Mathematical Modelling and Analysis 2005. Pages 349–354
Proceedings of the 10th International Conference MMA2005&CMAM2, Trakai
© 2005 Technika ISBN 9986-05-924-0

SPEEDING UP THE CONVERGENCE OF SIMPLE GRADIENT METHODS

O.KACHER[1], A. BOLTNEV[2] and N. KALITKIN[3]

[1,2] *Moscow Institute of Electronic Engineering*

K-498, Moscow (Zelenograd), 103498, Russia

[3] *Institute for Mathematical Modelling RAS*

Miuskaya pl.4A, Moscow, Russia, 125047

E-mail: olya_kacher@mail.ru

Abstract. This paper introduces a simple gradient method for solving of linear systems resulting from approximation of elliptic PDEs. The method has the same rate of convergence as conjugate gradient method but it is as simple and reliable as the minimal residual method.
Key words: gradient descent, minimal residual method, elliptic PDE

1. Introduction

Solving multidimensional differential equations of the elliptic type is very important problem in the theory of numerical methods. But finite-difference approximation of such equations leads to linear algebraic system of large dimension $M \sim 10^4 - 10^6$. The resulting matrices of these systems are very sparse: each row contains only five or seven nonzero elements. Such systems are usually solved by iterative methods. The most simple and reliable of them are the steepest descent method and minimal residual method [3]. But convergence of these methods is too slow. If the matrix spectrum has boundary values $\lambda_{\min}$, $\lambda_{\max}$ then number of iterations $K \sim \mu = \lambda_{\max}/\lambda_{\max}$. In this paper the modification for the minimal residual method is suggested. The modified method is as simple and reliable as the original one. But its convergence speed is much better $K \sim \sqrt{\mu}$, that corresponds to the speed of the conjugate gradient method.

2. The Gradient Descent Method

Both the steepest descent method and the minimum residual method have an interesting feature (see [2]): the residual diminishes very fast during the first

several iterations, however, the convergence rate reduces to the theoretical prediction at the later steps. This is illustrated in Fig. 1, where the horizontal axis shows the index k of each iteration step, while the vertical axis shows (on the logarithmic scale) the Euclidean norm of the residual.

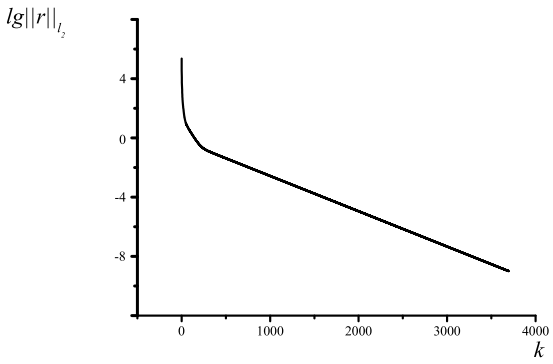

Figure 1. Euclidean norm of the residual dependence on the iteration number k.

In practical computations, it is desirable to keep the high rate of convergence that characterizes the first steps through the entire iterative process.

3. Relaxation Method

Let us consider the linear system

$$Ax = b \tag{3.1}$$

We implement one step of the iterative process in anti-gradient direction, according to the formula:

$$x^{(k+1)} = x^{(k)} + c\tau_k r^{(k)} \tag{3.2}$$

where k is the iteration number, $\tau = (r, Ar)/(Ar, Ar)$ is the step in the standard minimal residual method, $r = b - Ax$ is the residual of the system, c is a multiplicative constant, called the relaxation factor.

The main aim of this paper is to find optimal value for the constant c when the convergence speed of the gradient method is maximal. It is well known that the gradient method with $c = 2$ diverges (see [2]). It is also clear that the number of steps greatly increases as $c \to 0$ since the step size tends to zero. The typical dependence of iterations number K on c is shown in Fig. 2.

One can see that this dependence has a deep minimum near $c \approx 0.9$. The series of numerical investigation were carried out and it was shown that by

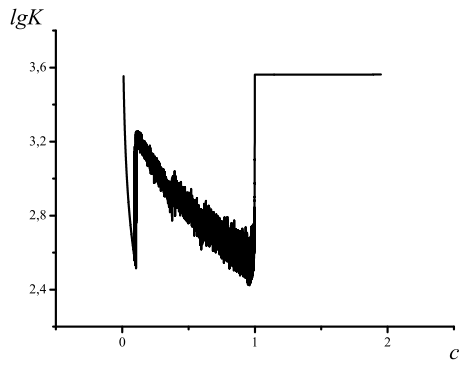

Figure 2. Typical dependence of iterations number K on coefficient c.

setting $c = 0.9$ in (3.2) one can increase convergence speed up to $K \sim \sqrt{\mu}$. The typical test matrices we have used in numerical experiments are the following: matrices resulting from the approximation of one- and two-dimensional elliptic PDEs on the uniform grid, matrix with eigenvalues uniformly distributed between the smallest eigenvalue $\lambda_{\min} = \alpha$ and the largest eigenvalue $\lambda_{\max} = \beta$ and some others.

4. Numerical Experiments

The convergence rate of the new method for one test problem is illustrated in Fig. 3.

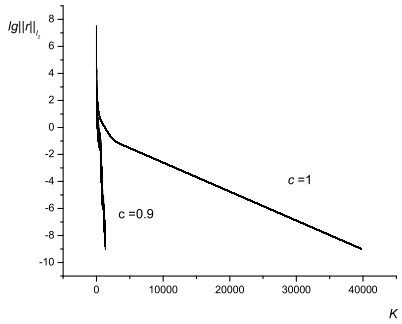

Figure 3. The convergence rate of the method with different parameter values.

Here, the error is represented as a function of the step number for the original minimal residual method ($c = 1$) and for the method (3.2) with the

recommended value $c = 0.9$. It is seen that both curves start in a similar way; this is the initial stage of anomalous convergence. The curve corresponding to $c = 0.9$ preserves this fast convergence, while, for the curve with $c = 1$, the convergence becomes slow.

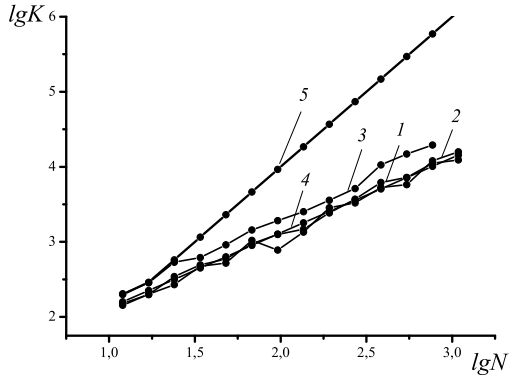

Figure 4. Dependence of the iteration number K on N, the number of grid points per space direction, in a double-logarithmic scale.

Fig. 4 illustrates the dependence of the number of iterations K on the number of grid points per space direction N in a double-logarithmic scale. In all tests, method (3.2) was used with $c = 0.9$. Curves 1, 2, and 3 correspond to the test matrices, resulting from the approximation of the one-, two- and three-dimensional elliptic PDEs on the uniform grid, respectively. Curve 4 is the theoretical curve for the conjugate gradient method, and curve 5 corresponds to the original minimal residual method. It is evident from Fig. 4 that for all test problems the number of steps of the new iterative method is practically a linear function of N and all curves pass near the theoretical line for the conjugate gradient method. For comparison purposes, we show how much slower the minimal residual method is.

5. Descent Trajectories

In the two-dimensional case, trajectories of gradient descent can be visualized. This is done in Fig. 5 for the system with the matrix A and the right-hand side vector b given by:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 100 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

hence, the exact solution is given by $x_e = \begin{pmatrix} 1 \\ 0.01 \end{pmatrix}$. Thus, Fig. 5 is the graphical representation of the process of solving system (3.1) by method (3.2) in the case $N = 2$.

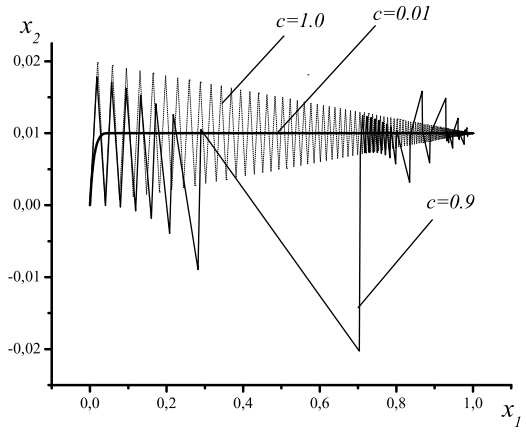

Figure 5. A solution of the system (3.1) obtained by method (3.2) in a two-dimensional case.

The steepest descent curve is replaced in Fig. 5 by the trajectory with $c = 0.01$.The classical steepest descent curve corresponds to the trajectory with $c = 1$. For the descent with $c = 0.9$, most of the steps are about the same as those for $c = 1$; however, once in a while, there occur jumps that bring us much closer to the solution. Observe that the most beneficial jump happens at the moment when the current approximation is close to the steepest descent curve. In the pure steepest descent method, no such jumps occur.

6. Results

Method (3.2) with $c = 0.9$ was successfully applied to solve 2D elliptic PDE [1], which describes a stationary problem for conductivity or electric fields in anisotropic medium:

$$\begin{cases} \dfrac{\partial^2 u}{\partial x^2} + \sigma \dfrac{\partial^2 u}{\partial x \partial y} + \dfrac{\partial^2 u}{\partial y^2} = 0, \quad 0 \le x, y \le 1, \\ u(0, y) = u(1, y) = u(x, 0) = u(x, 1) = 0. \end{cases} \tag{6.1}$$

Experiments have shown that $K \sim \sqrt{\mu}$ for any important in practice anisotropy parameter values $|\sigma| \le 2$. Therefore the proposed simple iterative

method is reliable for anisotropic problems and is recommended for practical usage.

7. Conclusions

We have proposed a simple one-step method of gradient descent that have the same rate of convergence as the conjugate gradient method. The high efficiency of the proposed methods is confirmed by numerous computational experiments. Choosing the values of the parameters does not require knowledge of the spectrum of the matrix. For matrices with elliptic spectra (and some other types of spectrum), we experimentally investigated the methods of under-relaxation and found the values of the relaxation factor that greatly improve the performance of these methods.

This work is supported by RFBR (projects 05-01-00152, 05-01-00144) and presidential supporting program for scientific schools (project 1918.2003.1)

References

[1] E.A.Al'shina, A.A.Boltnev and O.A.Kacher. Gradient methods with improved convergence rate. *Computational Mathematics and Mathematical Physics*, **45**(3), 356 – 365, 2005.

[2] V.V. Ermakov and N.N. Kalitkin. The two-step gradient descent. *Zh. Vychisl. Mat. Mat. Fiz.*, **20**(2), 1040 – 1045, 1980. (in Russian)

[3] D.K. Faddeev and V.N. Faddeeva. *Computational Methods of Linear Algebra, 2nd ed.* Fizmatgiz, Moscow, 1963. (in Russian)