Univ. u Novom Sadu
Zb. Rad. Prirod.–Mat. Fak.
Ser. Mat. 22, 2 (1992), 1-10

Review of Research
Faculty of Science
Mathematics Series

ON A NONSTATIONARY MODIFICATION OF AN ITERATIVE METHOD [1]

Dragoslav Herceg, Nataša Krejić
Institute of Mathematics, University of Novi Sad
Trg Dositeja Obradovića 4, 21000 Novi Sad, Yugoslavia

Abstract

The numerical solution of the system of linear equations $Ax = b$, in the sense $\min_x ||Ax - b||$, where the matrix A can be even rectangular is considered. By introducing some parameters or continuous function the Method of Optimal Basic Descent is accelerated. Sufficient convergence conditions for the modified method are given. Numerical examples confirm the eficiency of the proposed algorithm.

AMS Mathematics Subject Classification (1991): 65F10
Key words and phrases: iterative method.

1. Introduction

In this paper we shall investigate the numerical solution of the system

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

where A is $n \times n$ real matrix and b is a column vector from R^n. The only condition on system (1) is that the solution exists, which means that matrix A can even be singular. Such systems often arise in economic models and

[1] This research was partly supported by Science Fund of Serbia, grant number 0401A, through Matematički institut

have none of the usual features of physical systems, such as symmetry, diagonal dominance or nonnegativity, so research in this area is concentrated on determing the convergence conditions as generally as possible.

The solution of system (1) is investigated in the sense

$$\min_x ||Ax - b||, \tag{2}$$

where $||\cdot||$ is Euclidean vector norm, and that is equivalent to the solution of system

$$A^t Ax = A^t b. \tag{3}$$

Algorithm of Optimal Basic Descent was proposed by Pospelov, [6]. That algorithm generates a class of methods, but their convergence is often very slow. We propose some modifications of that algorithm which can accelerate convergence significantely.

One modification is given by introducing some parameters. Nonstationary modification is also considered for one special class of matrices. In both cases convergence conditions are given and theoretical results are confirmed by numerical experiments.

2. The Method of Optimal Basic Descent

Let $\{w_i\}, i = 1, \ldots, n$ be a base for the space R^n, which satisfies the conditions $(w_i, w_j) = \delta_{ij}$, $i, j = 1, \ldots, n$, where δ_{ij} are Croneker symbols. We also request that $Aw_j \neq 0$, for at least one j.

For solving problem (2) following algorithm is proposed, [6].

$$x^0 = 0, r^0 = b \tag{4}$$

$$j_k = arg \max_j \{ \frac{|(r^k, Aw_j)|}{||Aw_j||} : Aw_j \neq 0 \} \tag{5}$$

$$r^k = b - Ax^k$$

$$x^{k+1} = x^k + \frac{(r^k, Aw_j)}{||Aw_{j_k}||^2} w_{j_k}. \tag{6}$$

If there is more than one j which satisfies (5) then the choice of j_k is arbitrary.

For the algorithm defined by (4) - (6) the following convergence theorem has been proved.

Theorem 1. *[6] The sequence generated by iterations (4) - (6) is convergent, its limit point, x^∞, is the solution of system (3) and the inequality*

$$||x^k - x^\infty|| \le Cq^k,$$

where $0 \le q < 1$ and C are some constants, which depend on the set $\{w_i\}$ and matrix A, is satisfied.

Using different sets of $\{w_i\}$ we can get a class of methods, which we shall discuss later.

3. Modifications of the Method of Optimal Basic Descent

First, we shall prove one lemma, which is the generalization of the results from [6].

Lemma 1. *Let $g_1, \dots, g_m$ be a set of vectors from R^n with the property $||g_i|| = 1, i = 1, \dots, m$. Then for every vector $x \in L(g)$, where $L(g)$ is a linear space generated by the vectors $g_i, i = 1, \dots, m$, and for every continuous function $\varphi(x) \in (0,2)$, inequality*

$$||x - \varphi(x)(x, g_k)g_k|| \le q||x||, \; q < 1, \; k = arg\max_i |(x, g_i)|, \tag{7}$$

is satisfied.

Proof. Suppose that there exists no $q < 1$, so that inequality (7) is satisfied. That means that

$$||x - \varphi(x)(x, g_k)g_k|| \ge ||x||$$

because the functions $||x - \varphi(x)(x, g_k)g_k||$ and $||x||$ are continuous functions. So,

$$||x - \varphi(x)(x, g_k)g_k||^2 \ge ||x||^2$$

which is equivalent to

$$(x, x) - 2\varphi(x)(x, g_k)^2 + \varphi(x)^2(x, g_k)^2(g_k, g_k) \ge (x, x).$$

As $||g_k|| = 1$, we get

$$(-2\varphi(x) + \varphi(x)^2)(x, g_k)^2 \geq 0.$$

By the assumption, $k = arg\max_i |(x, g_i)|$, and $x \in L(g)$, so $(x, g_k)^2 > 0$, which gives us

$$\varphi(x)(\varphi(x) - 2) \geq 0. \tag{8}$$

If we choose $\varphi(x) \in (0,2)$ then inequality (8) cannot be satisfied, and we get contradiction with our assumption, so inequality (7) is satisfied, which completes the proof.

Now, consider the following algorithm.

$$x^0 = 0, r^0 = b \tag{9}$$

$$j_k = arg\max_j \{\frac{|(r^k, Aw_j)|}{||Aw_j||} : Aw_j \neq 0\} \tag{10}$$

$$r^k = b - Ax^k$$

$$x^{k+1} = x^k + \varphi(x^k)\frac{(r^k, Aw_j)}{||Aw_{j_k}||^2} w_{j_k}. \tag{11}$$

Using Lemma 1 convergence theorem for algorithm (9) - (11) can be given.

Theorem 2. *The sequence generated by iterations (9) - (11) is convergent, its limit point, x^∞, is the solution of system (3) and the inequality*

$$||x^k - x^\infty|| \leq Cq^k,$$

where $0 \leq q < 1$ and C are some constants, which depend on the set $\{w_i\}$ and the matrix A, is satisfied.

Proof. As $||Aw_{j_k}|| \neq 0$, it is obvious that there exists a constant δ such that

$$||Aw_{j_k}|| \geq \delta, \text{ for every } k.$$

Using the definition of iterative method (11) we get

$$b - Ax^{k+1} = b - Ax^k - \varphi(x^k)\frac{(r^k, Aw_{j_k})}{||Aw_{j_k}||^2} Aw_{j_k},$$

i.e.,

$$r^{k+1} = r^k - \varphi(x^k)\frac{(r^k, Aw_{j_k})}{||Aw_{j_k}||^2}Aw_{j_k}.$$

Now, with $g_i = \frac{Aw_i}{||Aw_i||}$, Lemma 1 can be applied and we achieve the relation

(12) $$||r^{k+1}|| \leq q||r^k||.$$

From (11) and (12) it follows that

$$||x^{k+1} - x^k|| \leq ||\varphi(x^k)||\frac{||r^k||}{||Aw_{j_k}||},$$

which means that

$$||x^{k+1} - x^k|| \leq C\frac{q^k}{1-q},$$

where C is a constant depending on A, r^0, δ and$\varphi(x^k)$. Using this fact we can conclude that the sequence $\{x^k\}$ is Cauchy sequence and hence convergent and its limit point, x^∞ solves system (3). This completes the proof.

Remark. The vectors $\{w_i\}$ need not be orthonormal and the number of these vectors could be less that n. The same remark stays valid for the original Method of Optimal Basic Descent.

It should be noticed that that proposed algorithm is very effective in such problems where we have a priori information about the solution's structure, for example if we know the basic functions with which the solution can be represented.

For different sets of $\{w_i\}$ and functions $\varphi(x)$ we can generate various methods for solving problem (2).

Let us denote by $e_i, i = 1, \ldots, n$ the usual orthonormal base for the space R^n, by $a_i, i = 1, \ldots, n$ columns of the matrix A, and by $b_i, i = 1, \ldots, n$ rows of the matix A.

We are going to consider three sets of $\{w_i\}$.

1. $w_i = e_i, i = 1, \ldots, n$.

In such a case iteration rule (11) becomes

(13) $$x^{k+1} = x^k + \varphi(x^k)\frac{(r^k, a_{j_k})}{||a_{j_k}||^2}e_{j_k}.$$

2. $w_i = a_i, i = 1, \dots, n.$

Now, we get iterations

(14) $$x^{k+1} = x^k + \varphi(x^k)\frac{(r^k, Aa_{j_k})}{||Aa_{j_k}||^2}a_{j_k},$$

i.e., a sort of column iterations, depending on the function φ.

3. $w_i = b_i, i = 1, \dots, n.$

For the vectors w_i the method is given by

(15) $$x^{k+1} = x^k + \varphi(x^k)\frac{(r^k, Ab_{j_k})}{||Ab_{j_k}||^2}b_{j_k}.$$

This can be considered as row relaxation method.

The simpliest way to define the function $\varphi(x)$ is

$$\varphi(x) = \beta, \ \ \beta \in (0, 2).$$

Using such function φ we get convergent column and row relaxation methods. This case was also analysed in paper [2]. Numerical examples show that the convergence of relaxed iterations is usually much faster than the convergence of iterations with $\beta = 1$, i.e., Optimal Basic Descent.

Here, we give numerical results for the system $Ax = b$, where A is the matrix of random numbers and vector b is choosen such that the exact solution is vector x with all components $x_i = 1, i = 1, \dots n$. Figure 1 shows the graph of $-Log(eps)$ versus β and k, where $eps = ||x^k - x||$, for the column iterations given by (14). In that case $n = 10$.

4. Nonstationary Modifications

If the function $\varphi(x^k)$ is not a constant, then the iterations given by (13), (14) and (15) become nonstationary. Nonstationary iterative methods have obvious attractions because they can be adapted to handle solution automatically, without any interventions by the user. This property is very important in economic models because systems which arise in these models can be impossible to see and analyse.

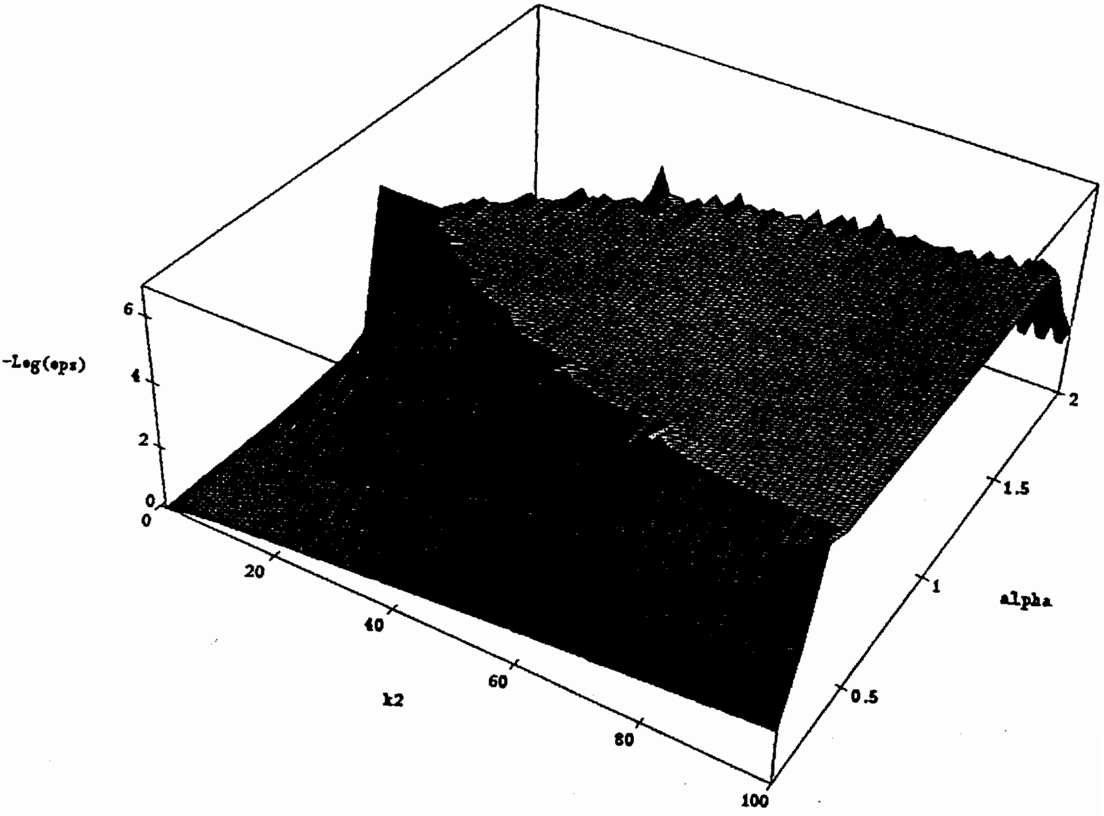

Figure 1

For these methods the function $\varphi(x)$ could be written as

$$\varphi(x^k) = \beta_k,$$

where $\beta_k \in (0,2)$, and function φ is continuos.

We are going to give one nonstationary modification of the Method of Optimal Basic Descent for a class of strictly diagonal dominant matrices. Such modification was given in the paper [1] for one row relaxation method. First of all, we need the following theorem.

Theorem 3. *[7] Let A be a strictly diagonal dominant matrix and let*

$$\alpha_0 = \min\{|a_{ii}| - \sum_{j\neq i} |a_{ij}|,\ i = 1, \ldots, n\}.$$

Then

$$||A^{-1}||_\infty \leq \frac{1}{\alpha_0}.$$

As the convergence theorem for modified method assumes that $\varphi(x^k) \in (0,2)$, using Theorem 3, we are able to consider iterations with

$$\varphi(x^k) = \beta_k = 2 - \omega + \omega f_k, \tag{16}$$

where $\omega \in (0,2)$ and

$$f_k = \frac{\alpha||x^k - x^{k-1}||_\infty}{||r^k||_\infty + ||r^{k-1}||_\infty}. \tag{17}$$

Theorem 4. *Let A be a strictly diagonal dominant matrix. Then for $\varphi(x^k) = 2 - \omega + \omega f_k$, where $\omega \in (0,2)$, f_k is defined by (17), and $\alpha < \alpha_0$, iterative method (9) - (11) is convergent.*

Proof. To apply Theorem 2 we need to prove that $f_k < 1, k = 1, 2, \ldots$ Using the definition of residual vectors, $r^k = b - Ax^k$, $r^{k-1} = b - Ax^{k-1}$, we get

$$r^k - r^{k-1} = Ax^{k-1} - A^k$$

$$x^k - x^{k-1} = -A^{-1}(r^k - r^{k-1}),$$

so

$$||x^k - x^{k-1}||_\infty \leq ||A^{-1}||_\infty ||r^k - r^{k-1}||_\infty.$$

Now, by Theorem 3, we have

$$||A^{-1}||_\infty \leq \frac{1}{\alpha_0} < \frac{1}{\alpha},$$

and it follows that

$$||x^k - x^{k-1}||_\infty \leq \frac{1}{\alpha}(||r^k||_\infty + ||r^{k-1}||_\infty),$$

what gives us

$$f_k < 1, k = 1, 2, ...$$

To finish the proof we simply apply Theorem 2.

Applying (17) and (18) on the iteration rule (11) in case 2 and 3 we get the nonstationary column relaxations

$$x^{k+1} = x^k + (2 - \omega + \omega f_k)\frac{(r^k, Aa_{j_k})}{||Aa_{j_k}||^2}a_{j_k},$$

and the nonstationary row relaxations

$$x^{k+1} = x^k + (2 - \omega + \omega f_k)\frac{(r^k, Ab_{j_k})}{||Ab_{j_k}||^2}b_{j_k}.$$

To illustrate this modification we present one test problem. Consider the system $Ax = b$, where

$$A = \begin{bmatrix} 4 & -1 & & & & \\ -1 & 4 & -1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & & -1 & 4 & -1 \\ & & & & -1 & 4 \end{bmatrix}, b = [3, 2, \ldots, 2, 3]^t.$$

The exact solution is $x = [1, 1, \ldots, 1]^t$. The error vector is defined by $eps(k) = x^k - x$. Termination criterium was $||eps(k)|| < 10^{-3}$. Results for different values of ω are given in Table 1. The last row in Table 1 shows the number of iterations acheived by the Method of Optimal Basic Descent. In this case the dimension of the system was $n = 10$.

ω	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.8	OBD
k	356	188	145	207	238	225	274	359	461	913

Table 1

References

[1] Herceg, D.: A note on an iterative process for solving linear systems of equations, Univ. u Novom Sadu Zb. Rad. Prirod.-Mat. Fak. Ser. Mat. 20,1(1990), 1-10.

[2] Herceg, D., Krejić, N.: A Remark on Iterative Method for a System of Linear Equations, Bull. Appl. Math. 779(1992)(LXI), 13-22.

[3] Dax, A.: Line Search Acceleration of Iterative Methods, Linear Algebra Appl. 130(1990), 43-63.

[4] Dax, A.: On Hybrid Acceleration of a Linear Stationary Iterative Process, Linear Algebra Appl. 130(1990), 99-110.

[5] Dax, A., Berkowitz, B.: Column Relaxation Methods for Least Norm Problems, SIAM J. Sci. Stat. comput. 11,5(1990), 975-989.

[6] Pospelov, V.V.: Metod optimalnogo spuska po bazisu dlya resheniya vyrozhdenuyih sistem lineĭnyh algebraicheskih uravneniĭ, Zh. Vychisl. Mat. Mat. Fiz. 31,7(1991), 962-969.

[7] Varah, J.M.: A lower bound for the smallest singular value of a matrix, Linear Algebra Appl. 11(1975), 3-5.

REZIME

O NESTACIONARNOJ MODIFIKACIJI ITERATIVNOG POSTUPKA

Posmatra se numeričko rešenje sistema linearnih jednačina $Ax = b$, u smislu $\min_x ||Ax - b||$, pri čemu matrica A može biti i pravougaona. Uvodjenjem parametra ili neprekidne funkcije ubrzan je metod optimalnog pada po bazi. Dati su dovoljni uslovi za konvergenciju tako modifikovanog postupka, a efikasnost modifikacije je potvrdjena numeričkim primerima.

Received by the editors April 8, 1992.