Applications of Mathematics

Jitka Křížková; Petr Vaněk

Multigrid method with preconditioning on coarse level

Applications of Mathematics, Vol. 39 (1994), No. 5, 365--373

Persistent URL: http://dml.cz/dmlcz/134265

Terms of use:

© Institute of Mathematics AS CR, 1994

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://project.dml.cz

MULTIGRID METHOD WITH PRECONDITIONING ON COARSE LEVEL

JITKA KŘÍŽKOVÁ, PETR VANĚK, Plzeň

(Received February 5, 1993)

Summary. An algorithm for using the preconditioned conjugate gradient method to solve a coarse level problem is presented.

Keywords: Conjugate gradient method, preconditioning, multigrid method

AMS classification: 65F10

1. INTRODUCTION

Let us consider a system of linear algebraic equations

$$Ax = b,$$

where A is a positive definite matrix of order n. The efficiency of a multigrid solver depends on the properties of a prolongation operator p. The multigrid solver is well constructed if the range of p contains all vectors that cannot be effectively eliminated by smoothing. These vectors will be called smooth. Non-smooth vectors from the range of p can be suppressed using the operator $M^\nu p$ instead of p, M being a smoothing operator, ν a positive integer. If we use this prolongation operator the coarse level problem with the matrix $A_\nu = p^T(M^\nu)^T AM^\nu p$ must be solved. Choosing $\nu \geqslant 1$ the rate of convergence is very good but the construction of the matrix A_ν becomes time consuming. In this paper an algorithm not requiring the construction of the matrix mentioned above is proposed. The properties of $A_0 = p^T Ap$ are similar to those of A_ν therefore A_0 is suitable for preconditioning in the conjugate gradient method. The rates of convergence of both the conjugate gradient method and the multigrid method are analysed.

2. Notation

Let m, n be positive integers, $m < n$. We will denote by (x, y) the usual scalar product in $\mathbb{R}_n$, the norm in $\mathbb{R}_n$ being $\|x\| = (x, x)^{\frac{1}{2}}$, $(x, y)_2$ will denote the standard scalar product in $\mathbb{R}_m$. Let H be a finite dimensional Hilbert space. For an arbitrary linear operator L on H, $\|L\|$ denotes the operator norm of L defined by the norm $\|.\|$, $\varrho(L)$ the spectral radius of L, L^* the adjoint operator. Every positive definite operator K on H defines the K-scalar product $(K., .)$, $\|.\|_K$ denotes the corresponding norm and $\|L\|_K$ denotes the corresponding operator norm. For K, L positive definite operators on H let us denote by $Q(x)$,

$$Q(x) = \frac{(Lx, x)}{(Kx, x)}$$

for every $x \in H$, $x \neq 0$. Let us define the so called relative condition number of K and L by

$$\operatorname{cond}(K, L) = \frac{\max\limits_{x \neq 0} Q(x)}{\min\limits_{x \neq 0} Q(x)}.$$

Lemma 2.1. *Let K, L be positive definite operators on H. Then*

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

P r o o f. It is not difficult to see that

$$\sigma(K^{-1}L) = \sigma(K^{-\frac{1}{2}}LK^{-\frac{1}{2}}),$$

therefore

$$\lambda_{\max}(K^{-1}L) = \lambda_{\max}(K^{-\frac{1}{2}}LK^{-\frac{1}{2}}) = \max_{x \neq 0} \frac{(K^{-\frac{1}{2}}LK^{-\frac{1}{2}}x, x)}{(x, x)}$$

and setting $y = K^{-\frac{1}{2}}x$ we get

$$\lambda_{\max}(K^{-1}L) = \max_{x \neq 0} \frac{(Ly, y)}{(Ky, y)}.$$

The statement of the lemma follows from the analogous expression for $\lambda_{\min}(K^{-1}L)$.
$\square$

3. Algorithm

Let us consider the iterative method

$$S(x) = Mx + Nb,$$

where M, N are linear operators on $\mathbb{R}_n$ satisfying the consistence condition

$$I = M + NA,$$

M is regular and $\varrho(M) < 1$. Let $p\colon \mathbb{R}_m \to \mathbb{R}_n$ be a linear injective operator. Let us note p is usually constructed so that $Mp \approx p$ (for technical details see [5], [9]). Let us denote by r the linear operator adjoint to p with respect to the standard scalar products on $\mathbb{R}_n$ and $\mathbb{R}_m$.

Definition 3.1. For every integer $i \geqslant 0$ let us define

$$\begin{aligned} p_i &= M^i p \\ r_i &= r(M^i)^* \\ A_i &= r_i A p_i. \end{aligned}$$

Remark 3.1. It is easy to see

1. $p_0 = p$, $r_0 = r$, $A_i = r(M^i)^* A M^i p$,
2. A_i is positive definite for all i.

Algorithm 3.1. For given x_i we set

$$\begin{aligned} \tilde{x} &= S^{(\xi_1)}(x_i) \quad (\xi_1\text{-times iterating } S) \\ d &= A\tilde{x} - b \\ d_2 &= r(M^*)^\nu d \end{aligned}$$

(3.1) v is determined so that $r(M^*)^\nu A M^\nu p v = d_2$

$$\begin{aligned} \overline{x} &= \tilde{x} - M^\nu p v \\ x_{i+1} &= S^{(\xi_2)}(\overline{x}), \end{aligned}$$

ξ_1, ξ_2, ν are positive integers, $\nu \approx 1 - 4$, $\xi_1 \approx 2\nu$. The matrix $A_\nu = r(M^*)^\nu A M^\nu p$ is not constructed, the problem (3.1) is solved by the preconditioned conjugate gradient method in the following form.

Algorithm 3.2.

Step 1. Given $v_0 = 0$, let $k = 0$ and

$$g_0 = d_2 - A_\nu v_0 = d_2,$$
$$h_0 = A_0^{-1} g_0,$$
$$s_0 = h_0.$$

Step 2. Repeat

$$\begin{aligned} \alpha_k &= \frac{(s_k, g_k)_2}{(A_\nu s_k, s_k)_2}, \\ v_{k+1} &= v_k + \alpha_k s_k, \\ g_{k+1} &= g_k - \alpha_k A_\nu s_k, \\ h_{k+1} &= A_0^{-1} g_{k+1}, \\ \beta_k &= \frac{(g_{k+1}, h_{k+1})_2}{(g_k, h_k)_2}, \\ s_{k+1} &= h_{k+1} + \beta_k s_k. \end{aligned} \tag{3.2}$$

Let us note that the preconditioning matrix $A_0 = rAp$. Let us define the error $e(v)$ by $e(v) = v - \widehat{v}$, where $\widehat{v}$ is the exact solution of (3.1).
Then for the error of the preconditioned conjugate gradient method the following formula can be derived—see [3]:

$$\|e(v_i)\|_{A_\nu} \leqslant 2\left(\frac{\sqrt{\operatorname{cond}(A_0, A_\nu)} - 1}{\sqrt{\operatorname{cond}(A_0, A_\nu)} + 1}\right)^i \|e(v_0)\|_{A_\nu}. \tag{3.3}$$

4. Coarse level problem convergence analysis

Definition 4.1. For every integer $i \geqslant 0$ let us define

$$S_i = R(p_i).$$

Lemma 4.1. *Let K be a regular selfadjoint operator on a Hilbert space H. Then*

$$\frac{\|K^2 x\|}{\|Kx\|} \geqslant \frac{\|Kx\|}{\|x\|}$$

for every $x \in H$, $x \neq 0$.

Proof.

$$\|Kx\|^2 = (K^2 x, x) \leqslant \|K^2 x\| \, \|x\|.$$

□

Definition 4.2. For every $x \in \mathbb{R}_n$, $i \geqslant 0$ let us define

$$\|x\|_i = (AM^i x, M^i x)^{\frac{1}{2}}.$$

Remark 4.1. Let us note that

$$\|.\|_0 = \|.\|_A.$$

Definition 4.3. Let us denote by c_ν, C_ν the constants of the norm equivalence between $\|.\|_\nu$ and $\|.\|_0$ on the subspace S_0, i.e.

$$c_\nu \|x\|_A \leqslant \|x\|_\nu \leqslant C_\nu \|x\|_A \quad \text{for every } x \in S_0.$$

Lemma 4.2. *If M is selfadjoint with respect to the A-scalar product, then*
1. $C_\nu \leqslant \varrho(M^\nu)$,
2. $c_\nu \geqslant c_0^\nu$.

Proof.

$$\|px\|_\nu \leqslant \|M^\nu px\|_A \leqslant \|M^\nu\|_A \|px\|_A = \varrho(M^\nu)\|px\|_A.$$

Using Lemma 4.1 we get

$$\frac{\|M^\nu px\|_A}{\|px\|_A} = \frac{\|M^\nu px\|_A}{\|M^{\nu-1}px\|_A} \cdot \frac{\|M^{\nu-1}px\|_A}{\|M^{\nu-2}px\|_A} \cdots \frac{\|Mpx\|_A}{\|px\|_A} \geqslant \left(\frac{\|Mpx\|_A}{\|px\|_A}\right)^\nu.$$

This inequality yields 2. □

Theorem 1. *Let us consider the conjugate gradient method for the system of linear algebraic equations with the matrix A_ν preconditioned by the matrix A_0 (Algorithm 3.1). Then*

$$\|e(v_i)\|_{A_\nu} \leqslant 2\left(\frac{C_\nu - c_\nu}{C_\nu + c_\nu}\right)^i \|e(v_0)\|_{A_\nu}.$$

Proof. For every $x \in \mathbb{R}_m$, $x \neq 0$

$$Q(x) = \frac{(r(M^\nu)^* AM^\nu px, x)_2}{(rApx, x)_2} = \frac{(AM^\nu px, M^\nu px)}{(Apx, px)} = \frac{\|M^\nu px\|_A^2}{\|px\|_A^2} = \frac{\|px\|_\nu^2}{\|px\|_A^2}.$$

Therefore

$$c_\nu^2 \leqslant Q(x) \leqslant C_\nu^2$$

and

$$\operatorname{cond}(A_0, A_\nu) \leqslant \left(\frac{C_\nu}{c_\nu}\right)^2.$$

Substituting this inequality into (3.3) we get the statement. □

Remark 4.2. 1. p is usually constructed so that $Mp \approx p$ and therefore $c_\nu \approx C_\nu \approx 1$. Due to this fact the rate of convergence will be good.

2. Lemma 4.2 yields that C_ν can be replaced by 1 and c_ν by c_0^ν if M is chosen so that M is A-selfadjoint (this is the case of the damped Jacobi method—see Section 5).

3. The spaces $\mathbb{R}_m$ with A_ν-scalar product and $R(p_\nu)$ with A-scalar product are isometrically isomorphic, therefore

$$\|e(v_i)\|_{A_\nu} = \|pe(v_i)\|_A.$$

5. Fine level problem convergence analysis

In this section M will be the operator of the damped Jacobi method, i.e.

$$M = I - \omega D^{-1}A, \ \omega \in (0,1), \ \mathrm{Ker}(M) = \{0\}.$$

Lemma 5.1. *AM is a selfadjoint operator.*

Proof.

$$M^*A = (I - \omega AD^{-1})A = A(I - \omega D^{-1}A) = AM.$$

□

Corollary. *M is selfadjoint with respect to the A-scalar product.*

Definition 5.1. For integer $i \geqslant 0$ let us define

$$T_i = \mathrm{Ker}(r_i A).$$

Remark 5.1. Lemma 5.1 implies

$$T_i = \mathrm{Ker}(rAM^i).$$

Lemma 5.2. *Let us consider the Algorithm 3.1, where*

$$S(x) = (I - \omega D^{-1}A)x + \omega D^{-1}b.$$

If the coarse level problem is solved exactly the following estimate holds:

$$\frac{\|e(x_{i+1})\|_A}{\|e(x_i)\|_A} \leqslant \|M_{T_\nu}^{\xi_1}\|_A \|M_{T_\nu}^{\xi_2}\|_A.$$

Proof. See [5]. □

Lemma 5.3. *T_0 and T_i are isomorphic, the corresponding isomorphism being M^i, i.e. $x \in T_i$ if and only if $M^i x \in T_0$.*

Proof. The statement is the immediate consequence of Definition 5.1. □

Due to Lemma 5.1 M is selfadjoint with respect to the A-scalar product. Therefore there exists an A-orthonormal basis v_j, $j = 1, \ldots, n$ of $\mathbb{R}_n$ consisting of eigenvectors of M belonging to the eigenvalues λ_j, $j = 1, \ldots, n$.

Definition 5.2. For $i \geqslant 0$ integer let us denote by T_i^c the linear space of coordinates of all vectors $x \in T_i$ with respect to the basis v_j, $j = 1, \ldots, n$, i.e.

$$T_i^c = \left\{ [c_1, \ldots c_n]^T,\ x = \sum_{j=1}^{n} c_j v_j,\ x \in T_i \right\}.$$

Lemma 5.4. *Every element of T_i^c, $i \geqslant 0$ integer is of the form*

$$\left[\frac{c_1}{\lambda_1^i}, \ldots, \frac{c_n}{\lambda_n^i} \right]^T, \text{ where } [c_1, \ldots, c_n]^T \in T_0^c.$$

Proof. Due to Lemma 5.4, $x \in T_0$ if and only if $M^{-i} x \in T_i$. Let

$$x = \sum_{j=1}^{n} c_j v_j,$$

then

$$M^{-i} x = \sum_{j=1}^{n} \frac{c_j}{\lambda_j^i} v_j.$$

□

Lemma 5.5. *Let i, ξ be positive integers, then*

$$\|M_{T_i}^{\xi}\|_A^2 = \max_{\substack{\mathbf{c} \in T_0^c \\ \mathbf{c} \neq 0}} \frac{\sum_{j=1}^{n} \lambda_j^{2\xi} \frac{c_j^2}{\lambda_j^{2i}}}{\sum_{j=1}^{n} \frac{c_j^2}{\lambda_j^{2i}}}, \quad \text{where } \mathbf{c} = [c_1, \ldots, c_n]^T.$$

Proof. For $x \in T_\nu$ we have

$$\|x\|_A^2 = \sum_{j=1}^{n} \frac{c_j^2}{\lambda_j^{2i}}, \quad \mathbf{c} = [c_1, \ldots, c_n]^T \in T_0^c$$

(see Lemma 5.4) and

$$\|M^{\xi}x\|_A^2 = \sum_{j=1}^{n} \lambda_j^{2\xi} \frac{c_j^2}{\lambda_j^{2i}}.$$

□

Theorem 2. *Let us consider the Algorithm* 3.1, *where*

$$S(x) = (I - \omega D^{-1}A)x + \omega D^{-1}b.$$

If the coarse level problem is solved exactly the following estimate holds:

$$\frac{\|e(x_{i+1})\|_A^2}{\|e(x_i)\|_A^2} \leqslant \max_{\substack{\mathbf{c}\in T_0^c\\ \mathbf{c}\neq 0}} \frac{\sum_{j=1}^{n} \lambda_j^{2\xi_1} \frac{c_j^2}{\lambda_j^{2i}}}{\sum_{j=1}^{n} \frac{c_j^2}{\lambda_j^{2i}}} \max_{\substack{\mathbf{c}\in T_0^c\\ \mathbf{c}\neq 0}} \frac{\sum_{j=1}^{n} \lambda_j^{2\xi_2} \frac{c_j^2}{\lambda_j^{2i}}}{\sum_{j=1}^{n} \frac{c_j^2}{\lambda_j^{2i}}}.$$

P r o o f. An immediate consequence of Lemmas 5.2 and 5.5. □

R e m a r k 5.2. If the transfer operators p_0, r_0 are well constructed then T_0 contains elements $\mathbf{c} = [c_1, \dots, c_n]^T$ for which the components c_j corresponding to the small eigenvalues λ_j, i.e. $|\lambda_j| \approx 0$ are large in comparison with the others. The stronger this property the smaller $\|M_{T_0}^{\xi}\|_A$ is (see Lemma 5.5). For small λ_j we have

$$\frac{c_j^2}{\lambda_j^{2\nu}} \gg c_j^2,$$

while for large λ_j, i.e. $|\lambda_j| \approx 1$,

$$\frac{c_j^2}{\lambda_j^{2\nu}} \approx c_j^2.$$

Therefore

$$\|M_{T_\nu}^{\xi}\|_A^2 = \max_{\substack{\mathbf{c}\in T_0^c\\ \mathbf{c}\neq 0}} \frac{\sum_{j=1}^{n} \lambda_j^{2\xi} \frac{c_j^2}{\lambda_j^{2\nu}}}{\sum_{j=1}^{n} \frac{c_j^2}{\lambda_j^{2\nu}}} \ll \|M_{T_0}^{\xi}\|_A^2 = \max_{\substack{\mathbf{c}\in T_0^c\\ \mathbf{c}\neq 0}} \frac{\sum_{j=1}^{n} \lambda_j^{2\xi} c_j^2}{\sum_{j=1}^{n} c_j^2}$$

can be expected.

Theorem 3. *Let us consider the Algorithm* 3.1, *where*

$$S(x) = (I - \omega D^{-1}A)x + \omega D^{-1}b.$$

Let $\xi_1, \xi_2 \geqslant \nu + 1$. *If the coarse level problem is solved exactly the following estimate holds:*

$$\frac{\|e(x_{i+1})\|_A^2}{\|e(x_i)\|_A^2} \leqslant \|M_{T_0}\|_A^{2\nu+2}.$$

Remark 5.3. Techniques for estimating $\|M_{T_0}\|_A$ can be found in [9].

Proof. For every $i \geqslant 1$, $x \in T_i$, $\xi \geqslant \nu + 1$ if and only if $Mx \in T_{i-1}$. Further,

$$\frac{\|M^\xi x\|_A}{\|x\|_A} = \frac{\|M^\xi x\|_A}{\|M^{\xi-1}x\|_A} \cdot \frac{\|M^{\xi-1}x\|_A}{\|M^{\xi-2}x\|_A} \cdot \frac{\|Mx\|_A}{\|x\|_A}.$$

Taking into account $\varrho(M) < 1$ we get

$$\|M_{T_\nu}^\xi\|_A \leqslant \|M_{T_\nu}\|_A \cdots \|M_{T_0}\|_A.$$

Lemma 4.1 implies

$$\|M_{T_i}\|_A \leqslant \|M_{T_0}\|_A, \quad i \geqslant 0.$$

Therefore

$$\|M_{T_\nu}^\xi\|_A \leqslant \|M_{T_0}\|_A^{\nu+1},$$

and the usage of Lemma 5.2 completes the proof. □

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 IX. 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. Appl. of Math. *37* (1992), no. 5.
[8] *O. Axelsson, V.A. Barker*: Finite Element Solution of Boundary Value Problems. Academic Press, 1984.
[9] *P. Vaněk*: Fast multigrid solver. To appear.

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