5] *P. Vaněk*: Acceleration of Algebraic Multigrid Method. Proceedings of the IXth Summer School SANM. 1991.

[6] *W. Hackbusch*: Multigrid Methods and Applications. Springer Verlag, 1985.

[7] *S. Míka, P. Vaněk*: The Acceleration of the Convergence of a Two-level Algebraic Algorithm by Aggregation in Smoothing Process. Applications of Math. *37* (1992), no. 5.

[8] *O. Axelsson, V.A. Barker*: Finite Element Solution of Boundary Value Problems. Academic Press, 1984.

Authors' address: *Jitka Křížková, Petr Vaněk*, Department of Math., University of West Bohemia, Americká 42, 306 14 Plzeň, Czech Republic, e-mail: **krizkova@scylla.zcu.cz**, **pvanek@scylla.zcu.cz**.