Journal of Applied Mathematics & Bioinformatics, vol.4, no.3, 2014, 65-73
ISSN: 1792-6602 (print), 1792-6939 (online)
Scienpress Ltd, 2014

A Second Order Method for Minimizing Unconstrained Optimization Problems

M.O. Oke[1]

Abstract

One of the commonly used second order methods for the minimization of quadratic functionals is the Newton's method. However, for nonquadratic functionals, the Newton's iterative scheme may not converge to the optimum minimum point. This paper is based on a second order method derived from the Newton's iterative scheme by incorporating a minimizing step length in the Newton's formula. The second order iterative scheme used in this paper minimizes both quadratic and nonquadratic functionals in one iteration. This makes it more efficient than other methods of minimization.

Mathematics Subject Classification: 49M15

Keywords: Unconstrained Minimization Problems; Newton's Method; Second Order Method; Quadratic Functionals; Nonquadratic Functionals

[1] Department of Mathematical Sciences, Ekiti State University, P.M.B 5363, Ado – Ekiti, Nigeria

Article Info: *Received* : August 10, 2014. *Revised* : September 29, 2014.
Published online : October 15, 2014.

1 Introduction

A second order method is a method that uses the second derivatives of the objective function for the minimization of functionals. The Newton's method is a well-known second order method for minimizing quadratic functionals. But if f(x) is a non-quadratic function, the Newton's method may converge to saddle points, relative maxima or it may diverge.

Many researchers have worked on minimization of functionals, David and Oregon (2010) considered the minimization of nonlinear problems in optimization using modified conjugate gradient method. Fletcher and Freeman (1977) worked on a modified Newton method for minimization. Anderson (2010) looked at minimization of constrained problems in optimization while Ramesh and Craven (2009) worked on computational methods for minimizing some engineering problems to mention a few. None of these researchers worked on methods that converge for both quadratic and nonquadratic functionals in one iteration. The second order method used in this paper converges to minimum points in one iteration for both quadratic and non-quadratic functionals.

2 Materials and Methods

The Newton's iterative scheme for minimizing unconstrained optimization problem is giving by the formula

$$X_{i+1} = X_i + D_i = X_i - A_i^{-1} g_i \quad (1)$$

where

g_i is the gradient of the given function at point X_i.

A_i is a non-singular Hessian matrix of the function evaluated at point X_i, Rao (1977) and Polak (1971).

The sequence of points generated from the iterative scheme in (1) converges to the optimum minimum point X^* from any initial starting point X_1 that is close to the solution for all quadratic functionals, Rao (1977) and Polak (1971). However, for nonquadratic functionals, the Newton's iterative scheme in (1) may not converge to the minimum point X^*, Rao (1977). Therefore, there is the need to modify the Newton's iterative scheme to find the minimum of both quadratic and nonquadratic functionals.

This paper is based on a second order method derived from the Newton's iterative scheme in (1) by incorporating a minimizing step length in the direction of D_i. The modified Newton's method will now be in the form

$$X_{i+1} = X_i + \lambda_i^* D_i = X_i - \lambda_i^* A_i^{-1} g_i \tag{2}$$

where λ_i^* is the minimizing step length in the direction $D_i = -A_i^{-1} g_i$. With this modification in (2), both quadratic and nonquadratic functionals will converge to the optimum minimum point in lesser number of iterations compared to the conjugate gradient and the Newton's method.

3 Results and Discussion

Example 1: Minimize $f(x_1, x_2) = -\dfrac{1}{(x_1^2 + 2x_2^2 - 1)}$ from the starting point

$$X_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

The gradient g of the function $f(x_1, x_2)$ is given by

$$g = \begin{pmatrix} \dfrac{\delta f}{\delta x_1} \\ \dfrac{\delta f}{\delta x_2} \end{pmatrix} = \begin{pmatrix} \dfrac{2x_1}{(x + x - 1)^2} \\ \dfrac{4x_2}{(x_1^2 + 2x_2^2 - 1)^2} \end{pmatrix}$$

The matrix of second order partial derivative of the function $f(x_1, x_2)$ is given by

$$A = \begin{pmatrix} \frac{-6x_1^2 + 4x_2^2 - 2}{(x_1^2 + 2x_2^2 - 1)^3} & \frac{-16x_1x_2}{(x_1^2 + 2x_2^2 - 1)^3} \\ \frac{-16x_1x_2}{(x_1^2 + 2x_2^2 - 1)^3} & \frac{4x_1^2 - 24x_2^2 - 4}{(x_1^2 + 2x_2^2 - 1)^3} \end{pmatrix}$$

At point X_1,

$$g_1 = \begin{pmatrix} 0.444444444 \\ 0 \end{pmatrix}$$

And

$$A_1 = \begin{pmatrix} -0.962962963 & 0 \\ 0 & 0.444444444 \end{pmatrix}$$

$$A_1^{-1} = \begin{pmatrix} -1.038461538 & 0 \\ 0 & 2.250000001 \end{pmatrix}$$

$$A_1^{-1}g_1 = \begin{pmatrix} -0.46153846 \\ 0 \end{pmatrix}$$

The direction of search $D_i = -A_i^{-1}g_i$

Therefore $D_i = -A_i^{-1}g_i = \begin{pmatrix} 0.46153846 \\ 0 \end{pmatrix}$

From the modified iterative scheme, we have

$$X_2 = X_1 + \lambda_1 D_1 = X_1 - \lambda_1 A_1^{-1} g_1 = \begin{pmatrix} 2 + 0.46153846\lambda_1 \\ 0 \end{pmatrix}$$

$$f(X_2) = f(X_1 + \lambda_1 D_1) = -\frac{1}{\left((2 + 0.46153846\lambda_1)^2 - 1\right)}$$

$$\frac{\delta f}{\delta \lambda_1} = \frac{2(2 + 0.46153846\lambda_1)(0.46153846)}{\left((2 + 0.46153846\lambda_1)^2 - 1\right)^2}$$

$\frac{\delta f}{\delta \lambda_1} = 0$ implies that

$$1.84615384 + 0.4260355\lambda_1 = 0$$

and

$\lambda_1 = -4.333333349$

Substituting the value of λ_1 in X_2, we have

$$X_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix} - 4.333333349 \begin{pmatrix} 0.46153846 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The value of the gradient at point X_2 is $g_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. This clearly shows that X_2 is the minimum point. $f_{Optimum} = 1$

Example 2: Minimize $f(x_1, x_2) = x_1^3 - 2x_1^2 x_2 + x_1^2 + x_2^2 - 2x_1 + 1$

from the starting point $X_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

The gradient g of the function $f(x_1, x_2)$ is given by

$$g = \begin{pmatrix} 3x_1^2 - 4x_1 x_2 + 2x_1 - 2 \\ -2x_1^2 + 2x_2 \end{pmatrix}$$

The matrix of second order partial derivative of the function $f(x_1, x_2)$ is given by

$$A = \begin{pmatrix} 6x_1 - 4x_2 + 2 & -4x_1 \\ -4x_1 & 2 \end{pmatrix}$$

At point X_1, $g_1 = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$ and

$$A_1 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \ A_1^{-1} = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}, \ A_1^{-1} g_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

The direction of search $D_i = -A_i^{-1} g_i = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

From the modified iterative scheme, we have

$$X_2 = X_1 + \lambda_1 D_1 = X_1 - \lambda_1 A_1^{-1} g_1 = \begin{pmatrix} \lambda_1 \\ 0 \end{pmatrix}$$

$f(X_2) = f(X_1 + \lambda_1 D_1) = \lambda_1^3 + \lambda_1^2 - 2\lambda_1 + 1$

$\frac{\delta f}{\delta \lambda_1} = 3\lambda_1^2 + 2\lambda_1 - 2,$

$\frac{\delta f}{\delta \lambda_1} = 0$ implies that

$\lambda_1 = 0.548537$ or -1.2152504

Substituting the value of λ_1 in X_2, we have

$X_2 = \begin{pmatrix} 0.5485837 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} -1.2152504 \\ 0 \end{pmatrix}$

The gradient $g_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ at these two values of X this indicates that the two values of X represents minimum points for the given function. $f_{Optimum}$ for these two values are 0.3688695 and 3.1126118 respectively.

Example 3: Minimize $f(x_1, x_2) = 4x_1^2 + 3x_2^2 - 5x_1x_2 - 8x_1$

from the starting point $X_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

The gradient g of the function $f(x_1, x_2)$ is given by

$g = \begin{pmatrix} 8x_1 - 5x_2 - 8 \\ 6x_2 - 5x_1 \end{pmatrix}$. At point X_1, $g_1 = \begin{pmatrix} -8 \\ 0 \end{pmatrix}$

The matrix of second order partial derivative of the function $f(x_1, x_2)$ is given by

$A = A_1 = \begin{pmatrix} 8 & -5 \\ -5 & 6 \end{pmatrix}$, $A_1^{-1} = \begin{pmatrix} 0.260869565 & 0.217391304 \\ 0.217391304 & 0.347826087 \end{pmatrix}$,

$A_1^{-1}g_1 = \begin{pmatrix} -2.086956522 \\ -1.739130435 \end{pmatrix}$

The direction of search $D_i = -A_i^{-1}g_i = \begin{pmatrix} 2.086956522 \\ 1.739130435 \end{pmatrix}$

From the modified iterative scheme, we have

$$X_2 = X_1 + \lambda_1 D_1 = X_1 - \lambda_1 A_1^{-1} g_1 = \begin{pmatrix} 2.086956522\lambda_1 \\ 1.739130435\lambda_1 \end{pmatrix}$$

$$f(X_2) = f(X_1 + \lambda_1 D_1) = 8.34782609\lambda_1^2 - 16.69565218\lambda_1$$

$$\frac{\delta f}{\delta \lambda_1} = 8.34782609\lambda_1 - 16.69565218,$$

$$\frac{\delta f}{\delta \lambda_1} = 0 \text{ implies that } \lambda_1 = 1$$

Substituting the value of λ_1 in X_2, we have

$$X_2 = \begin{pmatrix} 2.086956522 \\ 1.739130435 \end{pmatrix}$$

The value of the gradient at point X_2 is $g_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. This clearly shows that X_2 is the minimum point. $f_{Optimum} = -8.3478609$

Example 4: Minimize $f(x_1, x_2) = 2x_1^2 + x_2^2$ from the starting point $X_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

The gradient g of the function $f(x_1, x_2)$ is given by

$g = \begin{pmatrix} 4x_1 \\ 2x_2 \end{pmatrix}$. At point X_1, $g_1 = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$

The matrix of second order partial derivative of the function $f(x_1, x_2)$ is given by

$$A = A_1 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}, \quad A_1^{-1} = \begin{pmatrix} 0.25 & 0 \\ 0 & 0.5 \end{pmatrix}$$

$$A_1^{-1} g_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

The direction of search $D_i = -A_i^{-1} g_i = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$

From the modified iterative scheme, we have

$$X_2 = X_1 + \lambda_1 D_1 = X_1 - \lambda_1 A_1^{-1} g_1 = \begin{pmatrix} 1-\lambda_1 \\ 2-2\lambda_1 \end{pmatrix}$$

$$f(X_2) = f(X_1 + \lambda_1 D_1) = 6\lambda_1^2 - 12\lambda_1 + 6$$

$$\frac{\delta f}{\delta \lambda_1} = 12\lambda_1 - 12,$$

$\frac{\delta f}{\delta \lambda_1} = 0$ implies that $\lambda_1 = 1$

Substituting the value of λ_1 in X_2, we have

$$X_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The value of the gradient at point X_2 is $g_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. This clearly shows that X_2 is the minimum point. $f_{Optimum} = 0$.

The first and second examples are on nonquadratic functionals while the third and fourth examples are on quadratic functionals. From the results, we can easily see that that the minimum points were obtained for all the functionals in one iteration.

4 Conclusion

The second order method in this paper has been used to minimize both quadratic and nonquadratic functionals. The iterative scheme converges to the optimum minimum point in one iteration for all the functionals considered. This now makes this second order method to be more efficient than the conjugate gradient method, the Newton's method and the Quasi-Newton's method that are well known for minimizing unconstrained optimization problems.

References

[1] S.O. David and T.M. Oregon, Minimization of Nonlinear Problems in Optimization Using Modified Conjugate Gradient Method, DUVEN *Journal of Applied Sciences*, **5**(2), (2010), 35 – 47.

[2] R. Fletcher and T.L. Freeman, A Modified Newton Method for Minimization, Journal of Optimization Theory and Applications, **23**(3), (1977), 357 – 372.

[3] R.O. Anderson, Minimization of Constrained Problems in Optimization, *Mavon Journal of Mathematics and Statistics*, **8**(2), (2010), 11 – 26.

[4] S.S. Rao, *Optimization: Theory and Applications*, Wiley Eastern Limited, New Delhi, 1977.

[5] K.A. Ramesh and L.R. Craven, Computational Methods for Minimizing some Engineering Problems, *International Journal of Engineering*, **5**(3), (2009), 56 – 67.

[6] E. Polak, *Computational Methods in Optimization: A Unified Approach*, Academic Press Inc., New York, 1971.