Journal of the Nigerian Association of Mathematical Physics
***Volume* 20 (March, 2012), pp 151 – 164**

The Accuracy and Effectiveness of Search Method for Optimizing A Multivariable Unimodal Function.

Famutimi Alaba Bukola

Department of Mathematics,
Faculty of Science, University of Ibadan, Ibadan, Nigeria.

Abstract

This work discusses the accuracy and effectiveness of search method for optimizing a multivariable unimodal function using various updates.

Keywords: Quasi-Newton's updates, Davidon- Fletcher powell updates, Powell Symmetric, Broyden's updates, Broyeden –Fletcher-Goldfarb-Shanno's updates

.

1.0 Introduction

Optimization is the collection processes of determining a set of conditions required to achieve the best result from a given situation [2], for any individual or organization. All of us make many decisions in the course of our day-to-day events in order to accomplish various tasks. Although some choices will generally be better than others, Consciously or unconsciously, we must decide upon the best or optimal way to realize our objectives.

In design, Construction, and maintenance of any engineering system, engineers have to take many technological and managerial decisions at several stages. The ultimate goal of all such decisions is to either minimize the effort required or maximize the desired benefit, since the effort required or benefit desired in any practical situation can be expressed as the process of finding the conditions that give the maximum or minimum value of a function.

This study deals with the accuracy and effectiveness of search method for optimizing a unimodal function. A function is said to be unimodal if it has only one maximum or minimum in the region to be searched.

In getting an optimal value of a unimodal function, we first conduct a search. Search techniques could either be classified as either sequential or preplanned (simultaneous) search. A sequential search involves step-by-step procedure and the outcome of an experiment is determined before another experiment is made. In a preplanned search technique, the location of the experiment is specified and the outcome of the measurement is obtained at the same time.

1.1 Methodology

A search method is employed for functional evaluations.It is sometimes difficult to determine the exact optimal point of a unimodal function due to various assumptions and experiments carried out.

Therefore, there is a need to compare search techniques used for evaluation of the optimal point of a unimodal function. This will determine the most effective and accurate one with fewer experiments.

There are essentially six (6) types of procedures to solve constrained nonlinear optimization problems. The three considered most successful are successive linear programming, successive quadratic programming and the generalized reduced-gradient method. The other three have not proved as useful, especially on problems with a large number of variables (more than twenty).These are penalty and barrier function methods, augmented Lagrangian functions and the methods of feasible directions (or projections) which are sometimes called methods of restricted movement. The reader interested in constrained multivariable search method can see [3]

Corresponding author: E-mail: famubukky@yahoo.co.uk, Tel. +2348062645294, +2348058018102

2.0 Unconstrained Multivariable Search Methods

2.1 Quasi-Newton's Method[5]

This method begins the search along a gradient line and use gradient information to build a quadratic fit to the economic model (profit function). Consequently, to understand these methods it is helpful to discuss the gradient search algorithm and Newton's method as background for the extension to the quasi-Newton algorithms. All of the algorithms involve a line search given by the following equation.

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha \boldsymbol{H}_k \nabla y(\boldsymbol{x}_k). \qquad (2.1.2)$$

The quasi-Newton algorithm that employs the BFGS(Broyden, Fletcher, Golfarb, Shanno) formula for updating the Hessian matrix is considered to be the most effective of the unconstrained multivariable search techniques, according to Fletcher[5]. This method is an extension of the DFP (Davidon, Fletcher, Powell) method.

2.2 Davidon Fletcher Powell Methods (DFP) [4]

The DFP algorithm has the following form of equation for minimizing the function y(**x**).

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha \boldsymbol{H}_k \nabla y(\boldsymbol{x}_k).$$

Wherex_{k+1} is the parameter of the line through$\boldsymbol{x}_k, and \boldsymbol{H}_k$is given by the following equation(2.2.1).

$$\boldsymbol{H}_k = \boldsymbol{H}_{k-1} + \boldsymbol{A}_k + \boldsymbol{B}_k \qquad (2.2.1)$$

The matrices $\boldsymbol{A}_k and \boldsymbol{B}_k$ are given by the following equations.

$$A_k = \frac{(x_k - x_{k-1})(x_k - x_{k-1})^T}{(x_k - x_{k-1})^T(\nabla y(x_k) - \nabla y(x_{k-1})} \qquad (2.2.2)$$

$$B_k = \frac{-H_{k-1}(\nabla y(x_k) - \nabla y(x_{k-1})(\nabla y(x_k) - \nabla y(x_{k-1})^T H_{k-1}}{((\nabla y(x_k) - \nabla y(x_{k-1})^T H_{k-1}\left(\nabla y(x_k) - \nabla y(x_{k-1})\right)} \qquad (2.2.3)$$

The algorithm begins with a search along the gradient line from the starting point $\mathbf{x}_0$ as given by the following equation obtained from equation, with $k = 0$.

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 - \alpha_1 \boldsymbol{H}_0 \nabla y(\boldsymbol{x}_0). \qquad (2.2.4)$$

2.3 Powell's Symmetric Methods [6]

In Powell's algorithm [6] the procedure begins at a starting point $\mathbf{x}_0$, and each application of the algorithm consists of (n +2) successive exact line searches. The first (n +1)are along the *n* coordinate axes. The (n+2)nd line search goes from the point obtained from the first line search through the best point (obtained at the end of the (n +1) line searches). If the function is quadratic, this will locate the optimum. If it is not, then the search is continued with one of the first *n* directions replaced by the $(n + 1)^{th}$ direction, and the procedure is repeated until a stopping criterion is met. In Powell's method, the conjugate directions are the orthogonal coordinate axes initially, and in steep ascent pattern the conjugate directions are the gradient lines.

Powell's Method For A General Function:

0. Calculate α_1 so that $y(\boldsymbol{x}_1 + \alpha_1 \boldsymbol{s}_n)$ is a minimum, and define $\mathbf{x}_o = \mathbf{x}_I + \boldsymbol{\alpha_1 s_n}$
1. For j = 1,2,...,n:
   Calculate α_j so that $y(\boldsymbol{x}_{j-1} + \alpha_j s_j)$ is a minimum.
   Define $\boldsymbol{x}_j = \boldsymbol{x}_{j-1} + \alpha_j \boldsymbol{s}_j$
   Replace $\boldsymbol{s}_j \; with \; \boldsymbol{s}_{j+1}$
2. Replace $\boldsymbol{s}_n \; with \; \boldsymbol{x}_n - \boldsymbol{x}_o$
3. Choose α so that $y\,[\alpha\,(\boldsymbol{x}_n - \boldsymbol{x}_0)]$ is a minimum, and replace $\mathbf{x}_o$ with $\mathbf{x}_o + \alpha(\mathbf{x}_n - \mathbf{x}_o)$.
4. Repeat steps 1-3 until a stopping criterion is met.

For a quadratic function the method will arrive at the minimum on completing Step 3. For a general function Steps 1-3 are repeated until a stopping criterion is satisfied. Step 0 is required to start the method by having $\boldsymbol{x}_o$, the point beginning the iteration steps 1-3, be a minimum point on the contour tangent line $\boldsymbol{s}_n$. The following example illustrates the above procedure for a quadratic function with two independent variables.

3.0 Examples of Unconstrained Multivariable Search Methods

We are considering two examples, one equation for two unconstrained search methods in other to determine the most accurate and effective among the updates.

Example 1:

Search for the minimum of the following function using the DFP algorithm starting at

` $\boldsymbol{x}_0^T = (0,0,0)$.

$$Minimise \quad y(x) = 7X_1^2 + 2X_2^2 + 2X_3^2 + 2X_1X_2 + 2X_2X_3 - 2X_1X_3 - 8X_3$$

$$\nabla y(x) = \begin{bmatrix} 14X_1 + 2X_2 - 2X_3 \\ 2X_1 + 4X_2 + 2X_3 \\ -2X_1 + 2X_2 + 4X_3 - 8 \end{bmatrix}$$

$$H = \begin{bmatrix} 14 & 2 & -2 \\ 2 & 4 & 2 \\ -2 & 2 & 4 \end{bmatrix}$$

at the starting point $\nabla y(x)$ becomes:

$$X_1 = X_{10} + \alpha \frac{\partial y}{\partial x_1}(X_0)$$

$$X_2 = X_{20} + \alpha \frac{\partial y}{\partial x_2}(X_0)$$

$$X_3 = X_{30} + \alpha \frac{\partial y}{\partial x_3}(X_0)$$

$$\frac{\partial y}{\partial X_1}(X_0) = 0, \quad \frac{\partial y}{\partial X_2}(X_0) = 0, \frac{\partial y}{\partial X_3}(X_0) = -8$$

The formular for finding the gradient line is given by

$$X_{k+1} = X_k \quad \alpha H_k \nabla y(X_k)$$

(3.1)

$$X_1 = X_0 \quad \alpha H_0 \nabla y(X_0).$$

$$\begin{bmatrix} X_{11} \\ X_{21} \\ X_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \alpha \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 8\alpha_1 \end{bmatrix}$$

The optimal value of α_1 is determined by an exact line search using

$X_1 = 0,\ X_2 = 0,\ X_3 = 8\alpha_1$ as follows:

$$y(\alpha_1) = 2(8\alpha_1)^2 - 8(8\alpha_1) = 128\alpha_1^2 - 64\alpha_1$$

$$\frac{dy}{d\alpha_1} = 256\alpha_1 = 64 = 0$$

$$\alpha_1 = \frac{64}{256} = \frac{1}{4}$$

$$\begin{bmatrix} X_{11} \\ X_{21} \\ X_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} - \alpha \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

The value of X_1 is computed as follows:

$$X_0^T = (0,0,2)\ , \qquad \nabla y(X_1)^T = (-4, 4, 0),\ \ \nabla y(X_0)^T = (0,0,-8)$$

The algorithm continues using equation (3.1) for k=1

$$X_2 = X_1 - \alpha_2 H_1 \nabla y(X_1)$$

where $H_1 = H_0 + A_1 + B_1$ **(3.2)**

The general formular for A_k and B_k is given as follows:

$$A_k = \frac{(X_k - X_{k-1})(X_k - X_{k-1})^T}{(X_k - X_{k-1})^T(\nabla y(X_k) - \nabla y(X_{k-1}))} \qquad (3.3)$$

$$B_k = \frac{-H_{k-1}(\nabla y(X_k)) - \nabla y(X_{k-1})(\nabla y(X_k)) - \nabla y(X_{k-1})^T H_{k-1}^T}{(\nabla y(X_k) - \nabla y(X_{k-1}))^T H_{k-1}(\nabla y(X_k)) - \nabla y(X_{k-1})} \qquad (3.4)$$

for k=1 equations (3.3) and (3.4) becomes:

$$A_1 = \frac{(X_1 - X_0)(X_1 - X_0)^T}{(X_1 - X_0)^T(\nabla y(X_1)) - \nabla y(X_0)}$$

$$B_1 = \frac{H_0(\nabla y(X_1)) - \nabla y(X_0)(\nabla y(X_1)) - \nabla y(X_0)^T H_0^T}{\nabla y(X_1) - \nabla y(X_0)^T H_0(\nabla y(X_1)) - \nabla y(X_0)}$$

$$X_0 = (0,0,0)\ , \qquad X_1^T = (0,0,2)$$

$$\nabla y(X_1^T) = (-4,4,0)\ , \quad \nabla y(X_0^T) = (0,0,-8)$$

Consequently, we then obtain the following:

$$A_1 = \frac{\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} [0\ 0\ 2]}{[0\ 0\ 0] \begin{bmatrix} -4 \\ 4 \\ 8 \end{bmatrix}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}$$

$$B_1 = \frac{\begin{bmatrix} 01 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -4 \\ 4 \\ 8 \end{bmatrix} [-4\ 4\ 8] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}{[-4\ 4\ 8] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -4 \\ 4 \\ 8 \end{bmatrix}}$$

$$B_1 = \frac{\begin{bmatrix} 16 & -16 & -32 \\ -16 & 16 & 32 \\ -32 & 32 & 64 \end{bmatrix}}{96} = \begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

From (3.2)

$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix} + \begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

$$H_1 = \begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

$$\begin{bmatrix} X_{12} \\ X_{22} \\ X_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} - \alpha_2 \begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} -4 \\ 4 \\ 0 \end{bmatrix} \begin{bmatrix} \frac{4}{3}\alpha_2 \\ -\frac{4}{3}\alpha_2 \\ 2 - \frac{8}{3}\alpha_2 \end{bmatrix}$$

The optimal value of α_2 is determined by an exact line search using $X_1 = \frac{4}{3}\alpha_2$, $X_2 = -\frac{4}{3}\alpha_2$, $X_3 = 2 - \frac{8}{3}\alpha_2$ in the function being minimized to give:

$$y(\alpha_2) = \frac{368}{9}\alpha_2^2 - \frac{32}{3}\alpha - 8$$

$$\frac{dy}{d\alpha_1} = \frac{736}{9}\alpha_2 - \frac{32}{3} = 0$$

$$\alpha_2 = \frac{32}{3} \times \frac{9}{736} = \frac{3}{23}.$$

The Value of $x_2$is computed by substituting for α_2 in (3.2)

$$\begin{bmatrix} X_{12} \\ X_{22} \\ X_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} - \alpha_2 \begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} -4 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}\alpha_2 \\ -\frac{4}{3}\alpha_2 \\ 2 - \frac{8}{3}\alpha_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{23} \\ -\frac{4}{23} \\ \frac{38}{23} \end{bmatrix}$$

$$X_2^T = \left(\frac{4}{23}, -\frac{4}{23}, \frac{38}{23}\right), \quad \nabla y(X_2)^T = \left(-\frac{28}{23}, \frac{68}{23}, -\frac{48}{23}\right)$$

The computation of x_3 uses equation (3.1) as follows:
Where

$$\nabla y(X_2)^T = \left(-\frac{28}{23}, \frac{68}{23}, -\frac{48}{23}\right), \quad H_2 = H_1 + A_2 + B_2, \quad \nabla y(X_1)^T = (-4, 4, 0)$$

$$X_1 = (0,0,2), \quad X_2 = \left(\frac{4}{23}, -\frac{4}{23}, \frac{38}{23}\right)$$

When $K = 2$

$$A_2 = \frac{(X_2 - X_1)(X_2 - X_1)^T}{(X_2 - X_1)^T(\nabla y(X_2) - \nabla y(X_1))}$$

$$B_2 = \frac{H_1(\nabla y(X_2)) - \nabla y(X_1)(\nabla y(X_2)) - \nabla y(X_1)^T H_1^T}{\nabla y(X_2) - \nabla y(X_1)^T H_1(\nabla y(X_2 - \nabla y(X_1)))}.$$

$$A_2 = \frac{\begin{bmatrix} \frac{4}{23} \\ -\frac{4}{23} \\ -\frac{8}{23} \end{bmatrix} \begin{bmatrix} \frac{4}{23} & -\frac{4}{23} & -\frac{8}{23} \end{bmatrix}}{\begin{bmatrix} -\frac{4}{23} & -\frac{4}{23} & -\frac{8}{23} \end{bmatrix} \begin{bmatrix} \frac{64}{23} \\ -\frac{11}{7} \\ -\frac{48}{23} \end{bmatrix}} = \begin{bmatrix} \frac{28}{1373} & -\frac{28}{1373} & -\frac{568}{1373} \\ -\frac{28}{1373} & \frac{28}{1373} & \frac{56}{1373} \\ -\frac{56}{1373} & \frac{56}{1373} & \frac{112}{1373} \end{bmatrix}$$

$$B_2 = \frac{\begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{23}{12} \end{bmatrix} \begin{bmatrix} \frac{64}{23} \\ -\frac{11}{7} \\ -\frac{48}{23} \end{bmatrix} \begin{bmatrix} \frac{64}{23} & -\frac{11}{7} & -\frac{48}{23} \end{bmatrix} \begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{23}{12} \end{bmatrix}}{\begin{bmatrix} \frac{64}{23} & -\frac{11}{7} & -\frac{48}{23} \end{bmatrix} \begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{23}{12} \end{bmatrix} \begin{bmatrix} \frac{64}{23} \\ -\frac{11}{7} \\ -\frac{48}{23} \end{bmatrix}}$$

$$= \begin{bmatrix} \frac{1885129}{16391094} & -\frac{1885129}{16391094} & -\frac{3615109}{8195547} \\ -\frac{1885129}{16391094} & \frac{1885129}{16391094} & \frac{3615109}{8195547} \\ -\frac{3615109}{8195547} & \frac{3615109}{8195547} & \frac{13865378}{8195547} \end{bmatrix}$$

$$H_2 = \begin{bmatrix} \frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{23}{12} \end{bmatrix} + \begin{bmatrix} \frac{28}{1373} & -\frac{28}{1373} & -\frac{56}{1373} \\ -\frac{28}{1373} & \frac{28}{1373} & \frac{56}{1373} \\ -\frac{56}{1373} & \frac{56}{1373} & \frac{112}{1373} \end{bmatrix} \begin{bmatrix} \frac{1885129}{16391094} & -\frac{1885129}{16391094} & -\frac{3615109}{8195547} \\ -\frac{1885129}{16391094} & \frac{1885129}{16391094} & \frac{3615109}{8195547} \\ -\frac{3615109}{8195547} & \frac{3615109}{8195547} & \frac{13865378}{8195547} \end{bmatrix}$$

$$\begin{bmatrix} X_{13} \\ X_{23} \\ x_{33} \end{bmatrix} = \begin{bmatrix} \frac{4}{23} \\ \frac{-4}{23} \\ \frac{38}{23} \end{bmatrix} - \alpha_3 \begin{bmatrix} \frac{3399030713}{11252486031} & \frac{-3399030713}{11252486031} & \frac{-9173323966}{11252486031} \\ \frac{-3399030713}{11252486031} & \frac{3399030713}{11252486031} & \frac{9173323966}{11252486031} \\ \frac{-9173323966}{11252486031} & \frac{9173323966}{11252486031} & \frac{166089320603}{45009944124} \end{bmatrix} \begin{bmatrix} \frac{-28}{23} \\ \frac{68}{23} \\ \frac{-48}{23} \end{bmatrix}$$

$$\begin{bmatrix} X_{13} \\ X_{23} \\ x_{33} \end{bmatrix} = \begin{bmatrix} \frac{4}{23} \\ \frac{-4}{23} \\ \frac{38}{23} \end{bmatrix} - \alpha_3 \begin{bmatrix} \frac{27679680}{62832527} \\ \frac{-27679680}{62832527} \\ \frac{-270073500}{62832527} \end{bmatrix} = \begin{bmatrix} \frac{4}{23} - \frac{27679680}{62832527}\alpha_3 \\ \frac{-4}{23} + \frac{-27679680}{62832527}\alpha_3 \\ \frac{38}{23} + \frac{-270073500}{62832527}\alpha_3 \end{bmatrix}$$

The optimal value of α_3 is determined by an exact line search using:

$$X_1 = \frac{4}{23} - \frac{27679680}{62832527}\alpha_3$$
$$X_2 = \frac{-4}{23} + \frac{27679680}{62832527}\alpha_3$$

$$X_3 = \frac{38}{23} + \frac{270073500}{62832527}\alpha_3$$
$$y(\alpha_3) = \frac{1811447358247368000}{3947926449205729}\alpha_3^2 - \frac{10306278720}{1445148121}\alpha - \frac{200}{23} = 0.0777146$$

The value of α_3 is determined as previously by setting $\frac{dy}{d\alpha_3}(\alpha_3) = 0$ and the optimal value of $X_3^T = (0.139677, -0.139677, 1.98622)$ is computed which is the value of the function at the minimum.

Example 2:

Determine the minimum of the following function using the BFGS algorithm starting at
$X_0 = (0,0,0)$

$Minimise \quad y(x) = 7X_1^2 + 2X_2^2 + 2X_3^2 + 2X_1X_2 + 2X_2X_3 - 2X_1X_3 - 8X_3$

The first application of the algorithm is the same as the example above, which is a search along gradient line through $x_0 = (0,0,0)$. The results were:

$$X_1^T = (0,0,2), \qquad \nabla y(X_1)^T = (-4,4,0)$$
$$X_0^T = (0,0,0), \qquad \nabla y(X_0)^T = (0,0,-8)$$

The algorithm continues using equation (3.1) for $k = 1$

$$X_2 = X_1 - \alpha H_1 y(X_1)$$

where H_1 is given as,

$$H_1 = H_0 = \left[\frac{H_0 Y_0 \partial_0^T + \partial_0 Y_0^T H_0}{\partial_0^T Y_0}\right] + \left[1 + \frac{Y_0^T H_0 Y_0}{\partial_0^T Y_0}\right] \frac{\partial_0 \partial_0^T}{\partial_0^T Y_0} \tag{3.5}$$

Where

$$H_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \;, \; \delta_0 = X_1 - X_0 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

$$Y_0 = \nabla y(X_1) - \nabla y(X_0) = \begin{bmatrix} -4 \\ 4 \\ 8 \end{bmatrix} \;, \; \delta_0^T Y_0 = 16 \;, \; Y_0^T H_0 Y_0 = 96.$$

$$H_1 = \frac{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -4 \\ 4 \\ 8 \end{bmatrix} [0\ 0\ 2] + \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} [\ 4\ 4\ 8] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}{16} + \left[1 + \frac{96}{16}\right] \frac{\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} [0\ 0\ 2]}{16}$$

$$H_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & -\frac{1}{2} \\ 0 & 0 & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{7}{4} \end{bmatrix}$$

$$\begin{bmatrix} X_{12} \\ X_{22} \\ X_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \quad \alpha_2 \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} -4 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 4\alpha_2 \\ 4\alpha_2 \\ 2+4\alpha_2 \end{bmatrix}$$

The optimal value of α_2 is determined by an exact line search $X_1 = 4\alpha_2$,

$X_2 = 4\alpha_2$, $X_3 = 2 + 4\alpha_2$ in the function being minimized to give

$$\begin{aligned} y(\alpha_2) &= 80\alpha^2 - 32\alpha - 8 \\ \frac{dy}{d\alpha_2} &= 160\alpha - 32 = 0 \\ \alpha &= \frac{32}{160} = \frac{1}{5} \end{aligned}$$

$$\begin{bmatrix} X_{12} \\ X_{22} \\ X_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} - \alpha_2 \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} -4 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 4\alpha_2 \\ -4\alpha_2 \\ 2+4\alpha_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} \\ -\frac{4}{5} \\ \frac{14}{5} \end{bmatrix}$$

The value of X_2 is computed by substituting for α_2 in (3.2)

$$X_2^T = \left(\frac{4}{5}, -\frac{4}{5}, \frac{14}{5}\right) \ , \ \nabla y(X_2)^T = (4,4,0)$$

The computation of X_3 uses the application of the algorithm as follows:

$$X_3 = X_2 - \alpha_3 H_2 \nabla y(X_2).$$

where

$$H_2 = H_1 - \left[\frac{H_1 Y_1 \partial_1^T + \partial_1 Y_1^T H_1}{\partial_1^T Y_1}\right] + \left[1 + \frac{Y_1^T H_1 Y_1}{\partial_1^T Y_1}\right] \frac{\partial_1 \partial_1^T}{\partial_1^T Y_1}$$

where

$$H_0 = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{3}{4} \end{bmatrix}, \quad \delta_1 = X_2 - X_1 = \begin{bmatrix} \frac{4}{5} \\ -\frac{4}{5} \\ \frac{4}{5} \end{bmatrix}, \quad Y_1^T = \Delta y(X_2) - \Delta y(X_1) = (8,0,0)$$

$$\delta_1^T Y_1 = \frac{32}{5}, \quad Y_1^T H_1 Y_1 = 64$$

$$H_2 = \frac{\begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} 8 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \frac{4}{5} & -\frac{4}{5} & \frac{4}{5} \end{bmatrix} + \begin{bmatrix} \frac{4}{5} \\ -\frac{4}{5} \\ \frac{4}{5} \end{bmatrix} \begin{bmatrix} 8 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{3}{4} \end{bmatrix}}{\frac{32}{5}}$$

$$\left[1 + \frac{64}{\frac{32}{5}}\right] \frac{\begin{bmatrix} \frac{4}{5} \\ -\frac{4}{5} \\ \frac{4}{5} \end{bmatrix} \begin{bmatrix} \frac{4}{5} & -\frac{4}{5} & \frac{4}{5} \end{bmatrix}}{\begin{bmatrix} \frac{4}{5} & -\frac{4}{5} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} 8 \\ 0 \\ 0 \end{bmatrix}}$$

$$\begin{bmatrix} X_{13} \\ X_{23} \\ X_{33} \end{bmatrix} = \begin{bmatrix} \frac{4}{5} \\ -\frac{4}{5} \\ \frac{4}{5} \end{bmatrix} - \alpha_2 \begin{bmatrix} -\frac{107}{10} & \frac{53}{10} & -8 \\ -\frac{53}{10} & \frac{21}{10} & \frac{8}{5} \\ -8 & \frac{8}{5} & -\frac{91}{20} \end{bmatrix} \begin{bmatrix} 4 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} + \frac{108}{5}\alpha_3 \\ -\frac{4}{5} - \frac{148}{5}\alpha_3 \\ \frac{14}{5} + \frac{128}{5}\alpha_3 \end{bmatrix}$$

$$H_2 = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{3}{4} \end{bmatrix} - \begin{bmatrix} \frac{64}{5} & -\frac{32}{5} & \frac{48}{5} \\ -\frac{32}{5} & 0 & -\frac{16}{5} \\ \frac{48}{5} & -\frac{16}{5} & \frac{32}{5} \end{bmatrix} + \begin{bmatrix} \frac{11}{10} & -\frac{11}{10} & \frac{11}{10} \\ -\frac{11}{10} & \frac{11}{10} & -\frac{11}{10} \\ \frac{11}{10} & -\frac{11}{10} & \frac{11}{10} \end{bmatrix}$$

The optimal value of α_3 is determined by an exact line search using $X_1 = \frac{4}{5} + \frac{108}{5}\alpha_3$, $X_2 = -\frac{4}{5} - \frac{148}{5}\alpha_3$, $X_3 = \frac{14}{5} + \frac{128}{5}\alpha_3$ in the function being minimized to give:

$$y(\alpha_3) = \frac{12144}{5}\alpha^2 - 32\alpha - \frac{56}{5}$$

$$\frac{dy}{d\alpha_3} = \qquad \alpha_3 = \frac{5}{759}$$

Setting $\frac{dy}{d\alpha}(\alpha_3) = 0$ and solving for α_3 gives the optimal value of

$X_3^T = (0.942, -0.995, 2.969)$ is computed, which is the value of the function

at the minimum.

The value of the function at the minimum for BFGS and DFP in Tabular form

	$BFGS$	DFP
X_1	0.942	0.139677
X_2	−0.995	−0.139677
X_3	2.969	1.98622

The quasi-Newton algorithm that employs the BFGS(Broyden, Fletcher, Golfarb, Shanno) formula for updating the Hessian matrix is considered to be the most effective of the unconstrained multivariable search techniques, according to [5]. This method is an extension of the DFP (Davidon, Fletcher, Powell) method.

The choice of equations taken into consideration is of paramount importance. This paper was convered from Multivariable Optimization Procedure[1]

Determine the minimum of the following function using the DFP algorithm starting at $\boldsymbol{x}_0^T = (0,0,0)$.
Minimize: $5X_1^2 + 2X_2^2 + 2X_3^2 + 2X_1X_2 + 2X_2X_3 - 2X_1X_3 - 6X_3$
Performing the appropriate partial differentiation, the gradient vector $\nabla y(\mathbf{x})$ and the Hessian matrix are:

$$\nabla y(x) = \begin{bmatrix} 10x_1 + 2x_2 - 2x_3 \\ 2x_1 + 4x_2 + 2x_3 \\ -2x_1 + 2x_2 + 4x_3 - 6 \end{bmatrix}, \qquad H = \begin{bmatrix} 10 & 2 & -2 \\ 2 & 4 & 2 \\ -2 & 2 & 4 \end{bmatrix}$$

The optimal value of α_1 is determined by an exact line search using $x_1 = 0$, $x_2 = 0$, $x_3 = 6\alpha_1$ as follows:

$$y(\alpha_1) = 2(6\alpha_1)^2 - 6(6\alpha_1) = 72\alpha_1^2 - 36\alpha_1$$

$$\frac{dy}{d\alpha_1} = 144\alpha_1 - 36 = 0 \rightarrow \alpha_1 = 1/4$$

The value of $\mathbf{x}_1$ is computed by substituting for $_1$ in the previous equation.

$$\boldsymbol{X}_1^T = \left(0, 0, \frac{3}{2}\right), \nabla y(X_1)^T = (-3, 3, 0), \nabla y(\boldsymbol{X_0})^T = (0, 0, -6)$$

The algorithm continues using equations (3.1) for $k = 1$.

$$\boldsymbol{X_2} = \boldsymbol{X_2} - \alpha_2 \boldsymbol{H}_1 \nabla y(\boldsymbol{X}_1)$$

The optimal value of α_2 is determined by an exact line search as follows.

$y(\alpha_2) = 12\alpha_2^2 - 12\alpha_2 + 9/2$

$$\frac{dy}{d\alpha_2} = 24\alpha_2 - 12 = 0 \rightarrow \alpha_2 = ½$$

The value of $\mathbf{x}_2$ is computed by substituting for α_2 in (3.2).

$\mathbf{x}_2^T = (1, -1, 5/2)$, $\nabla y(\mathbf{x}_2)^T = (3, 3, 0)$

The computation of $\mathbf{x}_3$ uses equations (3.1) as follows:

$\mathbf{x}_3 = \mathbf{x}_2 - \alpha_3 \mathbf{H}_2 \nabla y(\mathbf{x}_2)$

andthe optimal value of α_3 is determined by an exact line search as follows:
$y(\alpha_3) = 5 + 2(1 + 12\alpha_3/5)^2 + 2(5/2 + 6\alpha_3/5)^2 - 2(1 + 12\alpha_3/5) - 2(1 + 12\alpha_3/5)(5/2 + 6\alpha_3\backslash 5) - 2(5/2 + 6\alpha_3/5) - 6(5/2 + 6\alpha_3/5)$
Setting $\frac{dy(\alpha_3)}{d\alpha_3} = 0$ and solving for α_3 gives $\alpha_3 = \frac{5}{12}$ and $\boldsymbol{x}_3^T = (1, -2, 3)$ which is the value of the function at the minimum.
Determine the minimum of the following function using the BFGS algorithm starting at $\mathbf{x}_0 = (0,0,0)$.
Minimize: $5x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_2x_3 - 2x_1x_3 - 6x_3$
The first application of the algorithm is the same as above, which is a search along the gradient line through $\mathbf{x}_0 = (0,0,0)$.
These results were:
$\mathbf{x}_1^T = (0, 0, 3/2)$, $\nabla y(\mathbf{x}_1)^T = (-3, 3, 0)$

$\mathbf{x}_0^T = (0, 0, 0)$, $\nabla y(\mathbf{x}_0)^T = (0, 0, -6)$

The algorithm continues using equations (3.1) for k=1.
$\mathbf{x}_2 = \mathbf{x}_1 - \alpha_2 \mathbf{H}_1 \nabla y(\mathbf{x}_1)$

The optimal value of α_2 is determined by an exact line search using $x_1 = 3\alpha_2$, $x_2 = -3\alpha_2$, $x_3 = 3/2 + 3\alpha_2$ in the function is minimized to give:

$y = 27\alpha_2^2 - 18\alpha_2 + 4½$

$$\frac{dy}{d\alpha_2} = 54\alpha_2 - 18 = 0 \rightarrow \alpha_2 = 1/3$$

The value for $\mathbf{x}_2$ is computed by substituting for α_2 in
$\mathbf{x}_2^T = (1, -1, 5/2), \nabla y(\mathbf{x}_2)^T = (3, 3, 0)$

The computation of $\mathbf{x}_3$ repeats the application of the algorithm as follows:
$\mathbf{x}_3 = \mathbf{x}_2 - \alpha_3 \mathbf{H}_2 \nabla y(\mathbf{x}_2)$

The optimal value of α_3 is determined by an exact line search using $x_{13} = 1$, $x_{23} = -1\text{-}6\alpha_3$, $x_{33} = 5/2 + 3\alpha_3$ in the function being minimized to give $y(\alpha_3)$. The value of $\alpha_3 = 1/6$ is determined as previously by setting $dy(\alpha_3)/d\alpha_3 = 0$, and the optimal value of $\mathbf{x}_3^T = (1,-2,3)$ is computed, which is the value of the function at the minimum.

4.0 Conclusion

The quasi-Newton algorithm that employs the BFGS(Broyden, Fletcher, Golfarb, Shanno) formula for updating the Hessian matrix is considered to be the most effective of the unconstrained multivariable search techniques, according to [5]. This method is an extension of the DFP (Davidon, Fletcher, Powell) method.

Deterministic problems have no experimental error or random factors present. Example is the mathematical model of a process, where the outputs are calculated by a computer program from specified inputs. Stochastic problems have random error present, usually in the form of experimental errors from measurement of the process variables. An example is a plant where there are experimental errors associated with laboratory and instrument measurements of the process flow rates and composition.

Search plans can be classified as either simultaneous or sequential. In a simultaneous search plan the locations of all of the experiments are specified, and the outcome of the measurements is obtained at the same time. An example is the location of a set of thermocouples installed along the length of a fixed-bed reactor to determine the position of the hot spot (maximum temperature) in the catalyst bed while the process is in operation. In a sequential search plan the outcome of an experiment is determined before another experiment is made. Being able to base future experiments on past outcomes is a significant advantage. In fact, it can be shown that the advantage of sequential search plans over simultaneous search plans increases exponentially with the number of experiments.

References

[1] Avriel, Mordecai, Nonlinear Programming: Methods and Analysis, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1976).

[2] Cooper, Leon, and David Steinberg, Introduction to Methods of Optimization, W. B. Saunders Company, Philadelphia (1970).

[3] Griffith, R. E. and R. A. Stewart, " A Nonlinear Programming Technique for the Optimization of Continous Processing Systems", Management Science, Vol. 7, p. 379 (1961).

[4] Fletcher, Roger, Practical Methods of Optimization, Vol. I, Unconstrained Optimization, John Wiley and Sons, Inc., New York(1981).

[5] Fletcher, Roger, Practical Methods of Optimization, Vol. II, Constrained Optimization, John Wiley and Sons, Inc., New York (1981).

[6] Powell, M. J. D., "An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives," The Computer Journal, Vol. 7, p. 155 (1964).