A Genetic Algorithm for Solving General System of Equations

Győző Molnárka, Edit Miletics

Department of Mathematics, Széchenyi István University, Győr, Hungary
molnarka@sze.hu, miletics@sze.hu

Abstract: For solving linear system of equations is known several algorithms. Iteration algorithms are recommended for the large linear systems with sparse matrix. But in the case of general n x m matrices the classic iterative algorithms are not applicable with a few exceptions. For example in some cases the Lanczos type algorithms are adequate. The algorithm presented here based on the minimization of residuum of solution and it has some genetic character. Therefore this algorithm seems to be applicable for construction of parallel algorithm. Here we describe a sequential version of proposed algorithm and give its theoretical analysis. Moreover we show some numerical test results of the sequential algorithm.

Keywords: linear system of equations, iterative algorithms, genetic algorithms, parallel algorithm.

1 Introduction

Let A be a general nxm matrix. The basic problem is to solve the following linear system of equations:

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

where $u \in R^m$ and $b \in R^n$ are the solution and the given right hand side vector. The existence and uniqueness of the solution of (1) can be determined from matrix A and the vector b. Theoretically the Gaussian or Gauss-Jordan elimination algorithm is appropriate tool to solve the system (1) and decide the question of solvability. But practically for large systems when we use floating point arithmetic, these direct algorithms are inapplicable. For these cases the iteration algorithms are suitable. Effective iteration algorithms are known for symmetric positive definite linear systems called Hermitian. Most of known iterative algorithms in general form can be written in the form of:

$$u^{(n)} = Gu^{(n-1)} + k, \;\; n = 1, 2, \ldots, \tag{2}$$

where G and k are such matrix and vector that (2) for its stationary solution must be equivalent with (1), see ([1]). The most effective algorithms for such system are the preconditioned conjugate gradient (CG) ones see ([2]). These iterative algorithms can be applied for general nonsymmetric linear systems too if we solve the following normal system

$$A^T Ax = A^T b = v \tag{3}$$

instead of the original one. A disadvantage of this approach is that the resulting linear system (3) for matrices with full rank will be Hermitian ones but its condition number will be the square of the original condition number, therefore the convergence will be very slow. For general non-Hermitian linear systems instead of generalization of some variant of the CG algorithms one of the most successful scheme is the generalized minimal residual algorithm (GMRES) see ([5]) and the biconjugate gradient algorithm (BCG) see ([6]). A more effective approach was suggested by Freund and Nachtigal ([7]) for the case of general nonsingular non-Hermitian systems which calls the quasi-minimal residual algorithm (QMR).

In the following we describe an iterative minimal residual algorithm which is slightly different from the above ones, but this difference can be very important for further development of these algorithms for parallel implementation.

2 An Iterative Minimal Residual Algorithm

It is known, that the most of the iterative algorithms for the solution of linear systems based on some minimization algorithm see ([2]). The normal system (3) can be obtained by the least square minimization in the following way. We have to solve the following problem:

$$\min_{x \in R^n} \|Ax - b\|_2^2 = \min_{x \in R^n} (Ax - b,\ Ax - b) = \min_{x \in R^n} (r,\ r)\,, \tag{4}$$

where $r = Ax - b$ is the residual belonging to the vector x.

The normal system (3) can be derived easily from (4) by a simple calculation. More precisely we obtain, that the necessary condition of the existence and uniqueness of the solution of (4) is the fulfilment of (3). A sufficient condition for the uniqueness is the Hermitian property of the normal matrix $A^T A$. For general non-Hermitian matrices this condition will not fulfil in general.

To solve the problem (4) we chose the following way. One possible algorithm can be obtained from the observation formulated by the following theorem.

Theorem 1 Let $A \in R^m \to R^n$ and $b \in R^n$ be an arbitrary matrix and vector. Moreover let $x^\alpha \in R^m$ and $x^\beta \in R^m$ be arbitrary, but such different vectors for which $A(x^\alpha - x^\beta) \neq 0$. Let us introduce the following notations: $r^s = Ax^s - b,\ s = \alpha, \beta$, and

$$r^{\alpha\beta} = cr^\alpha + (1-c)r^\beta, \qquad x^{\alpha\beta} = cx^\alpha + (1-c)x^\beta, \tag{5}$$

where $c \in R$. It is easy to see, that $Ax^{\alpha\beta} - b = r^{\alpha\beta}$.

Then the solution of the constrained minimization problem of (4) along the line defined by the vectors x^α and x^β is the vector $x^{\alpha\beta}$ with c, where

$$c = \frac{(r^\beta - r^\alpha,\ r^\beta)}{\left\| r^\alpha - r^\beta \right\|^2} \tag{6}$$

Moreover

$$\left\| r^{\alpha\beta} \right\| < \min_{\alpha,\beta} \left\{ \left\| r^\alpha \right\|, \left\| r^\beta \right\| \right\},$$

where we use the Euclidean norm.

Proof of Theorem 1 The constrained minimization problem of (4) with (5) is a one-dimensional problem. So we have to solve the following one-dimensional problem:

$$\min_c f(c) = \min_c (cr^\alpha + (1-c)r^\beta,\ cr^\alpha + (1-c)r^\beta) = \min_c \left\| r^{\alpha\beta}(c) \right\|^2. \tag{7}$$

From here the result (6) as the necessary condition can be obtained after a simple calculation. The sufficient condition from (7) is:

$$\frac{d^2 f(c)}{dc^2} = 2\left\| r^\alpha - r^\beta \right\|^2 > 0, \tag{8}$$

which condition is fulfilled if r^α and r^β are different vectors.

The meaning of the theorem 1 is the following: If we have some approximate solution x^α for the problem (1), then by arbitrary x^β vector satisfying the conditions of theorem (1) we can get a better approximate solution in the form of $x^{\alpha\beta}$ vector, where the appropriate constant c is defined by (4).

Remark: As against the classical iterative algorithms, such as gradient and the conjugate gradient ones, here the directions of minimization are chosen by chance. The algorithm can be continued in every step by choosing a new independently

chosen arbitrary x^{β} vector. This is the main new property of the proposed algorithm, because this property allows, that in parallel arbitrary number of approximation vector can be generated independently, and all these vectors can be used to improve the earlier best approximation of the solution.

The realization of the theoretical results as an algorithm can be made in several ways. Here we describe a simple (maybe the most simple) version of a sequential algorithm.

2.1 The Algorithm 1

Using the results of the Theorem 1 we can formulate an algorithm, which generate an approximate solution sequence x^k, $k = 1,2,3,...$ and in parallel its residual vectors r^k, $k = 1,2,3,...$.

1 Let x^1 be an arbitrary vector and ε the error tolerance.

2 Calculate $r^1 = Ax^1 - b$.

3 Generate an arbitrary vector , x^2 such that $r^1 - r^2 \neq 0$.

4 $$c^{12} := \frac{(r^2 - r^1,\ r^2)}{\left\| r^1 - r^2 \right\|^2},$$

5 Calculate the new $x^{12} := c^{12}x^1 + (1-c^{12})x^2$ and $r^{12} := c^{12}r^1 + (1-c^{12})r^2$ vectors.

6 $x^1 := x^{12}$ és $r^1 := r^{12}$

7 If $\left\| r^1 \right\| < \varepsilon$ then go to 8. else go to 3.

8 The approximate solution is x^1, end of algorithm.

The algorithm 1 is the simplest one which can be formulated by the result of Theorem 1. Therefore the convergence of this algorithm is not better than the convergence of the classical ones. The novelty of this algorithm is not in this property.

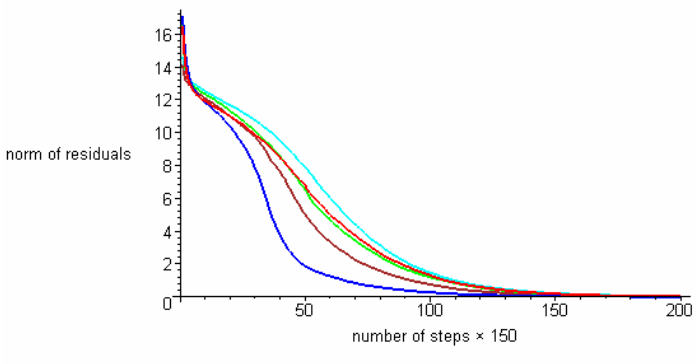

Figure 1
The convergence of the Algorithm 1. on the 16×17 dimension problems.

2.2 Test Results for the Convergence

Several arbitrary 16×17 dimension $A \in R^{16} \to R^{17}$ matrix and $b \in R^{17}$ was generated as test problems. The Algorithm 1 was realized using the Maple8 Linear Algebra library. The results of the algorithm are shown on the Fig. 1.

On the Fig. 1 one can see that the convergence of the algorithm is strictly monoton but very slow. The attained residual norms are:

$$\|r\| = 0.02531, \quad 0.004331, \quad 0.021461, \quad 0.02777, \quad 0.01322 \quad .$$

These results are such ones what we could expect because the test problems are such linear systems where the matrix is not a square matrix, and the test vectors were chosen by hasard. By analysing of the algorithm it is possible to elaborate better strategy for the generation of test vectors, but this can be a topic of a new work. On the Fig. 2 we show the speed of convergence depending on the condition number of the 10×10 test matrices. One can see, that the convergence is very slow for the test problems with large condition numbers, which is not a surprising result too.

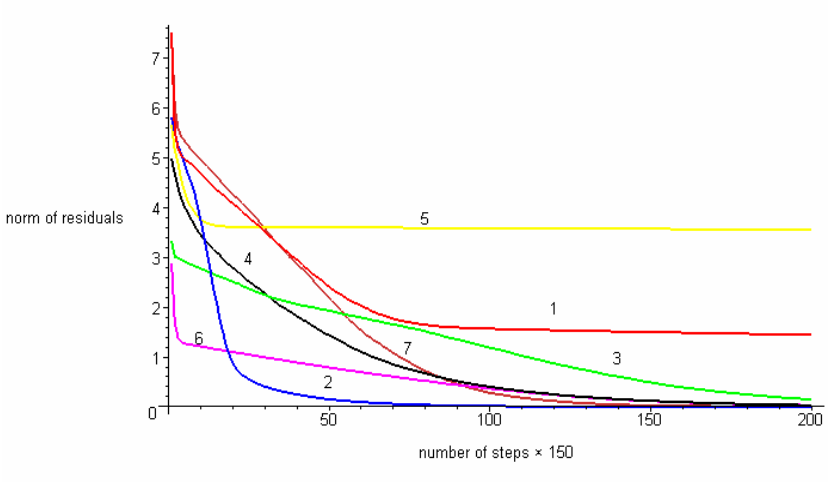

Figure 2
The convergence results of the Algorithm 1. for 10×10 dimension test problems with different condition numbers.

On the Table 1 we show the results for the different test problems. From this table we can conclude that the Algorithm 1 because of its hazard character can produce quite different speed of convergence, but this speed depends on the condition number of the test matrices.

Table 1

Numeration of the curve on Fig. 1.	1	2	3	4	5	6	7
Condition number of the matrix.	623	75	573	123	1544	90	227
Residual norm attained.	1.4	0.000089	0.14	0.031	3.5	0.032	0.00 25

3 Summary

We have formulated a new genetic like algorithm for the solution of general linear systems of equations, which based on the residual minimization technique. The test results confirm the theoretical results. The convergence speed of the sequential algorithm proposed is very slow, but the idea suggested is appropriate for constructing more effective parallel algorithm too combining this idea with the results published in ([8-11]). This problem can be the subject of some forthcoming paper.

References

[1] Louis A. Hageman, Davis M. Joung,: Applied Iterative Methods, Computer Science and Applied Mathematics, Academic Press, (1981)

[2] P. G. Ciarlet, Introduction à l'analyse numérique matricielle aet à l'optimisation, MASSON, Paris, 1982

[3] G. Golub, A. Greenbaum, M. Luskin, eds., Recent Advances is Iterative Methods, The IMA Volumes in Math. and its Applications Vol. 60, Springer Verlag, 1994

[4] J. Gilbert, D. Kershaw, Large-Scale Matrix Problems and the Numerical Solution of Partial Differential Equations, Advances in Numerical Analysis, Vol. III, Clarendon Press, Oxford, 1994

[5] J. Saad and M. H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Stat. Comput., 7 (1986) pp. 856-869

[6] C. Lanczos, Solution of systems of linear equations by minimized iterations, J Res. Nat. Bur. Standards, 49 (1952) pp. 33-53

[7] R. W. Freund and N. M. Nachtigal, QMR: a quasi-minimal residual method for nonhermitian linear systems, Numer. Math., 60 (1991) pp. 315-339

[8] G. Molnárka and B. Török, Residual Elimination Algorithm for Solving Linear Equations and Application to Sparse Systems. Zeitschrift für Angewandte Mathematik und Mechanik (ZAMM), Issue 1, Numerical Analysis, Scientific Computing, Computer Science. pp. 485-486, 1996

[9] J. K Tar**,** I. J. Rudas, J. F. Bitó, L. Madarász: An Emerging Branch of Computational Cybernetics Dedicated to the Solution of Reasonably Limited Problem Classes, AT & P Journal Plus2 2001, pp. 19-25

[10] J. K. Tar, I. J. Rudas, L. Madarász, J. F. Bitó: "Simultaneous Optimization of the External Loop Parameters in an Adaptive Control Based on the Co-operation of Uniform Procedures", Journal of Advanced Computational Intelligence, Vol. 4, No. 4, 2000, pp. 279-285

[11] A. Lotfi, An H-matrix Type Preconditioner for Frictional Contact Problems, PAMM Proc. Appl. Math. Mech. 4, 438-439 (2004)

This work was supported by OTKA N^{o} T043258 and

CBC PHARE 2002/000-317-02-20