ISSN 2250-3153

Comparative study of Optimization methods for Unconstrained Multivariable Nonlinear Programming Problems

Neha Varma*, Dr. Ganesh Kumar**

*(NET- JRF, Department of Mathematics, B.R.A. Bihar University, Muzaffarpur, Bihar, India)
**(University Professor& Ex. H.O.D. of Mathematics, B.R.A. Bihar University, Muzaffarpur, Bihar, India)

Abstract- In this paper we propose to discuss unconstrained multivariable search methods that are used for optimization of nonlinear programming problems. Earlier Dr.William P.Fox and Dr.William H.Richardson [1] have attempted to solve such problems by using MAPPLE. But we have preferred to solve these problems by OR methods as our present area of research is OR. Although several OR methods are known but we have confined our discussions to Gradient search method, Newton's method and Quasi- Newton methods. We propose to conclude the discussion by taking a suitable example.
.

Key Words- Multivariable,Optimization, Quasi -Newton methods, steepest ascent/descent

I. Introduction

Problems containing more than one variable are called multivariate. The univariate results have multivariate analogues. In the multivariate case, again the necessary and sufficient condition for optimality is given by the system of equations obtained by setting the respective partial derivatives equal to zero. The first and second order partial derivatives are again key to the optimization process, except that a vector of the first derivatives and a matrix of second derivatives are involved.

II. Gradient Search Method

First of all, we take up the gradient search method also known as steepest ascent/descent method which is the simplest and most fundamental method for unconstrained optimization. When the negative gradient is used as its descent direction , the methd is known as steepest descent method and when the positive gradient is used as its ascent direction, the method is called steepest ascent method. The gradient search method can be summarized in the following steps:

1) Choose an initial starting point x_0. Thereafter at the point x_k :

2) Calculate (analytically or numerically) the partial derivatives

   $\partial f(x)/\partial x_j$, j= 1,2,…………n

3) Calculate the maximum/minimum of the new function $f(x_k \pm t_k \nabla f(x_k))$ by using one dimensional search procedure (or calculus method) to find $t=t_k$ that maximizes/minimizes $f(x_k \pm t_k \nabla f(x_k))$ over $t_k \geq 0$ (for maximization) or $t_k \leq 0$ (for minimization).

4) Reset $x_{k+1} = x_k + t_k \nabla f(x_k)$ (for maximization) , $x_{k+1} = x_k - t_k \nabla f(x_k)$ (for minimization). Then go to the stopping rule.

   Stopping rule:

   At the maximum/ minimum, the value of the elements of the gradient vector will be each equal to zero. So we evaluate $\nabla f(x_k)$ at $x=x_k$. If $|\partial f/\partial x_j| \leq \epsilon$ for all j= 1,2,3……………n
   we must stop with the current x_k as the desired approximation of an optimal solution. If not, repeat above steps and set k= k+1.

III. Newton's method

Newton's method for multivariable optimization is analogues to Newton's single variable algorithm for obtaining the roots and Newton –Raphson method for finding the roots of first derivative, given a x_0 , iterates

$x_{k+1}= x_k- f'(x_k)/f''(x_k)$ until $|x_{k+1}- x_k|<\epsilon$

The algorithm is expanded to include partial derivatives w.r.t. each variable's dimension. The algorithm is developed by locating the stationary point of equation

$$f(x)= f(x_k)+ \nabla f(x_k)^T(x-x_k)+ ½(x-x_k)^T H(x-x_k)$$

where H is the Hessian matrix and $(x-x_k)^T$ is the row vector, which is the transpose of the column vector of the difference between the vector of independent variables x and the point x_k used for Taylor's series expansionby setting the first partial derivatives with respect to x_1, x_2,……….x_n equal to zero i.e.

$$\nabla f(x_k)+ H(x- x_k)= 0$$

Then solving for x , the optimum of the quadratic approximation, Newton's method algorithm is obtained as

$$X_{k+1}= x_k- H^{-1}\nabla f(x_k)$$

The Newton's method consists of the following steps:

1) Choose an initial point x_0 .

2) Calculate (analytically or numerically) the partial derivatives $\partial f(x)/\partial x_j$, j= 1,2,3……………,n

3) Calculate the Hessian matrix H_k , the matrix of second partial derivatives at the point x_k.

4)Find the inverse of H_K i.e. H_K^{-1}.

5) Set $x_{k+1}= x_k- H^{-1}\nabla f(x_k)$ 6) If $\|\nabla f(xk)\| < \epsilon$ then stop otherwise repeat the above steps.

IV. Quasi -Newton methods

We have noticed that Newton's method $x_{k+1} = x_k + H_K^{-1} \nabla f(x_k)$ (for maximization) is successful because it uses the Hessian which offers the useful curvature information. However for various practical problems, the computing efforts of the Hessian matrices are very expensive or the evaluation of Hessian is very difficult. Sometimes the Hessian is not available analytically. To overcome these disadvantages Quasi-Newton methods were developed. These methods use the functional values and the gradients of the objective function and also at the same time maintains a fast rate of convergence. There are several updates of Quasi-Newton method. Here we confine our discussion only to DFP update and BFGS update which are both rank two updates.

DFP update

DFP update is a rank- two update and has become the best known of the Quasi-Newton algorithms. The DFP update formula was originally proposed by Davidon [6] and developed later by Fletcher and Powell [7]. Hence it is called DFP update. This formula preserves the positive definiteness in case of minimization and negative definiteness in case of maximization and also symmetry of matrices H_k . DFP method is superlinearly convergent. But for a strictly convex function, under exact line search, DFP method is globally convergent. The DFP algorithm has the following form of equation for maximizing f(x):

$$x_{k+1} = x_k + t_k H_K \nabla f(x_k)$$

where $H_k = H_{k-1}+ A_k+ B_k$ and

the matrices A_k and B_k are given by

$$A_k= \{(x_k- x_{k-1}) (x_k- x_{k-1})^T \}/\{(x_k- x_{k-1})^T (\nabla f(x_k)- \nabla f(x_{k-1})\}$$

$$B_k= \{-H_{k-1}(\nabla f(x_k)- \nabla f(x_{k-1}))(\nabla f(x_k)- \nabla f(x_{k-1}))^T H_{k-1}^T\}/\{ (\nabla f(x_k)- \nabla f(x_{k-1}))^T H_{k-1}(\nabla f(x_k)- \nabla f(x_{k-1}))\}$$

The algorithm begins with a search along the gradient line from the starting point x_0 as given by the following equation

$$x_1= x_0+ t_0 H_0 \nabla f(x_0)$$

where H_0 =I is the unit matrix.

BFGS update

The other famous Quasi-Newton update – BFGS update overcomes all the drawbacks ofoter methods discussed previously and performs better than DFP update. The BFGS update formula developed simultaneously by Broyden[8] , Fletcher[9], Goldfarb[10], Shanno[11] in the year 1970 and hence known as BFGS formula. The BFGS algorithm for maximizing f(x) is given by

$$x_{k+1}=x_k + t_k H_k \nabla f(x_k)$$

where $H_k = H_{k-1}- A_k+ B_k$

The matrices A_k and B_k are given by

$$A_k = \{H_{k-1}\Upsilon_k\delta_k^T+ \delta_k\Upsilon_k^T H_{k-1}\}/ \{\delta_k^T\Upsilon_k\}$$

$$B_k= \{1+ \Upsilon_k^T H_{k-1}\Upsilon_k/ \delta_k^T\Upsilon_k\}\{ \delta_k\delta_k^T/ \delta_k^T\Upsilon_k\}$$

where $\delta_k = x_{k+1} – x_k$ and $\Upsilon_k = \nabla f(x_{k+1}) - \nabla f(x_k)$

Here also x_1 is calculated as in DFP update formula.

V. Illustration:

Maximize $f(x) = 2x_1 x_2 + 2x_2 - x_1^2 - 2x_2^2$ using **Gradient search** starting at point $x_0 = (0, 0)$.

Here $f(x) = 2x_1 x_2 + 2x_2 - x_1^2 - 2x_2^2$

$\nabla f(x_1, x_2)^T = (-2x_1 + 2x_2, 2x_1 - 4x_2 + 2)$

$\nabla f(x_0)^T = f(0,0)^T = (0,2)$

$x_1 = x_0 + t_0 \nabla f(x_0)$

$= \begin{bmatrix} 0 \\ 0 \end{bmatrix} + t_0 \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

$x_1^T = (0, 2t_0)$

Now $f(t) = 4t_0 - 8t_0^2$

For optimum value, $f'(t) = 0$

$t_0 = 1/4$

Also $f''(t) = -16 < 0$

$f(t)$ is maximum at $t_0 = 1/4$

$x_1^T = (0, 1/2)$

Proceeding as above we get a sequence of iterates as $x_2=(1/2,1/2)$, $x_3 = (1/2,3/4)$, $x_4 =(3/4,3/4)$, $x_5 =(3/4,7/8)$, $x_6=(7/8,7/8)$, $x_7 =(7/8,15/16)$ and so on.

We observe that these sequence of trials are converging to $x^* = (1,1)$ which is the optimal solution as verified by the analytical method.

By **Newton's method**

We have $f(x) = 2x_1 x_2 + 2x_2 - x_1^2 - 2x_2^2$

$\nabla f(x_1, x_2)^T = (-2x_1 + 2x_2, 2x_1 - 4x_2 + 2)$

$\nabla f(0,0)^T = (0,2)$

Now $H_0 = \begin{bmatrix} -2 & 2 \\ 2 & -4 \end{bmatrix}$

H_0 is negative definite.

Also $|H_0| = (8-4) = 4$

$H_0^{-1} = \begin{bmatrix} -1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$

$x_1^T = x_0 - H_0^{-1} \nabla f(x_0) = (1,1)$ is the required point of maxima. Clearly the optimum of this quadratic function is obtained in one step as stated in the method.

By **DFG update formula**

$f(x) = 2x_1 x_2 + 2x_2 - x_1^2 - 2x_2^2$

$\nabla f(x_1, x_2)^T = (-2x_1 + 2x_2, 2x_1 - 4x_2 + 2)$

$\nabla f(x_0) = \nabla f(0,0)^T = (0,2)$

Now, $x_1 = x_0 + t_0 H_0(I) \nabla f(x_0)$

$= \begin{bmatrix} 0 \\ 0 \end{bmatrix} + t_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

$x_1^T = (0, 2t_0)$

Now $f(t) = 4t_0 - 8t_0^2$

For optimum value, $f'(t) = 0$

$t_0 = 1/4$

Also $f''(t) = -16 < 0$

$f(t)$ is maximum at $t_0 = 1/4$

$x_1^T = (0, 1/2)$

$\nabla f(x_1)^T = (1,0)$

Now $x_2 = x_1 + t_1 H_1 \nabla f(x_1)$

where $H_1 = H_0 + A_1 + B_1$

Also $A_1 = \frac{(x_1 - x_0)(x_1 - x_0)^T}{(x_1 - x_0)^T(\nabla f(x_1) - \nabla f(x_0))}$

$B_1 = \frac{-H_0(\nabla f(x_1) - \nabla f(x_0))(\nabla f(x_1) - \nabla f(x_0))^T H_0}{(\nabla f(x_1) - \nabla f(x_0))^T H_0(\nabla f(x_1) - \nabla f(x_0))}$

$$= \begin{bmatrix} -\frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & -\frac{4}{5} \end{bmatrix}$$

$$H_1 = \begin{bmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & -\frac{1}{20} \end{bmatrix}$$

$x_2^T = (4t_1/5, 1/2 + 2t_1/5)$

Obtaining t_1 by an exact line search, we get $x_2^T = (1, 1)$

which is the required point of maxima.

We observe that for a quadratic function with two independent variables,the method converges to the optimum after two iterations.

By **BFGS update formula**

Here $f(x) = 2x_1 x_2 + 2x_2 - x_2^2 - 2x_2^2$

$\nabla f(x_1, x_2)^T = (-2x_1 + 2x_2, 2x_1 - 4x_2 + 2)$

$\nabla f(x_0)^T = \nabla f(0,0) = (0, 2)$

Now $x_1^T = x_0 + t_o H(I) \nabla f(x_0)$

$= (0, 2t_0)$

Let $f(t) = 4t_0 - 8t_0^2$

For optimum value, f'(t)=0

$t_0 = 1/4$

Also f''(t) =-16< 0

f(t) is maximum at $t_0 = 1/4$

$x_1^T = (0, 1/2)$

$\nabla f(x_1) = (1,0)$

Now $x_2 = x_1 + t_1 H_1 \nabla f(x_1)$

where $H_1 = H_0 - A_1 + B_1$

$A_1 = \frac{H_0(\nabla f(x_1) - \nabla f(x_0))(x_1 - x_0)^T + (x_1 - x_0)(\nabla f(x_1) - \nabla f(x_0))^T H_0}{(x_1 - x_0)^T(\nabla f(x_1) - \nabla f(x_0))}$

$$= \begin{bmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 2 \end{bmatrix}$$

$B_1 = \left\{1 + \frac{(\nabla f(x_1) - \nabla f(x_0))^T H_0(\nabla f(x_1) - \nabla f(x_0)}{(x_1 - x_0)^T(\nabla f(x_1) - \nabla f(x_0))}\right\} X \frac{(x_1 - x_0)(x_1 - x_0)^T}{(x_1 - x_0)^T(\nabla f(x_1) - \nabla f(x_0))}$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$H_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$$

Now $x_2 = \begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix} + t_1 \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$x_2^T = (t_1, 1/2 + 1/2\, t_1)$

Obtaining t_1 by an exact line search, we get

$x_2^T = (1, 1)$

which is the required point of maxima.

VI. Conclusion

Surprisingly not much attention has been given to maximization of multivariable nonlinear programming problems by the scholars. So we have chosen to study maximization problems and to our pleasant surprise, the results obtained are compatible with theoretical observations. In gradient search methods, the rate of convergence is slow and the result obtained is approximate to the optimal value. But in Newton's method and Quasi-Newton methods, the rate of convergence is faster and the results are accurate.

Acknowledgement

To all research scholars of this field and to Dr.William P.Fox and Dr.William H.Richardson in particular, whose works motivated us in preparing this research paper.

References

[1] Dr.William P. Fox and Dr.William H.Richardson, Department of Mathematics, Francis Marion University Florence,SC 29501; ' Comparison of NLP search methods in Nonlinear Optimization using MAPPLE'.
[2] Frederick S. Hillier, Gerald J.Lieberman; 'Introduction to Operations Research', Tata Mc Graw-Hill Publishing company limited.
[3] Wenyu Sun and YA-Xiang Yuan; 'Optimization theory and methods, Nonlinear Programming', Springer.
[4] M.Avriel; 'Nonlinear Programming: Analysis and methods', Prentice-Hall, Inc, 1976.
[5] R.Fletcher, John Willey and sons, Chichester; 'Practical methods of Optimization' , Vol.1, Unconstrained Optimization , 1981.
[6] W.C.Davidon; 'Variable metric methods for minimization', Argonne National Labs Report, ANL-5990, (1959).
[7] R.Fletcher and M.J.D.Powell; 'A rapid convergent descent method for minimization', Computer Journal 6(1963) 163-168.
[8] 'The convergence of a class double-rank minimization algorithms', Journal of the Institute of Mathematics and its Applications 6(1970) 76-90.
[9] R.Fletcher ; 'A new approach to variable metric algorithms', Computer J.13(1970) 317-322.
[10] D.Goldfarb.; 'A family of variable metric methods derived by variation mean', Mathematics of Computation 23(1970) 23-26 .
[11] D.F.Shanno; 'Conditioning of Quasi-Newton methods for function minimization', Math. Comput.24(1970) 647-656 .

Authors

Neha Varma, NET- JRF, Department of Mathematics, B.R.A. Bihar University, Muzaffarpur, Bihar, India
raginineha.v2@gmail.com
Dr. Ganesh Kumar, University Professor& Ex. H.O.D. of Mathematics, B.R.A. Bihar University, Muzaffarpur, Bihar, India.

Correspondence Author –. **Neha Varma**, NET- JRF, Department of Mathematics, B.R.A. Bihar University, Muzaffarpur, Bihar, India
raginineha.v2@gmail.com, sp.varma2shailesh@gmail.com, 09835815550