Error norm estimation and stopping criteria in preconditioned conjugate gradient iterations

Article *in* Numerical Linear Algebra with Applications · June 2001

Impact Factor: 1.32 · DOI: 10.1002/nla.244

2 authors:

Owe Axelsson

Uppsala University

172 PUBLICATIONS **3,843** CITATIONS

SEE PROFILE

I. E. Kaporin

Dorodnicyn Computing Centre of RAS

66 PUBLICATIONS **525** CITATIONS

SEE PROFILE

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NIJMEGEN The Netherlands

ERROR NORM ESTIMATION AND STOPPING CRITERIA IN PRECONDITIONED CONJUGATE GRADIENT ITERATIONS

Owe Axelsson and Igor Kaporin

Report No. 0006 (March 2000)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NIJMEGEN
Toernooiveld
6525 ED Nijmegen
The Netherlands

Error norm estimation and stopping criteria in preconditioned conjugate gradient iterations

Owe Axelsson[*]
Faculty of Mathematics and Informatics, University of Nijmegen,
Nijmegen, The Netherlands

Igor Kaporin[†]
Center for Supercomputer and Massively Parallel Applications,
Computing Center of Russian Academy of Sciences,
Vavilova 40, Moscow 117967, Russia

March 30, 2000

Abstract

Some techniques suitable to the control of the solution error in the Preconditioned Conjugate Method are considered and compared. The estimation can be performed both in the course ot the iterations and after their termination. The importance of such techniques follows from the non-existence of some reasonable *a priori* error estimate for very ill-conditioned linear systems when information about the right-hand side vector is lacking. Hence, some *a posteriori* estimates are required, which make it possible to verify the quality of the solution obtained for a *prescribed* right hand side. The performance of the considered error control procedures is demonstrated using real-world large-scale linear systems arising in computational mechanics.

KEY WORDS Sparse linear systems; Symmetric Positive Definite matrices; iterative solvers; Incomplete Factorization preconditionings; Conjugate Gradients; a posteriori error estimates

1 Introduction

The present paper is concerned with further developments of robust iterative methods for the solution of linear systems with symmetric positive definite (SPD) matrices. We consider some approaches to the estimation of the solution error norms (different from the residual norm which latter can be calculated directly) for the approximations obtained from the use

[*] E-mail: axelsson@sci.kun.nl
[†] E-mail: kaporin@sci.kun.nl, kaporin@ccas.ru

of the preconditioned conjugate gradient (PCG) method. Some special aspects arising when using certain Incomplete Cholesky (IC) factorizations as a preconditioning, are outlined.

We consider the problem

$$Ax = b, \tag{1}$$

where A is a sparse $n \times n$ SPD matrix, and present some versions of (mainly known) techniques for the estimation of such error norms as the A-norm and the Euclidean norm of the PCG iteration error

$$z_k = x - x_k \tag{2}$$

defined as

$$\|z_k\|_A = \sqrt{z_k^T A z_k} \tag{3}$$

and

$$\|z_k\| = \sqrt{z_k^T z_k}, \tag{4}$$

respectively, where k is the iteration number.

As we shall see, a good preconditioning not only considerably accelerates the convergence of the PCG iterations applied to (1.1) but also provides for better chances to obtain an estimate of the solution error Euclidean norm far better than the well known standard estimate for the relative error,

$$\|x - x_k\|/\|x\| \leq C(A)\|b - Ax_k\|/\|b\|. \tag{5}$$

Let v_1 and v_n be eigenvectors of A corresponding to the smallest and the largest eigenvalues, respectively. Estimate (1.5) is sharp when the residual $b - Ax_k$ is collinear to v_1 and the exact solution x is collinear to v_n. In practice, rather the the opposite situation occurs, i.e. the resudual is dominated by higher eigenvalue modes and the solution by lower eigenvalue modes. Besides being overly pessimistic, another drawback of the latter estimate is that it involves $C(A)$, the spectral condition number of A (cf.(2.9) below) which is difficult to estimate accurately for real-life large-scale ill-conditioned problems.

In this paper we derive more accurate estimates of the iteration error in spectral (Euclidean) norm and A-norm. Some of the estimates turn out to be related to estimates in [5] based on moments and Gaussian quadratures. Our estimates, however, are derived more directly and simpler from relations which hold for the conjugate gradient method.

The paper is organized as follows. In Section 2 we recall the algorithm and some convergence results for the PCG method. In Section 3 we present the essential properties of a real-life test problem and discuss its convergence obtained with preconditionings of different quality. In Sections 4 and 5, some simple error estimates are presented, e.g., involving the pseudoresidual norm and the smallest eigenvalue of the preconditioned matrix, and the behaviour of the latter estimate is illustrated using the numerical example. In Section 6, some new lower and upper estimates for A-norm of the iteration error are presented, and their performance is demonstrated using the same numerical example. In Section 7, some general remarks concerning the use of stopping criteria, are given, while Section 8 contains short conclusive remarks.

2 Some convergence results for the PCG method

Let us introduce an SPD preconditioning matrix H, which should approximate, in a proper sense, the inverse of the coefficient matrix A. The choice of the matrix H is subject to the requirement that a vector $w = Hr$ be easily calculated for any r. For instance, one of the best choices is the approximate Cholesky preconditioning, where

$$H = (U^T U)^{-1}, \tag{1}$$

and $U^T U \approx A$ with the upper triangular matrix U being much sparser than the exact Cholesky factor of A.

The PCG iterations [1] for the solution of the problem (1) can be written as follows:

$$\begin{aligned} r_0 &= b - Ax_0, \\ p_0 &= Hr_0; \end{aligned}$$

$$\textbf{for} \;\; i = 0, 1, \dots :$$

$$\alpha_i = \frac{r_i^T H r_i}{p_i^T A p_i}, \tag{2}$$

$$x_{i+1} = x_i + p_i \alpha_i, \tag{3}$$

$$r_{i+1} = r_i - A p_i \alpha_i, \tag{4}$$

$$\beta_i = \frac{r_{i+1}^T H r_{i+1}}{r_i^T H r_i}, \tag{5}$$

$$p_{i+1} = H r_{i+1} + p_i \beta_i. \tag{6}$$

As is known [1], the following estimate for the A-norm of the error holds:

$$\frac{\sqrt{r_i^T A^{-1} r_i}}{\sqrt{r_0^T A^{-1} r_0}} \le \frac{2}{\sigma^i + \sigma^{-i}}, \qquad \sigma = \frac{\sqrt{C} + 1}{\sqrt{C} - 1}, \tag{7}$$

and the corresponding standard upper bound for the iteration number needed for the ε times reduction of the error norm $\sqrt{r_i^T A^{-1} r_i}$ is therefore

$$i_C(\varepsilon) = \left\lceil \frac{1}{2} \sqrt{C} \log \frac{2}{\varepsilon} \right\rceil, \tag{8}$$

where

$$C = C(HA) = \lambda_{\max}(HA) / \lambda_{\min}(HA) \tag{9}$$

is the spectral condition number of the preconditioned matrix HA.

There are also known some results on the superlinear convergence of the PCG method. For instance, it was shown in [8, 9] that

$$\frac{\sqrt{r_i^T H r_i}}{\sqrt{r_0^T H r_0}} \le \left(K^{1/i} - 1 \right)^{i/2}, \tag{10}$$

so that the number of iterations needed for the ε times reduction of the error norm $\sqrt{r_i^T H r_i}$ can be bounded from above (roughly, for a more precise result see [9]) as

$$i_K(\varepsilon) = \log_2 K + \log_2 \frac{1}{\varepsilon}, \tag{11}$$

where

$$K = K(HA) = \left(\frac{1}{n}\mathrm{trace}(HA)\right)^n / \det(HA) \tag{12}$$

is the so-called K-condition number of the preconditioned matrix HA, see [2]. Estimate (2.11) may give a tighter bound for the PCG iteration number as compared to (2.8) in the case when the matrix HA still has rather large spectral condition number despite the preconditioning applied, but its eigenspectrum is already strongly clustered near $\lambda = 1$. For an alternate frequently more accurate estimate explicitly involving the set of smallest eigenvalues, see [2].

The quality of a preconditioner is usually measured by the reduction of an estimate of the total computational cost which can be taken as the cost of one iteration multiplied by an estimate of the number of iterations plus the costs of the computation of the preconditioner. However, we will see that the improvement in the condition number C obtained by the preconditioning, coupled with a moderate value of $||H||$, also makes it possible to ensure good a posteriori bounds for the iteration error.

3 An illustration of the PCG convergence behavior using a test problem from computational mechanics

Throughout this paper, we will illustrate the relevance of our error estimates to the actual PCG behavior using the results obtained with the new Incomplete Cholesky 2nd order (IC2) preconditioning [10], as well as the results obtained for the standard Jacobi-CG method with $H = D_A^{-1}$. We use the notations $A = D_A + U_A^T + U_A$ for the additive splitting of the coefficient matrix into its diagonal, strictly lower triangular, and strictly upper triangular parts, respectively.

We present the results for a system with ill-conditioned matrix chosen from a set of test problems obtained from [11]. The properties of the coefficient matrix are given in Table 1. Some tests with other matrices were also done but did not give any further insight and therefore are omitted here for the sake of brevity.

Table 1: Problem set **kw114_2D**: matrix properties

n	$\mathrm{nz}(A)$	$nz(D_A + U_A)$	$\lambda_{\min}(A)$	$\lambda_{\max}(A)$	$C(A)$
92862	1254552	673707	0.176-02	0.160+07	0.912+09

Here n and $\mathrm{nz}(A)$ denote the order of the matrix A and the number of its nonzero elements, respectively.

The iterations were started from $x_0 = 0$ and continued until the stopping criterion

$$||r_k|| \le \varepsilon_0 ||b||, \qquad \varepsilon_0 = 10^{-10},$$

was satisfied for the recursively computed ”iteration residual” r_k, and the value of k is indicated in Table 1 in the column “iter”. The choice of the IC2 threshold parameters was made exactly as was recommended in [10], i.e. $\zeta = \tau^2$, $\sigma = 2\tau^2$, and the values $\tau = 0.01$ and $\tau = 0.001$ were tested in order to illustrate the performance of a weaker and a stronger preconditioning, respectively. In all cases, we used the reverse permutation of the originally generated matrices (coarsest mesh nodes form the last block, then the first refined mesh nodes form the last-but-one block, then the nodes of the second refined mesh, etc.)

In the first line of Table 2 we present the results obtained with the use of the simplest possible (and the weakest) preconditioning by the inverse diagonal of A, i.e. $H = D_A^{-1}$. In the column “fill-in” the quantity $nz(U)/nz(D_A + U_A)$ is shown which indicates the (relative) memory volume needed for the preconditioning. Finally, we present an estimate of the spectral condition number of the preconditioned matrix $C(M)$ and for the minimum and maximum eigenvalues of the preconditioned matrix $M = U^{-T}AU^{-1}$ obtained in a standard way from the extreme eigenvalues of the tridiagonal matrix T constructed from the PCG coefficients α_i and β_i, cf. Section 6 below.

Table 2: Results for the 4th **kw114_2D** matrix ($n = 92862$)

precond.	fill-in	iter.	$\lambda_{\min}(M)$	$\lambda_{\max}(M)$	$C(M)$
Jacobi	0.14	28014	0.566-08	3.372	0.596+09
IC(0.01)	1.95	1073	0.157-05	1.394	0.890+06
IC(0.001)	3.73	137	0.150-03	1.487	0.993+04

Note that since for our test problem the condition number is of the order 10^{+9}, the standard error estimate (1.1) does not guarantee more than one correct digit in any component of the approximate solution so obtained.

However, the actual results appear to be much more favorable and this situation needs some theoretical explanation, which is just the subject of the present paper.

The iterates of the PCG method were compared with the ”exact” solution x_* computed by a direct method and then improved by one iterative refinement step. The direct solution routine was based on the One-Way Dissection method [4]. In Figs.1-3 we present the behavior of the Euclidean norms of the true residual $b - Ax_k$, the iterative residual r_k, and the true error $x_* - x_k$.

First, it is clearly seen that the attained solution error is rather small in all three cases (almost independently of the preconditionings used), and is reduced more than 10^{10} times. Second, one can observe that (due to too small residual reduction parameter ε_0) the true residual $b - Ax_k$ and the iterative residual from the PCG recurrence (2.4) are different in the very end of the convergence history, and as soon as this happens, the decrease in the error norm is already stopped. Actually, such PCG behavior is rather advantageous, as otherwise

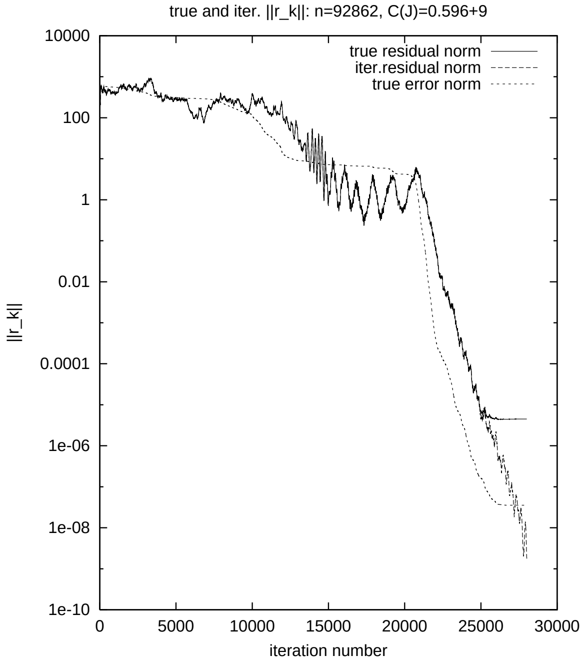


Figure 1: The behavior of residuals and iteration error: Jacobi

one can make senseless iterations up to the iteration number limit, and this will be discussed in more detail in Section 7.

4 Estimates for the relative solution error norm involving the right hand side data

The standard result is as follows. In order to estimate the relative norm of the error via the (relative) residual norm, provided that $x_0 = 0$ and therefore $r_0 = b$, it holds

$$\frac{||x - x_k||}{||x||} = \frac{||A^{-1}(b - Ax_k)||}{||A^{-1}b||} \leq ||A^{-1}||||A||\frac{||b - Ax_k||}{||b||} = C(A)\frac{||r_k||}{||r_0||}.$$

However, very often such an estimate is too pessimistic. A simple but efficient approach to derive a tighter bound is to use the *relative* condition number related to the solution vector

$$\kappa(A; x) = ||A^{-1}||\frac{||Ax||}{||x||}, \tag{1}$$

which is obviously a lower bound of $C(A)$.

Using the same argument, one can readily see that

$$\frac{||x - x_k||}{||x||} = \frac{||A^{-1}(b - Ax_k)||}{||x||} \leq ||A^{-1}||\frac{||Ax||}{||x||}\frac{||b - Ax_k||}{||b||} = \kappa(A; x)\frac{||r_k||}{||r_0||}.$$

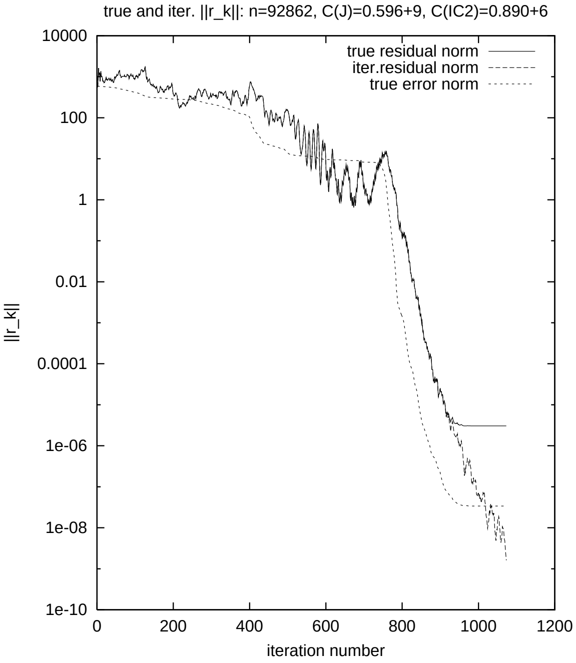


Figure 2: The behavior of residuals and iteration error: IC2(0.01)

Remark If A is an SPD matrix, then the following bound

$$\kappa(A;x) \leq \frac{1}{\sqrt{\cos^2\phi + C(A)^{-2}\sin^2\phi}},$$

is readily derived from the definition (4.1). Here ϕ is the acute angle between the right hand side $b = Ax$ and v_1, the normalized eigenvector of A corresponding to its smallest eigenvalue λ_1, so that $\cos\phi = |b^T v_1|/||b||$.

Now the question is, how to make this estimate computable. Noting that

$$\kappa(A;x) = ||A^{-1}||\frac{||b||}{||x||},$$

one must be able to estimate the quantities $||A^{-1}||$ and $||x||$. In order to get rid of the unknown quantity $||x||$, we can simply replace the definition of the relative error of the solution, introducing it as $||x - x_k||/||x_k||$. Clearly, using the inverse triangle inequality, $||x|| \geq ||x_k|| - ||x - x_k||$, one finds

$$\frac{||x-x_k||}{||x||} \leq \frac{||x-x_k||/||x_k||}{1 - ||x-x_k||/||x_k||}.$$

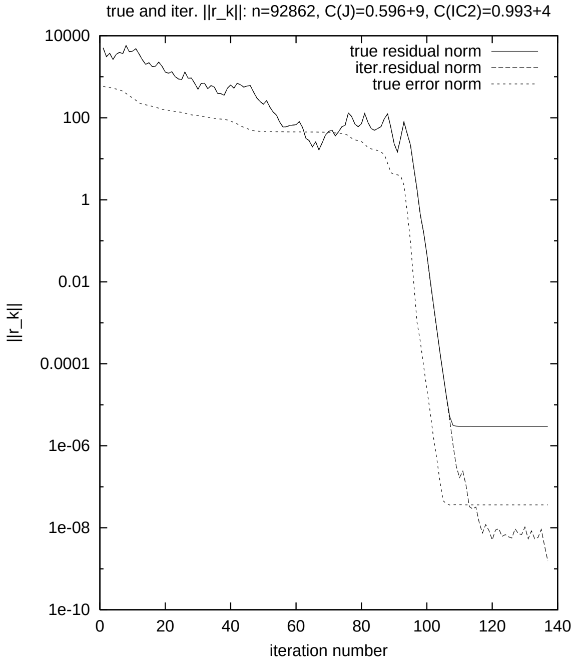


Figure 3: The behavior of residuals and iteration error: IC2(0.001)

Hence, this is asymptotically equivalent to the above definition (as x_k tends to x), and the modified estimate takes the form

$$\begin{aligned}\frac{||x-x_k||}{||x_k||} &\leq \frac{||A^{-1}||}{||x_k||}||b-Ax_k|| \\ &= \kappa(A;x_k)\frac{||b-Ax_k||}{||Ax_k||},\end{aligned} \tag{2}$$

so that the relative residual is also properly redefined.

However, the quantity $||A^{-1}||$ can be (relatively) large for ill-conditioned matrices. Also, its estimation can take (at least) no less computational effort than the solution of the linear system itself. Therefore, further we will discuss how to estimate the error norm $||x - x_k||$ using less straightforward approaches which use some information accumulated in the course of the PCG iteration, thus presenting an alternative to the estimate

$$||x-x_k|| \leq ||A^{-1}||||b-Ax_k|| \tag{3}$$

used above.

5 Some simple solution error estimates involving pseudoresiduals

In this section we present some estimates using only a small portion of the data generated by the PCG iterations, such as an estimate for the quantity

$$\mu_1 = \lambda_{\min}(HA). \tag{1}$$

After terminating the CG iterations, let us calculate the true residual

$$r_k^{\mathrm{true}} = b - Ax_k = Az_k$$

and the corresponding pseudo-residual (i.e. the preconditioned residual)

$$w_k^{\mathrm{true}} = Hr_k^{\mathrm{true}}.$$

Here we recall that if the termination of the PCG iterations is subject to the standard stopping criterion $\|r_k\| \leq \varepsilon \|r_0\|$ with too small ε, then due to round-off errors one obtains $r_k \neq r_k^{\mathrm{true}}$, cf. Section 7 below.

For the error, one has

$$\|z_k\| = \|(HA)^{-1}HAz_k\| = \|(HA)^{-1}w_k^{\mathrm{true}}\| \leq \|(HA)^{-1}\|\|w_k^{\mathrm{true}}\|.$$

In general, $\|(HA)^{-1}\| \geq \mu_1^{-1}$ but, supposing that

$$\|(HA)^{-1}\| \approx \mu_1^{-1},$$

one obtains an *a posteriori* error estimate in the form

$$\|z_k\| \lessapprox \mu_1^{-1}\|w_k^{\mathrm{true}}\|. \tag{2}$$

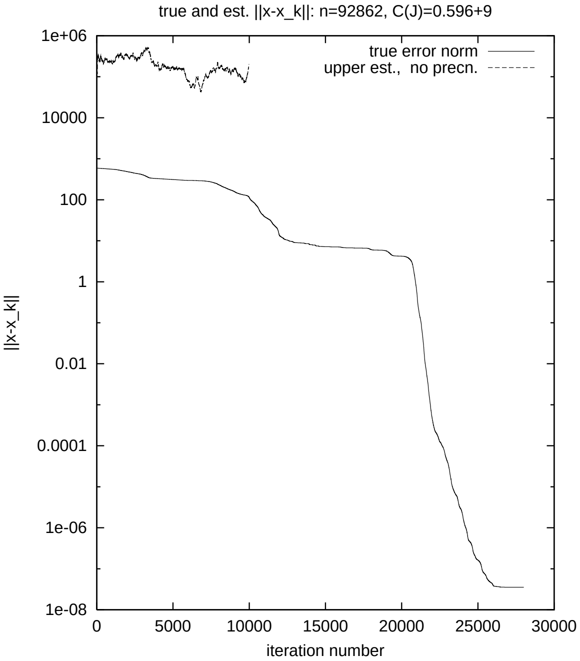
Though being not strictly justified, this estimate proved to be sufficiently accurate for our purposes. Note that it can present a very essential improvement as compared to the standard estimate

$$\|z_k\| \leq \lambda_1^{-1}\|r_k^{\mathrm{true}}\| \tag{3}$$

used in the preceding section, where

$$\lambda_1 = \lambda_{\min}(A), \qquad \lambda_1^{-1} = \|A^{-1}\|.$$

This is illustrated below in Figs. 4-6, where, for the same numerical experiments as described above in Section 3, we present the plots of the true error norm $\|x_k - x_*\|$ and its two upper estimates, (5.2) and (5.3). One can readily see that, being almost coinciding in the case of the weakest Jacobi preconditioning, these estimates behave rather differenly for stronger preconditionings, and the error estimate (5.2) based on the use of the pseudoresidual norm gives considerably better upper bound for the true error norm as compared to its unpreconditioned counterpart (5.3). Also, it is clearly seen from the comparison of Fig.5 and Fig.4,


true and est. ||x-x_k||: n=92862, C(J)=0.596+9
true error norm
upper est., no precn.
||x-x_k||
1e+06
10000
100
1
0.01
0.0001
1e-06
1e-08
0
5000
10000
15000
20000
25000
30000
iteration number