Journal of Expert Systems (JES)
Vol. 2, No. 1, 2013, ISSN 2169-3064

www.worldsciencepublisher.org

On the solution of ill-conditioned systems of linear equations

Zahra Vahedipour, Behrouz Daneshian

Department of Mathematics, Lahijan branch, Islamic Azad University, Lahijan, Iran

Email: zahra_vahedipour@yahoo.com

Abstract – In this paper we propose novel technique to solve the ill-conditioned system of linear equations $Ax = b$. For given a matrix A, we find two invertible diagonal matrices D_1 and D_2 by Simulated Annealing method, such that

$$cond(D_1AD_2) = \inf_{\Delta_1,\Delta_2\in\Upsilon} cond(\Delta_1A\Delta_2)$$

(Υ denoting the set of all diagonal matrices) relative to a given matrix norm. With this step carried out, the solution of the system $Ax = b$ is effected by solving the system $(D_1AD_2)y = D_1b$ and then calculating $x = D_2y$.

Keywords – Simulated Annealing; Ill-conditioned matrix; Stochastic optimization.

1. Introduction

A system of equations is considered to be ill-conditioned if a small change in the coefficient matrix or a small change in the right hand side results in a large change in the solution vector. It is well known that for a system of equations with an ill-conditioned matrix, an erroneous solution can be obtained which seems to satisfy the system quite well. Various measures of the ill-conditioning of a matrix have been proposed. For example, the condition number associated with the linear equation $Ax = b$ gives a bound on how inaccurate the solution x will be after approximation. Note that this is before the effects of round-off error are taken into account; conditioning is a property of the matrix, not the algorithm or floating point accuracy of the computer used to solve the corresponding system. In particular, one should think of the condition number as being (very roughly) the rate at which the solution, x, will change with respect to a change in b. Thus, if the condition number is large, even a small error in b may cause a large error in x. On the other hand, if the condition number is small then the error in x will not be much bigger than the error in b.

The condition number is defined more precisely to be the maximum ratio of the relative error in x divided by the relative error in b.

$$\frac{\|A^{-1}e\| / \|A^{-1}b\|}{\|e\| / \|b\|}$$

This is easily transformed to

$$\left(\|A^{-1}e\| / \|e\|\right).\left(\|b\| / \|A^{-1}b\|\right)$$

The maximum value (for nonzero b and e) is easily seen to be the product of the two operator norms:

$$\kappa(A) = \|A\| . \|A^{-1}\|$$

The same definition is used for any consistent norm, i.e. one that satisfies $\kappa(A) \geq 1$.

When the condition number is exactly one, then the algorithm may find an approximation of the solution with an arbitrary precision. However it does not mean that the algorithm will converge rapidly to this solution, just that it won't diverge arbitrarily because of inaccuracy on the source data (backward error), provided that the forward error introduced by the algorithm does not diverge as well because of accumulating intermediate rounding errors.

The condition number may also be infinite, in which case the algorithm will not reliably find a solution to the problem, not even a weak approximation of it (and not even its order of magnitude) with any reasonable and provable accuracy.

We shall now describe how to modify a given ill-conditioned matrix in order to make it a better conditioned one. Then we shall apply a new stochastic procedure to the modified system to obtain a solution. In this paper we propose novel techniques to solve the system of linear equations $Ax = b$. For given a matrix A, we find two invertible diagonal matrices D_1 and D_2 by Simulated Annealing method, such that

$$cond(D_1AD_2) = \inf_{\Delta_1,\Delta_2 \in \Upsilon} cond(\Delta_1 A \Delta_2)$$

(Υ denoting the set of all diagonal matrices) relative to a given matrix norm. With this step carried out, the solution of the system $Ax = b$ is effected by solving the system $(D_1AD_2)y = D_1b$ and then calculating $x = D_2y$ [1, 2, 5, 10].

2. Simulated annealing algorithm

Simulated annealing (SA) is a stochastic approach for minimizing multivariate functions based on the principles of thermodynamics. Simulated annealing is proposed in Kirkpatrick, Gelatt and Vecchi (1983) and Cerny (1985) for finding the global minimum of an objective function that may possess several local minima [7].

SA is motivated by an analogy to annealing in solids. The idea of SA comes from a paper published in 1953 by Metropolis et a1. [9], and motivated by an analogy to the behavior of physical systems in the presence of a heat bath. The algorithm in that paper simulated the cooling of material in a heat bath. This is a process known as annealing.

If we heat a solid past melting point and then cool it, the structural properties of the solid depend on the rate of cooling. If the liquid is cooled slowly enough, large crystals will be formed. However, if the liquid is cooled quickly (quenched) the crystals will contain imperfections. Metropolis's algorithm simulates the material as a system of particles. The algorithm simulates the cooling process by gradually lowering the temperature of the system until it converges to a steady, *frozen* state.

At each iteration of a simulated annealing algorithm applied to a discrete optimization problem, the objective function generates values for two solutions (the current solution and a newly selected solution) are compared. Improving solutions are always accepted; while a fraction of non-improving (inferior) solutions are accepted in the hope of escaping local optima in search of global optima. The probability of accepting non-improving solutions depends on a temperature parameter, which is typically non-increasing with each iteration of the algorithm.

The key algorithmic feature of simulated annealing is that it provides a means to escape local optima by allowing *hill-climbing* moves (i.e., moves which worsen the objective function value). As the temperature parameter is decreased to zero, hill climbing moves occur less frequently, and the solution distribution associated with the inhomogeneous Markov chain that models the behavior of the algorithm converges to a form in which all the probability is concentrated on the set of globally optimal solutions (provided that the algorithm is convergent; otherwise the algorithm will converge to a local optimum, which may or not be globally optimal).

3. Numerical Results

Example 1: Consider the following system of two linear equations in two unknowns.

$$\begin{bmatrix} 400 & -201 \\ -800 & 401 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 200 \\ -200 \end{bmatrix}$$

The solution is $x_1 = -100$ and $x_2 = -200$.

Now, let us make a slight change in one of the elements of the coefficient matrix. Change A_{11} from 400 to 401 and this small change affects the solution. This time the solution is $x_1 = \mathbf{40000}$ and $x_2 = \mathbf{79800}$.

With a modest change in one of the coefficients one would expect only a small change in the solution. However, in this case the change in solution is quite significant. It is obvious that in this case the solution is very

sensitive to the values of the coefficient matrix *A*, and this system is ill-conditioned.

Now, we try to obtain tow diagonal matrices D_1 and D_2 by Simulated Annealing method, such that the condition number of D_1AD_2 have the minimum value. The results are

$$D_1 = \begin{bmatrix} -143.2083 & 0 \\ 0 & 71.6182 \end{bmatrix} \text{ and}$$

$$D_2 = \begin{bmatrix} 181.3483 & 0 \\ 0 & 361.2922 \end{bmatrix}.$$

In this example, we have *cond(A)*= 2503 and $cond\,(D_1AD_2) = 1606$.

Example 2: Hilbert matrix is a square matrix with entries being the unit fractions

$$H_{ij} = \frac{1}{i + j - 1}$$

The Hilbert matrices are canonical examples of ill-conditioned matrices, making them notoriously difficult to use in numerical computation. For example The condition number of the *n*-by-*n* Hilbert matrix grows as $O((1+\sqrt{2})^{4n} / \sqrt{n})$. For $n = 50$, the condition number is equal to 1.3070e+19.

We applied the Simulated Annealing to obtain better condition number and we find $cond\,(D_1AD_2) = 4.4289\text{e+}17$. The diagonal elements of the coefficient matrix D_1 and D_2 are arranged in the following tables.

Table1: diagonal elements of the coefficient matrix D_1

Diagonal elements from 1 to 10	Diagonal elements from 11 to 20	Diagonal elements from 21 to 30	Diagonal elements from 31 to 40	Diagonal elements from 41 to 50
0.0744	0.168	2.2609	3.0791	-1.4584
0.065	0.7633	1.2473	1.355	2.116
0.0901	1.3563	-2.2467	1.9006	-2.251
0.2151	0.9845	-2.0986	1.4262	2.5664
-0.5945	-0.0082	1.8209	-1.8275	2.4349
-0.7718	0.0149	1.5567	1.3727	2.1465
-1.4754	-1.8157	2.3818	1.4433	2.1073
0.8923	-2.1838	1.7051	2.3483	-2.5665
0.7916	2.4525	1.5333	2.7381	1.3035
-0.8221	0.9635	-1.709	1.9328	-2.6583

Table2: diagonal elements of the coefficient matrix D_2

Diagonal elements from 1 to 10	Diagonal elements from 11 to 20	Diagonal elements from 21 to 30	Diagonal elements from 31 to 40	Diagonal elements from 41 to 50
0.9534	-0.5799	1.7436	1.5195	1.1977
-0.3542	-2.1745	0.8322	-2.5789	1.0842
-0.6122	0.6776	-2.2245	3.1977	-1.7965
1.1814	0.5723	3.1721	-1.3412	2.0111
0.0947	-1.2941	2.5523	1.6147	2.3489
-0.3552	1.7056	2.2819	1.0562	1.487
-1.0628	0.2142	1.6846	3.6905	1.1119
0.2022	2.9537	-2.0816	2.2463	2.0515
0.5813	-1.719	1.8723	1.173	3.3483
1.1553	0.8532	-0.7156	2.1264	-1.2599

References

[1] F. L. BAUER, Optimally scaled matrices, Numerische Mathematik, 5(1963) 7387.

[2] P.A. Businger, Matrices which can be optimally scaled, Numerische Mathematik 12 (1968) 346–348.

[3] Chin-Chieh Chiang, John P. Chandler, An Approximate Equation for the Condition Numbers of Well-scaled Matrices, Proceedings of The 2008 IAJC-IJME International Conference.

[4] B. Fathi Vajargah, M Moradi, M Kanafchian, Monte Carlo optimization for reducing the condition number of ill conditioned matrices, Advances in Computational Mathematics and its Applications 1 (3), 169-173.

[5] B. Fathi Vajargah, M Moradi, Diagonal Scaling of Ill-Conditioned Matrixes by Genetic Algorithm, Journal of Applied Mathematics, Statistics and Informatics 8 (1), 49-53.

[6] T.I. Fenner, G. Loizou, Some new bounds on the condition numbers of optimally scaled matrices, Journal of the ACM 21 (3) (1974) 514–524.

[7] S. Kirkpatrick, C.C. Gelatt, M.P. Vecchi, Optimization by simulated annealing, Science 220 (1983) 671–679.

[8] A.W. Marshall, I. Olkin, Scaling of matrices to achieve specified row and column sums, Numerische Mathematik 12 (1968) 83–90.

[9] Metropolis, A.Rosenbluth, M. Rosenbluth, A. Teller and E. Teller, *Equation of State Calculations by Fast Computing Machines*, J. of Chem. Physics, 21 (1953) 1087 -1092.

[10] M. Moradi**,** B. Fathi Vajargah, B. Daneshian, Z. Vahedipour, Poisson mixed distributions in the branching processes, Advances in Computer Science and its Applications 1 (4), 267-269.

[11] D.P. O'Leary, Scaling symmetric positive definite matrices to prescribed row sums, Linear Algebra & Its Applications 370 (2003) 185–191.

[12] R. Y. Rubinstein, Simulation and the Monte Carlo method, John Wiley and Sons, New York, 1981.

[13] A.van der Sluis, Condition numbers and equilibration of matrices, Numerische Mathematik 14 (1969) 14–23.

[14] Yudong Zhang, Lenan Wu, A Robust Hybrid Restarted Simulated Annealing Particle Swarm Optimization Technique, A C S A ,Vol 1, No 1 (2012).