

A Geometric Study of Convergence: The Nonlinear Conjugate Gradient Method

Ana Clara de Souza Oliveira*, Sandra Augusta Santos.

Abstract

The Conjugate Gradient Method (CG) is a numerical method for solving quadratic problems, which can be used to minimize any differentiable function. In this study, we have analyzed the efficiency of CG for quadratic problems and implemented a version of this method for nonlinear systems in the least-squares formulation. We have also prepared a geometric analysis of the convergence of its nonlinear version, in order to evaluate the dependence of its efficiency with the starting point.

Key words:

conjugate gradient method, nonlinear problems, geometric analysis of convergence

Introduction

The numerical Conjugate Gradient Method has as main inspiration the solution of linear systems of the form $Ax=b$ (1), where A is a symmetric positive definite matrix, equivalent to minimize quadratic forms $f(x) = 1/2x^T Ax - b^T x + c$ which gradient equals to zero results in (1). The method can also be used to minimize any smooth function $f(x)$ whose gradient is available.

Based on proposals of Shewchuk in [3], we have implemented algorithms to evaluate the efficiency of the method to the linear and nonlinear problems. For the implementations, we used the software *Matlab*. In the linear cases, we used the function PCG, of this software, to validate our results. For the nonlinear cases, we implemented the suggested algorithm and introduced modifications in order to get better results. We have analyzed the efficiency of the final version for problems proposed by Moré, Garbow and Hillstrom in [1] and conducted a geometric study of the convergence of this method, inspired by the study of Santos in [2].

Results and Discussion

For the linear problems, the CG presented high efficiency. We observed two conditions that determine the convergence of the method: the dimension of the problem and the spectral condition number of the matrix A . We concluded that the efficiency is independent of the starting point. Besides, the convergence is guaranteed in n iterations, n being the dimension of the problem. As expected, the results also showed that the method is favored when the eigenvalues of the matrix A are clustered.

For the nonlinear problems, we observed that conjugation loses efficiency because of the dissimilarity of the function with its quadratic expansion. For better results, we applied line searches with a condition of sufficient decrease (Armijo condition).

For these cases, we also noted that the convergence of the method is dependent of the starting point, different from the linear case. Given this dependence, we studied the convergence of the method for some functions, analyzing the number of iterations for the CG and for the line search. We illustrate here with the level curves of the Beale function (Figure 1(a)).

For the line search, we obtained the Figure 1(b), where the color scale (increasing from magenta to red) is proportional to the number of performed iterations. To

illustrate the convergence of the method, we prepared the Figures 1(c) and 1(d), where black points indicate local where the method did not progress; beige, where the number of maximum iterations was not sufficient to converge; and, again, increasing from magenta to red, points where the method converged with a number of iterations proportional to the scale. We have observed a direct relationship between the convergence of the method and the geometry of the function.

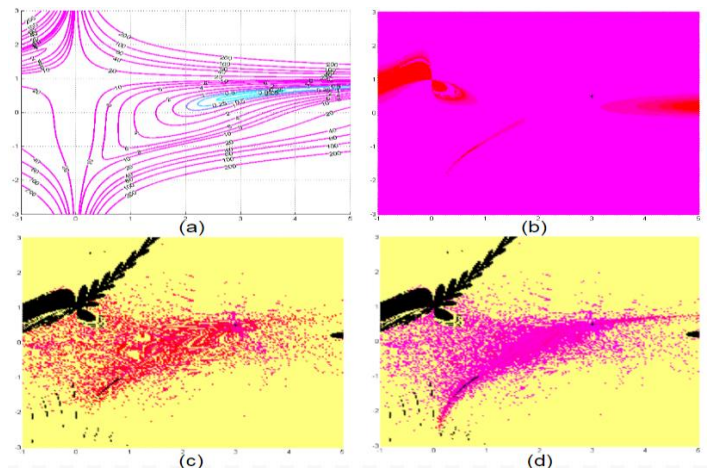


Image 1. Study of the convergence of the Conjugate Gradient Method for (a) the Beale function. (b) Analysis of the line search and Analysis of the CG convergence for maximum number of iterations of (c) 20 and (d) 100.

Conclusions

We concluded that the Conjugate Gradient Method is efficient for linear problems, but does not present the same behavior for the nonlinear case; its convergence depends on the starting point, being directly linked to the geometry of the function.

Acknowledgement

I thank my advisor Sandra, for her unconditional support and inspiration. I thank PICME program for the opportunity.

- [1] Garbow, B. S. Hillstrom, K. E. Moré, J. J. *Testing Unconstrained Optimization Software*. Março 1981. ACM Transactions on Mathematical Software, Vol 7, No 1, 136-140.
- [2] Santos, L. T. *Sistemas não Lineares e Fractais*. Dezembro 1993. Matemática Universitária, No 15, 102-116.
- [3] Shewchuk, J. R. *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, edição 1 1/4, Agosto 1994.