



**Optimization**  
**Fourth Class**  
**2020 - 2021**



**By**

***Dr. Jawad Mahmoud Jassim***

**Dept. of Math.**

**Education College for Pure Sciences**

***University of Basrah***

**Iraq**



# Chapter Five



# Quasi – Newton Methods

## Lecture 1

## 1: Introduction:

Let  $f: R^n \rightarrow R$  have continuous second partial derivatives in a convex set  $D \subset R^n$  and let the Hessian matrix  $G(X)$  be positive definite *for all*  $X \in D$ . Also let  $X^*$  be a critical point of  $f$  in  $D$ .

Then the general quasi – Newton method for minimizing  $f$  which is given in the following algorithm.



## Algorithm (1):

It is assumed that an estimate  $X_0$  of a minimizer  $X^*$  of  $f$  and a matrix  $H_0$  are given.

1: Set  $k = 0$ .

2: Compute  $f(X_k)$  and  $g_k = g(X_k)$ , where  $g_k$  is the gradient vector of  $f$  at  $X_k$ .

3: Compute the search direction  $p_k$  as:  $p_k = -H_k g_k$ .

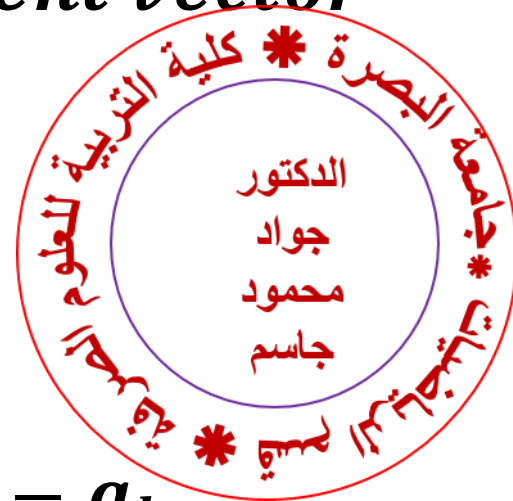
4: Compute  $\alpha_k$  such that  $f(X_k + \alpha_k p_k) = \min_{\alpha} f(X_k + \alpha p_k)$ .

5: Compute  $X_{k+1}$  as:  $X_{k+1} = X_k + \alpha_k p_k$ .

6: Compute  $g_{k+1} = g(X_{k+1})$ ,  $s_k = X_{k+1} - X_k$  and  $Y_k = g_{k+1} - g_k$ .

7: Compute  $H_{k+1}$  such that  $H_{k+1} = H_k + C_k$ , where  $C_k$  is such that  $H_{k+1}$  is positive definite and satisfy *the quasi - Newton equation which is defined by  $H_{k+1} Y_k = s_k, k \geq 0$ .*

8: Set  $k = k + 1$  and go to step 3.



## Note (1):

In Algorithm (1),  $C_k$  is not defined precisely by the requirements listed in step 7. Therefore Algorithm (1) contains a class of methods rather than a single quasi – Newton method. Many updating formula of type  $H_{k+1} = H_k + C_k$  have been proposed.



## Theorem (1):

The search direction  $p_k = -H_k g_k$  is downhill direction if  $H_k$  is positive definite matrix.

## Proof: (H.W.)

Some of the most first quasi – Newton methods will be described in the following sections.

## 2: Davidon – Fletcher – Powell Method (DFP)

It is one of the best methods for unconstrained minimization in which the gradient vector of the objective function is required. The DFP method is interesting also because it is at once a conjugate direction method and a quasi-Newton method, as will be established in this section.

As stated in section 1,  $H_{k+1}$  is not uniquely defined by the requirement that it satisfy the quasi – Newton equation.

Indeed it is easily shown that a class of updating formula for  $H_k$  every member of which satisfies the quasi – Newton equation is defined by



$$H_{k+1} = H_k + s_k u_k^T - H_k y_k v_k^T \quad \dots \dots \dots (1)$$

Where  $u_k$  and  $v_k$  are any nonzero vectors such that

$$u_k^T y_k = 1 \quad \text{and} \quad v_k^T y_k = 1 \quad \dots \dots \dots (2)$$

$$\begin{aligned} \therefore H_{k+1} y_k &= H_k y_k + s_k u_k^T y_k - H_k y_k v_k^T y_k \\ &= H_k y_k + s_k - H_k y_k = s_k . \end{aligned}$$

∴ The formula (1) satisfies the quasi – Newton equation.

Let

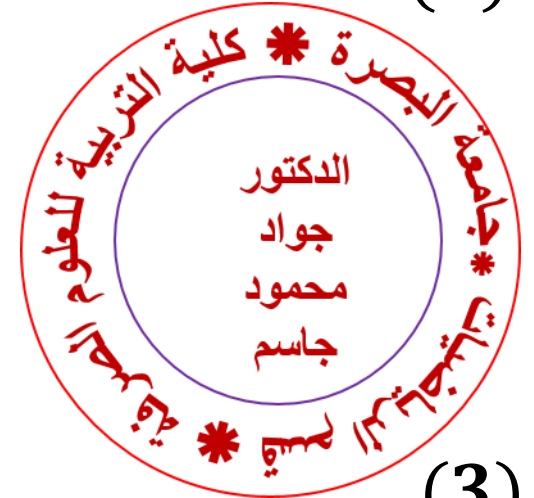
$$u_k = \frac{s_k}{s_k^T y_k} \quad \text{and} \quad v_k = \frac{H_k y_k}{y_k^T H_k y_k} \quad \dots \dots \dots (3)$$

∴  $u_k$  and  $v_k$  defined by (3) satisfy (2).

Therefore the matrix  $H_{k+1}$  given by updating formula ((Equation (1)))

$$H_{k+1} = H_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{(H_k y_k)(H_k y_k)^T}{y_k^T H_k y_k} \quad \dots \dots \dots (4)$$

∴  $H_{k+1}$  in (4) satisfies the quasi – Newton equation.



**Now, we give the DFP Algorithm.**

**Algorithm (2):**

It is assumed that an estimate  $X_0$  of a minimizer  $X^*$  of  $f: R^n \rightarrow R$  and a symmetric positive definite matrix  $H_0$  are known, where  $H_0 = I$  is an identity matrix.

**1: Set  $k = 0$ .**

**2: Compute  $g_k$  from  $g_k = g(X_k)$ .**

**3: Compute  $p_k$  from  $p_k = -H_k g_k$ .**

**4: Compute  $\alpha_k, X_{k+1}$  from  $f(X_k + \alpha_k p_k) = \min_{\alpha} f(X_k + \alpha p_k)$  and**

$$X_{k+1} = X_k + \alpha_k p_k .$$

**5: Compute  $g_{k+1}$  from  $g_{k+1} = g(X_{k+1})$ .**

**6: Compute  $s_k, y_k$  from  $s_k = \alpha_k p_k$  and  $y_k = g_{k+1} - g_k$ .**





**7: If convergence is considered to have been obtained, then go to step 11.**

**8: Compute  $z_k$  from  $z_k = H_k y_k$ .**

**9: Compute  $H_{k+1}$  from  $H_{k+1} = H_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{z_k z_k^T}{y_k^T z_k}$ .**

**10: Set  $k = k + 1$  and go to step 3.**

**11: Set  $X^* = X_{k+1}$ .**

**12: Stop.**

