NUMERICAL MATHEMATICS
A Journal of Chinese Universities (English Series)

Vol. **15** No. 4, pp. 299-305
November 2006

A Rapidly Convergence Algorithm for Linear Search and its Application

Jianliang Li[1,*], Hua Zhu[1], Xianzhong Zhou[2] and Wenjing Song[1]

[1] *Department of Applied Mathematics, Nanjing University of Science and Technology, Nanjing 210094, China.*

[2] *School of Management Science Engineering, Nanjing University, Nanjing 210093, China.*

Received May 9, 2004; Accepted (in revised version) March 4, 2005

Abstract. The essence of the linear search is one-dimension nonlinear minimization problem, which is an important part of the multi-nonlinear optimization, it will be spend the most of operation count for solving optimization problem. To improve the efficiency, we set about from quadratic interpolation, combine the advantage of the quadratic convergence rate of Newton's method and adopt the idea of Anderson-Björck extrapolation, then we present a rapidly convergence algorithm and give its corresponding convergence conclusions. Finally we did the numerical experiments with the some well-known test functions for optimization and the application test of the ANN learning examples. The experiment results showed the validity of the algorithm.

Key words: Linear Search; nonlinear optimization; accelerating convergence; learning algorithm.

AMS subject classifications: 40A25, 65B05, 65K10

1 Introduction

As an important part of many variable nonlinear optimization, most of operation count for solving optimization problem will be spend by linear search. So the accelerating convergence of the linear search is important to improve the efficiency of optimization algorithm. For a linear search problem

$$\min_{\eta>0} \boldsymbol{f}(\boldsymbol{x}_k + \eta \boldsymbol{d}_k) \tag{1}$$

where $f : D \subset R^n \to R$, its essence is the minimization of the one-dimensional nonlinear function $J(\eta) = \boldsymbol{f}(\boldsymbol{x}_k + \eta \boldsymbol{d}_k)$, which is also the extraction of root of the single nonlinear equation

$$J'(\eta) = \nabla \boldsymbol{f}(\boldsymbol{x}_k + \eta \boldsymbol{d}_k)^T \boldsymbol{d}_k = 0 \tag{2}$$

There are many methods for finding roots of a single nonlinear equation [1]. But if Newton's method

$$\eta_{k+1} = \eta_k - \frac{J'(\eta_k)}{J''(\eta_k)} \tag{3}$$

*Correspondence to: Jianliang Li, Department of Applied Mathematics, Nanjing University of Science and Technology, Nanjing 210094, China. Email: `li16006@hotmail.com`

which has the quadratic convergence rate is used, not only the derivative of $J'(\eta_k)$ at the point η_k, that is the gradient $\nabla \boldsymbol{f}(\boldsymbol{x}_k+\eta \boldsymbol{d}_k)$, but also the Hessian matrix $\nabla^2 \boldsymbol{f}(\boldsymbol{x}_k+\eta \boldsymbol{d}_k)$ of $\boldsymbol{f}(\boldsymbol{x})$ have to be calculated. The operation count for calculating the gradient and Hessian matrix is very large. So we set about from quadratic interpolation, consider the advantage of the quadratic convergence rate Newton's method, and adopt the idea of extrapolation, we will present a rapidly convergence algorithm which not only avoid the calculation of the second order derivative of $\boldsymbol{f}(\boldsymbol{x})$ but also keep the same number of derivative counts of $\boldsymbol{f}(\boldsymbol{x})$. So the convergence rate is improved.

2 Algorithm and convergence

Assume η_{n-1}, η_k are two approximations of the root of the nonlinear equation (2) η^*, a new approximation of η^* can be achieved with the quadratic interpolation [2]

$$\eta_{k+1} = \eta_k - \frac{\eta_k - \eta_{k-1}}{J'(\eta_k) - J'(\eta_{k-1})} J'(\eta_k). \tag{4}$$

In fact, formula (4) is exactly the secant method to equation (2).

With the approximations $\eta_{2k}, \eta_{2k-1}, \eta_{2k-2}$ of exact solution η^* and their derivatives, we can construct a quadratic interpolation polynomial of $J'(\eta)$

$$\begin{aligned} p(\eta) \quad = \quad & \frac{(\eta-\eta_{2k-1})(\eta-\eta_{2k})}{(\eta_{2k-2}-\eta_{2k-1})(\eta_{2k-2}-\eta_{2k})} J'(\eta_{2k-2}) + \frac{(\eta-\eta_{2k})(\eta-\eta_{2k-2})}{(\eta_{2k-1}-\eta_{2k-2})(\eta_{2k-1}-\eta_{2k})} J'(\eta_{2k-1}) \\ & + \frac{(\eta-\eta_{2k-1})(\eta-\eta_{2k-2})}{(\eta_{2k}-\eta_{2k-2})(\eta_{2k}-\eta_{2k-1})} J'(\eta_{2k}). \end{aligned} \tag{5}$$

Then replace $J''(\eta_{2k})$ of formula (3) by $\phi(\eta_{2k})$ of the derivative of $p(\eta)$ at η_{2k}

$$\begin{aligned} \phi(\eta_{2k}) \quad = \quad & \frac{\eta_{2k}-\eta_{2k-1}}{(\eta_{2k-2}-\eta_{2k-1})(\eta_{2k-2}-\eta_{2k})} J'(\eta_{2k-2}) + \frac{\eta_{2k}-\eta_{2k-2}}{(\eta_{2k-1}-\eta_{2k-2})(\eta_{2k-1}-\eta_{2k})} J'(\eta_{2k-1}) \\ & + \frac{2\eta_{2k}-\eta_{2k-1}-\eta_{2k-2}}{(\eta_{2k}-\eta_{2k-2})(\eta_{2k}-\eta_{2k-2})} J'(\eta_{2k}) \end{aligned} \tag{6}$$

and combine Anderson-Björck extrapolation with quadratic interpolation [3]. The interpolation formulas are given as follows

$$\begin{cases} \eta_{2k} = \eta_{2k-1} - \dfrac{\eta_{2k-1}-\eta_{2k-2}}{J'(\eta_{2k-1}) - J'(\eta_{2k-2})} J'(\eta_{2k-1}), & \text{(7a)} \\ \eta_{2k+1} = \eta_{2k} - \dfrac{J'(\eta_{2k})}{\phi(\eta_{2k})}. & \text{(7b)} \end{cases}$$

The iterative method with linear search based on the Anderson-Björck extrapolation are called ABE algorithm for short. The steps are given as follows.

Algorithm (ABE algorithm)

Step 1 Given a tolerance $\varepsilon > 0$ and an iteration upper limit N.

Step 2 With forward-backward algorithm, find the interval $[a, b]$ containing the minimum point of $J(\eta)$ on $[0, +\infty]$, and let $k = 1$, $\eta_{2k-2} = a$, $\eta_{2k-1} = b$.

Step 3 Calculate η_{2k} by formula (7a). If $|J'(\eta_{2k})| \le \varepsilon$, then stop and output η_{2k}.