Jitka Křížková; Petr Vaněk

Multigrid method with preconditioning on coarse level


Persistent URL: [http://dml.cz/dmlcz/134265](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](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, \( \nu \) a positive integer. If we use this prolongation operator the coarse level problem with the matrix \( A_\nu = p^T (M^\nu)^T A M^\nu p \) must be solved. Choosing \( \nu \geq 1 \) the rate of convergence is very good but the construction of the matrix \( A_\nu \) becomes time consuming. In this paper an algorithm not requiring the construction of the matrix mentioned above is proposed. The properties of \( A_0 = p^T A p \) are similar to those of \( A_\nu \) 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 \( \|\cdot\| \), \( \varrho(L) \) the spectral radius of \( L \), \( L^* \) the adjoint operator. Every positive definite operator \( K \) on \( H \) defines the \( K \)-scalar product \((K.,.)\), \( \|\cdot\|_K \) 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

\[
\text{cond}(K, L) = \frac{\max_{x \neq 0} Q(x)}{\min_{x \neq 0} Q(x)}.
\]

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

\[
\text{cond}(K, L) = \frac{\lambda_{\max}(K^{-1}L)}{\lambda_{\min}(K^{-1}L)}.
\]

**Proof.** 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 \( p: \mathbb{R}^m \to \mathbb{R}^n \) 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 \( i \geq 0 \) let us define

\[ p_i = M^i p, \]
\[ r_i = r(M^i)^*, \]
\[ A_i = r_i A p_i. \]

**Remark 3.1.** It is easy to see

1. \( p_0 = p, r_0 = r, A_0 = r(M^1)^*AM^1p, \)
2. \( A_i \) is positive definite for all \( i. \)

**Algorithm 3.1.** For given \( x_i \) we set

\[ \tilde{x} = S^{(\xi_1)}(x_i) \quad (\xi_1\text{-times iterating } S) \]
\[ d = A\tilde{x} - b \]
\[ d_2 = r(M^*)^{\nu}d \]

(3.1) \( v \) is determined so that \( r(M^*)^{\nu}AM^{\nu}pv = d_2 \)

\[ \overline{x} = \tilde{x} - M^{\nu}pv \]
\[ x_{i+1} = S^{(\xi_2)}(\overline{x}), \]

\( \xi_1, \xi_2, \nu \) are positive integers, \( \nu \approx 1 - 4, \xi_1 \approx 2\nu. \) The matrix \( A_\nu = r(M^*)^{\nu}AM^{\nu}p \) 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
\[
\alpha_k = \frac{(s_k, g_k)}{(A_\nu s_k, s_k)}, \\
v_{k+1} = v_k + \alpha_k s_k, \\
g_{k+1} = g_k - \alpha_k A_\nu s_k, \\
h_{k+1} = A_0^{-1}g_{k+1}, \\
\beta_k = \frac{(g_{k+1}, h_{k+1})}{(g_k, h_k)}, \\
s_{k+1} = h_{k+1} + \beta_k s_k.
\]

Let us note that the preconditioning matrix $A_0 = rAp$. Let us define the error $e(v)$ by $e(v) = v - \hat{v}$, where $\hat{v}$ 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]:

\[
\|e(v_i)\|_{A_\nu} \leq 2 \left( \frac{\sqrt{\text{cond}(A_0, A_\nu)} - 1}{\sqrt{\text{cond}(A_0, A_\nu)} + 1} \right)^i \|e(v_0)\|_{A_\nu}.
\]

4. Coarse level problem convergence analysis

Definition 4.1. For every integer $i \geq 0$ let us define
\[S_i = R(p_i).\]

Lemma 4.1. Let $K$ be a regular selfadjoint operator on a Hilbert space $H$. Then
\[
\frac{\|K^2 x\|}{\|K x\|} \geq \frac{\|K x\|}{\|x\|}
\]
for every $x \in H$, $x \neq 0$.

Proof.
\[\|K x\|^2 = (K^2 x, x) \leq \|K^2 x\| \|x\|.
\]
**Definition 4.2.** For every $x \in \mathbb{R}^n$, $i \geq 0$ let us define

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

**Remark 4.1.** Let us note that

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

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

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

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

**Proof.**

$$\|px\|_\nu \leq \|M^\nu px\|_A \leq \|M^\nu\|_A \|px\|_A = \varphi(M^\nu) \|px\|_A.$$ 

Using Lemma 4.1 we get

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

This inequality yields 2. 

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

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

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

$$Q(x) = \frac{(r(M^\nu)^* AM_\nu^\nu px, x)_2}{(rAp_x, x)_2} = \frac{(AM_\nu^\nu p_x, M_\nu^\nu px)}{(Ap_x, px)} = \frac{\|M_\nu^\nu px\|_A^2}{\|px\|_A^2} \geq \frac{\|px\|_A^2}{\|px\|_A^2}.$$ 

Therefore

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

and

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

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

$$\square$$
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 $C_\nu$ can be replaced by 1 and $c_\nu$ by $c_\nu^0$ 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 $A_\nu$-scalar product and $R(p_\nu)$ with $A$-scalar product are isometrically isomorphic, therefore

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

5. Fine Level Problem Convergence Analysis

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

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

Lemma 5.1. $AM$ is a selfadjoint operator.

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

\hfill \square

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

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

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

Remark 5.1. Lemma 5.1 implies

$$T_i = \text{Ker}(r A M^i).$$

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:

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

Proof. See [5]. \hfill \square
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 $\lambda_j$, $j = 1, \ldots, n$.

Definition 5.2. For $i \geq 0$ 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 $T_i^c$, $i \geq 0$ integer is of the form

$$\left[ \frac{c_1}{\lambda_1^i}, \ldots, \frac{c_n}{\lambda_n^i} \right]^T$$

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 $i, \xi$ be positive integers, then

$$\|M_i^\xi\|_A^2 = \frac{\sum_{j=1}^n \lambda_j^{2\xi} \frac{c_j^2}{\lambda_j^{2i}}}{\sum_{j=1}^n \frac{c_j^2}{\lambda_j^{2i}}}, \quad \text{where } c = [c_1, \ldots, c_n]^T.$$

Proof. For $x \in T_i$, we have

$$\|x\|^2_A = \sum_{j=1}^n \frac{c_j^2}{\lambda_j^{2i}}, \quad c = [c_1, \ldots, c_n]^T \in T_0^c.$
(see Lemma 5.4) and

\[ \|M^\xi x\|_A^2 = \sum_{j=1}^{n} \lambda_j^{2\xi} \frac{c_j^2}{\lambda_j^{2\nu}}. \]

\[ \square \]

**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:

\[ \frac{\|e(x_{i+1})\|_A^2}{\|e(x_i)\|_A^2} \leq \max_{c \in T_0} \frac{\sum_{j=1}^{n} \lambda_j^{2\xi} \frac{c_j^2}{\lambda_j^{2\nu}}}{\sum_{j=1}^{n} \frac{c_j^2}{\lambda_j^{2\nu}}} \cdot \]

\[ \frac{\sum_{j=1}^{n} \lambda_j^{2\xi} \frac{c_j^2}{\lambda_j^{2\nu}}}{\sum_{j=1}^{n} \frac{c_j^2}{\lambda_j^{2\nu}}} \]

**Proof.** An immediate consequence of Lemmas 5.2 and 5.5. \[ \square \]

**Remark 5.2.** If the transfer operators \( p_0, r_0 \) are well constructed then \( T_0 \) contains elements \( c = [c_1, \ldots, c_n]^T \) for which the components \( c_j \) corresponding to the small eigenvalues \( \lambda_j \), i.e. \( |\lambda_j| \approx 0 \) are large in comparison with the others. The stronger this property the smaller \( \|M^\xi_{T_0}\|_A \) is (see Lemma 5.5). For small \( \lambda_j \) we have

\[ \frac{c_j^2}{\lambda_j^{2\nu}} \gg c_j^2, \]

while for large \( \lambda_j \), i.e. \( |\lambda_j| \approx 1 \),

\[ \frac{c_j^2}{\lambda_j^{2\nu}} \approx c_j^2. \]

Therefore

\[ \|M^\xi_{T_0}\|_A^2 = \max_{c \in T_0, c \neq 0} \frac{\sum_{j=1}^{n} \lambda_j^{2\xi} \frac{c_j^2}{\lambda_j^{2\nu}}}{\sum_{j=1}^{n} \frac{c_j^2}{\lambda_j^{2\nu}}} \ll \|M^\xi_{T_0}\|_A^2 = \max_{c \in T_0, c \neq 0} \frac{\sum_{j=1}^{n} \lambda_j^{2\xi} \frac{c_j^2}{\lambda_j^{2\nu}}}{\sum_{j=1}^{n} \frac{c_j^2}{\lambda_j^{2\nu}}}, \]

can be expected.

372
Theorem 3. Let us consider the Algorithm 3.1, where
\[ S(x) = (I - \omega D^{-1}A)x + \omega D^{-1}b. \]
Let \( \xi_1, \xi_2 \geq \nu + 1 \). If the coarse level problem is solved exactly the following estimate holds:
\[ \frac{\|e(x_{i+1})\|_A^2}{\|e(x_i)\|_A^2} \leq \|M_{T_0}\|_A^{2\nu + 2}. \]

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

Proof. For every \( i \geq 1, x \in T_i, \xi \geq \nu + 1 \) if and only if \( Mx \in T_{i-1} \). Further,
\[ \frac{\|M^{\xi}x\|_A}{\|x\|_A} = \frac{\|M^{\xi}x\|_A}{\|M^{\xi-1}x\|_A} \cdot \frac{\|M^{\xi-1}x\|_A}{\|M^{\xi-2}x\|_A} \cdot \frac{\|M^{\xi-2}x\|_A}{\|x\|_A}. \]
Taking into account \( \varrho(M) < 1 \) we get
\[ \|M^{\xi}_{T_i}\|_A \leq \|M_{T_{i-1}}\|_A \cdots \|M_{T_0}\|_A. \]
Lemma 4.1 implies
\[ \|M_{T_i}\|_A \leq \|M_{T_0}\|_A, \quad i \geq 0. \]
Therefore
\[ \|M^{\xi}_{T_i}\|_A \leq \|M_{T_0}\|_A^{\nu + 1}, \]
and the usage of Lemma 5.2 completes the proof. \( \square \)

References


Authors' address: Jitka Krížková, 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.

373