Diagonal – BFGS Update method

Shahad Jassim Hammoud **Dr. Saad Shakir**
Almustansiriyah University, College of Education, Department of Mathematics

shjassem21@gmail.com
saadshakirmahmood@yahoo.com

Abstract:
The aim of this research is to introduce a new technique to solve the un-constrained optimization problem with large dimension where the Hessian matrix becomes sparse matrix. The BFGS update is modified to a diagonal update, so called the $Diagonal - BFGS$ update. Moreover that it is symmetric and positive definite are given. Give a numerical example to illustrate the effectiveness of technical.

Keyword: Quasi Newton method, Hessian matrices and $BFGS$ method.

طريقة تحديث $BFGS$ القطرية

شهد جاسم حمود **د. سعد شاكر محمود**
الجامعة المستنصرية، كلية التربية، قسم الرياضيات

الخلاصة :
الهدف من هذا البحث هو تقديم تقنية جديدة لحل مشكلة التحسين غير المقيدة بأبعاد كبيرة حيث تصبح مصفوفة هيسين مصفوفة متفرقة. تم تطوير تحديث BFGS إلى تحديث قطري ، ويسمى ذلك بتحديث BFGS القطري . علاوة على ذلك، يتم إعطاء التناظر و تعريف إيجابي. إعطاء مثال عددي لتوضيح فاعلية التقنية.

الكلمات المفتاحية : طريقة شبه نيوتن، مصفوفات هيس، طريقة BFGS.

1. Introduction

Today, Quasi-Newton methods are considered one of the most Efficient methods for solving nonlinear unconstrained or bounded constrained optimization problems. These methods are often used when the second derivative matrix of the objective function is either inaccessible or too expensive to compute. They are quite Similar to Newton's method, But escape the need of computing Hessian matrices by repeating, from Iteration to Iteration, a symmetric Matrix, which can be considered as an approximation of the Hessian. They allow, therefor, the curvature of the problem to be exploited in the numerical Algorithm, despite the fact that only first derivative (gradients) and function values are required [2]. The Quasi-Newton methods are very useful and Efficient methods for solving the unconstrained minimization problem

$$\min f(x) \; ; \; x \in \mathbb{R}^n$$

Where $f: \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable. Starting from point x_0 and a symmetric and positive definite Matrix $\mathcal{B}_0$, a quasi-Newton

In this sense, they may describe BFGS method .This project includes a development, execution and testing of new method like which use data from several recent steps in carrying out the updating. The formula of BFGS update method which is

$$\mathcal{B}_{k+1} = \mathcal{B}_k - \frac{\mathcal{B}_k s_k s_k^T \mathcal{B}_k}{s_k^T \mathcal{B}_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} \quad \ldots. (1.1)$$

Some researchers have also presented some important research see [3],[4],[5],[6]

We introduce $\text{Diagonal} - \text{BFGS}$ update is propose to solve the unconstrained optimization problem with large dimension, moreover the prove of symmetric and positive definite are given.

2. Diagonal – BFGS update method

Given $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$, f is assumed to be continuous and differentiable on D. The problem is to generate a sequence of hessian matrix which has a diagonal form to minimize the compilation of element of hessian matrix and to solve the unconstrained optimization problem with large dimension.

Let

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \text{ and } y^T = (y_1 \quad \ldots \quad y_n) \quad \ldots. (2.1)$$

$$\text{Now } yy^T = \begin{pmatrix} y_1^2 & \cdots & y_1 y_j & \cdots & y_1 y_n \\ \vdots & & & & \vdots \\ y_i y_1 & \cdots & y_i^2 & \cdots & y_i y_j \\ \vdots & & & & \vdots \\ y_n y_1 & \cdots & y_j y_i & \cdots & y_n^2 \end{pmatrix} \quad \ldots. (2.2)$$

$$\text{Let } \acute{y} = \begin{pmatrix} \acute{y}_1 \\ \vdots \\ \acute{y}_n \end{pmatrix} \text{ and } \acute{y}^T = (\acute{y}_1 \quad \ldots \quad \acute{y}_n)$$

$$\acute{y}_k \acute{y}_k^T = \begin{pmatrix} \acute{y}_1\acute{y}_1 & \cdots & \acute{y}_1\acute{y}_j & \cdots & \acute{y}_1\acute{y}_n \\ \vdots & & & & \vdots \\ \acute{y}_i\acute{y}_1 & \cdots & \acute{y}_i\acute{y}_i & \cdots & \acute{y}_i\acute{y}_j \\ \vdots & & & & \vdots \\ \acute{y}_n\acute{y}_1 & \cdots & \acute{y}_j\acute{y}_i & \cdots & \acute{y}_n\acute{y}_n \end{pmatrix} \quad \ldots. (2.3)$$

where $i = 1,2, \ldots, n \;\; and \;\; j = 1,2, \ldots, n$

$$\acute{y}_i\acute{y}_j = \begin{cases} \alpha_{ij} y_i y_j & , if \; i \neq j \\ y_i y_j & , if \; i = j \; (i.e. \; \acute{y}_i\acute{y}_i = y_i y_i = y_i^2) \end{cases} \quad \ldots. (2.4)$$

We substitute (2.4) in (2.3), to get

$$\acute{y}_k\acute{y}_k^T = \begin{pmatrix} y_1^2 & .. & \alpha_{1j}y_1y_j & .. & \alpha_{1n}y_1y_n \\ \vdots & & & & \vdots \\ \alpha_{ij}y_iy_j & .. & y_i^2 & .. & \alpha_{ij}y_iy_j \\ \vdots & & & & \vdots \\ \alpha_{n1}y_ny_1 & .. & \alpha_{nj}y_ny_j & .. & y_n^2 \end{pmatrix} \quad \text{…. (2.5)}$$

Where α_{ij} is the value can be determined to guarantee the symmetric and positive definite property of $\mathcal{B}_{k+1}$ for the $Diagonal - BFGS$ update

Now, Replace $\acute{y}\acute{y}^T$ by yy^T in (1.1), we have

$$\mathcal{B}_{k+1} = \mathcal{B}_k + \frac{\acute{y}\acute{y}^T}{\mathcal{S}_k^T y_k} - \frac{\mathcal{B}_k\mathcal{S}_k\mathcal{S}_k^T\mathcal{B}_k}{\mathcal{S}_k^T\mathcal{B}_k\mathcal{S}_k} \quad \text{….. (2.6)}$$

Now, we will explain how to extract the value α_{ij} that we used to obtain the formula $\mathcal{B}_{k+1}$ in (1.1) as shown below:

Where $\mathcal{B}_k\mathcal{S}_k\mathcal{S}_k^T\mathcal{B}_k = \begin{pmatrix} c_1c_1 & .. & c_1c_j & .. & c_1c_n \\ \vdots & & & & \vdots \\ c_ic_j & .. & c_ic_i & .. & c_ic_j \\ \vdots & & & & \vdots \\ c_nc_1 & .. & c_nc_i & .. & c_nc_n \end{pmatrix}$

$$\mathcal{B}_{k+1} = \mathcal{B}_k + \frac{\begin{pmatrix} y_1^2 & .. & \alpha_{1j}y_1y_j & .. & \alpha_{1n}y_1y_n \\ \vdots & & & & \vdots \\ \alpha_{ij}y_iy_j & .. & y_i^2 & .. & \alpha_{ij}y_iy_j \\ \vdots & & & & \vdots \\ \alpha_{n1}y_ny_1 & .. & \alpha_{nj}y_ny_j & .. & y_n^2 \end{pmatrix}}{\mathcal{S}_k^T y_k} - \frac{\begin{pmatrix} c_1c_1 & .. & c_1c_j & .. & c_1c_n \\ \vdots & & & & \vdots \\ c_ic_j & .. & c_ic_i & .. & c_ic_j \\ \vdots & & & & \vdots \\ c_nc_1 & .. & c_jc_i & .. & c_nc_n \end{pmatrix}}{\mathcal{S}_k^T\mathcal{B}_k\mathcal{S}_k} \quad \text{….(2.7)}$$

To find α_{ij}

$$\alpha_{ij}\ \frac{y_iy_j}{\mathcal{S}_k^T y_k} - \frac{c_ic_j}{\mathcal{S}_k^T\mathcal{B}_k\mathcal{S}_k} = 0 \quad \text{…. (2.8)}$$

$$\alpha_{ij}\ \frac{y_iy_j}{\mathcal{S}_k^T y_k} = \frac{c_ic_j}{\mathcal{S}_k^T\mathcal{B}_k\mathcal{S}_k}$$

$$\alpha_{ij} = \frac{\mathcal{S}_k^T y_k}{\mathcal{S}_k^T\mathcal{B}_k\mathcal{S}_k}\ \frac{c_ic_j}{y_iy_j} \quad \text{….(2.9)}$$

Now, since

$$\acute{y}_i\acute{y}_j = \alpha_{ij}\ y_iy_j \quad \text{…. (2.10)}$$

We substitute (2.9) in (2.10), we get

$$\acute{y}_i\acute{y}_j = \frac{\mathcal{S}_k^T y_k}{\mathcal{S}_k^T\mathcal{B}_k\mathcal{S}_k}\ \frac{c_ic_j}{y_iy_j}\ \ y_iy_j$$

$$\acute{y}_i\acute{y}_j = \frac{\mathcal{S}_k^T y_k}{\mathcal{S}_k^T\mathcal{B}_k\mathcal{S}_k}\ c_ic_j \quad \text{……(2.11)}$$

Therefore $\mathcal{B}_{k+1}$ in (2.6) is the formal of $Diagonal - BFGS$ where $\mathcal{B}_{k+1}$ always be a diagonal matrix symmetric and positive definite such that

$$\mathcal{B}_{k+1} = \begin{pmatrix} \mathfrak{b}_1 & .. & 0 & .. & 0 \\ : & & & & : \\ 0 & .. & \mathfrak{b}_i & .. & 0 \\ : & & & & : \\ 0 & .. & 0 & .. & \mathfrak{b}_n \end{pmatrix} \quad \ldots. (2.12)$$

Now, we will give two theories about symmetric and positive definite of $Diagonal - BFGS$

Theorem $(\mathbf{2.1})$**:** The Hessian matrix $\mathcal{B}_{k+1}$ produced by $Diagonal - BFGS$ update $\mathcal{B}_{k+1} = \mathcal{B}_k + \frac{\acute{y}_k \acute{y}_k^T}{\mathcal{S}_k^T y_k} - \frac{\mathcal{B}_k \mathcal{S}_k \mathcal{S}_k^T \mathcal{B}_k}{\mathcal{S}_k^T \mathcal{B}_k \mathcal{S}_k}$ is symmetric.

Proof:

$\mathcal{B}_{k+1} = \mathcal{B}_k + \frac{\acute{y}_k \acute{y}_k^T}{\mathcal{S}_k^T y_k} - \frac{\mathcal{B}_k \mathcal{S}_k \mathcal{S}_k^T \mathcal{B}_k}{\mathcal{S}_k^T \mathcal{B}_k \mathcal{S}_k}$, so $\mathcal{B}_{k+1}^T = \mathcal{B}_k^T + \frac{\acute{y}_k^{T^T} \acute{y}_k^T}{\mathcal{S}_k^T y_k} - \frac{\mathcal{B}_k^T \mathcal{S}_k^{T^T} \mathcal{S}_k^T \mathcal{B}_k^T}{\mathcal{S}_k^T \mathcal{B}_k \mathcal{S}_k}$

Since $\mathcal{B}_k^T = \mathcal{B}_k$, $\acute{y}_k^{T^T} = \acute{y}_k$ and $\mathcal{S}_k^{T^T} = \mathcal{S}_k$, we have

$\mathcal{B}_{k+1}^T = \mathcal{B}_k + \frac{\acute{y}_k \acute{y}_k^T}{\mathcal{S}_k^T y_k} - \frac{\mathcal{B}_k \mathcal{S}_k \mathcal{S}_k^T \mathcal{B}_k}{\mathcal{S}_k^T \mathcal{B}_k \mathcal{S}_k}$

Therefor $\mathcal{B}_{k+1} = \mathcal{B}_{k+1}^T$. Hence $\mathcal{B}_{k+1}$ is symmetric

Theorem $(\mathbf{2.2})$**:** The Hessian matrix $\mathcal{B}_{k+1}$ produced by $Diagonal - BFGS$ update is positive definite if and only if $\mathcal{S}_k^T y_k > 0$.

Proof:

Suppose that $\mathcal{S}_k^T y_k > 0$ to prove $\mathcal{B}_{k+1}$ is positive definite

$\mathcal{B}_{k+1} = \mathcal{B}_k + \frac{\acute{y}_k \acute{y}_k^T}{\mathcal{S}_k^T y_k} - \frac{\mathcal{B}_k \mathcal{S}_k \mathcal{S}_k^T \mathcal{B}_k}{\mathcal{S}_k^T \mathcal{B}_k \mathcal{S}_k}$

And let z $\neq 0$, then $z^T \mathcal{B}_{k+1} z = z^T \mathcal{B}_k z + \frac{z^T \acute{y}_k \acute{y}_k^T z}{\mathcal{S}_k^T y_k} - \frac{z^T \mathcal{B}_k \mathcal{S}_k \mathcal{S}_k^T \mathcal{B}_k z}{\mathcal{S}_k^T \mathcal{B}_k \mathcal{S}_k}$

Since $\frac{z^T \acute{y}_k \acute{y}_k^T z}{\mathcal{S}_k^T y_k} = \frac{\left(\acute{y}_k^T z\right)^T \acute{y}_k^T z}{\mathcal{S}_k^T y_k} = \frac{\|\acute{y}_k^T z\|^2}{\mathcal{S}_k^T y_k} \geq 0$, because $\mathcal{S}_k^T y_k > 0$,

And so, we get $z^T \mathcal{B}_{k+1} z = z^T \mathcal{B}_k z - \frac{z^T \mathcal{B}_k \mathcal{S}_k \mathcal{S}_k^T \mathcal{B}_k z}{\mathcal{S}_k^T \mathcal{B}_k \mathcal{S}_k} + \frac{\|\acute{y}_k^T z\|^2}{\mathcal{S}_k^T y_k}$

$z^T \mathcal{B}_{k+1} z = z^T \left(\mathcal{B}_k - \frac{\mathcal{B}_k \mathcal{S}_k \mathcal{S}_k^T \mathcal{B}_k}{\mathcal{S}_k^T \mathcal{B}_k \mathcal{S}_k} \right) z + \frac{\|\acute{y}_k^T z\|^2}{\mathcal{S}_k^T y_k}$

Since $\mathcal{B}_k \in S_{++}^n$, then there exist $\mathcal{L}_k \in \mathbb{R}^{n\times n}$ be Lower triangle matrix such that $\mathcal{B}_k = \mathcal{L}_k {\mathcal{L}_k}^T$

Now, we get $z^T\mathcal{B}_{k+1}z = z^T\mathcal{L}_k{\mathcal{L}_k}^T z - \frac{z^T\mathcal{L}_k{\mathcal{L}_k}^T\mathcal{S}_k\mathcal{S}_k^T\mathcal{L}_k{\mathcal{L}_k}^T z}{\mathcal{S}_k^T\mathcal{L}_k{\mathcal{L}_k}^T\mathcal{S}_k} + \frac{\|\acute{y}_k^T z\|^2}{\mathcal{S}_k^T y_k}$

Let $\mathfrak{a}_k = {\mathcal{L}_k}^T z \implies {\mathfrak{a}_k}^T = z^T\mathcal{L}_k$ and $\mathfrak{b}_k = {\mathcal{L}_k}^T\mathcal{S}_k \implies {\mathfrak{b}_k}^T = \mathcal{S}_k^T\mathcal{L}_k$

Therefore, we have $z^T\mathcal{B}_{k+1}z = {\mathfrak{a}_k}^T\mathfrak{a}_k - \frac{{\mathfrak{a}_k}^T\mathfrak{b}_k{\mathfrak{b}_k}^T\mathfrak{a}_k}{{\mathfrak{b}_k}^T\mathfrak{b}_k} + \frac{\|\acute{y}_k^T z\|^2}{\mathcal{S}_k^T y_k}$

$$= \frac{({\mathfrak{a}_k}^T\mathfrak{a}_k)\,({\mathfrak{b}_k}^T\mathfrak{b}_k) - ({\mathfrak{a}_k}^T\mathfrak{b}_k)^2}{{\mathfrak{b}_k}^T\mathfrak{b}_k} + \frac{\|\acute{y}_k^T z\|^2}{\mathcal{S}_k^T y_k}$$

By Cauchy-Schwartz inequality, we have $({\mathfrak{a}_k}^T\mathfrak{b}_k)^2 \le ({\mathfrak{a}_k}^T\mathfrak{a}_k)({\mathfrak{b}_k}^T\mathfrak{b}_k)$

so $\frac{({\mathfrak{a}_k}^T\mathfrak{a}_k)\,({\mathfrak{b}_k}^T\mathfrak{b}_k) - ({\mathfrak{a}_k}^T\mathfrak{b}_k)^2}{{\mathfrak{b}_k}^T\mathfrak{b}_k} > 0$,then $z^T\left(\mathcal{B}_k - \frac{\mathcal{B}_k\mathcal{S}_k\mathcal{S}_k^T\mathcal{B}_k}{\mathcal{S}_k^T\mathcal{B}_k\mathcal{S}_k}\right)z > 0$

Therefore $z^T\mathcal{B}_{k+1}z > 0$. Hence $\mathcal{B}_{k+1}$ is positive definite

Now, suppose that $\mathcal{B}_{k+1}$ is positive definite to prove $\mathcal{S}_k^T y_k > 0$

Since $\mathcal{B}_{k+1}\mathcal{S}_k = y_k$, so $\mathcal{S}_k^T\mathcal{B}_{k+1}\mathcal{S}_k = \mathcal{S}_k^T y_k$

Since $\mathcal{B}_{k+1}$ is positive definite, $then\ \mathcal{S}_k^T\mathcal{B}_{k+1}\mathcal{S}_k > 0$

Therefore, $\mathcal{S}_k^T y_k > 0$

3. $Diagonal - BFGS\ Algorithm$

Let x_0 be an initial point be given as well as an initial $n \times n$ diagonal matrix $\mathcal{B}_0$.

Step1: Set $k = 0$ and compute $\mathcal{g}_k = \nabla f(x_k)$

Step2: Find $\mathcal{P}_k$ by using $\mathcal{P}_k = -\mathcal{B}_k^{-1}\mathcal{g}_k$

Step3: Set $x_{k+1} = x_k + \lambda_k\mathcal{P}_k$, and compute λ_k by exact or inexact line search such that $f(x_1) \le f(x_0)$

Step4: Find a vector $\mathcal{S}_K$ where $\mathcal{S}_K = x_{k+1} - x_k$

Step5: Compute $\mathcal{g}_{k+1} = \nabla f(x_{k+1})$

Step6: Find a vector y_K where $y_K = \mathcal{g}_{k+1} - \mathcal{g}_k$

Step7: Update the matrix $\mathcal{B}_k$ to obtain $\mathcal{B}_{k+1}$ by using (2.6)

Step8: If $\|\mathcal{g}_k\| < \epsilon$ stop. Other wise $k = k + 1$, go to step 1

Example$(\mathbf{3.1})$**:**

Let $min\, f = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_3 - 2x_2x_3$, form the starting point $x_0 = \begin{pmatrix} 0.01 \\ 0 \\ 0.02 \end{pmatrix}$ and using $\mathcal{B}_0 = I$, $\epsilon = 0.0004$

Solution:

Iteration **1** ($\boldsymbol{k} = \mathbf{0}$)

The gradient of $f(x)$ is $\nabla f = \begin{pmatrix} 4x_1 - 2x_3 \\ 4x_2 - 2x_3 \\ 4x_3 - 2x_1 - 2x_2 \end{pmatrix}$

$g_0 = \nabla f(x_0) = \begin{pmatrix} 0 \\ -0.04 \\ 0.06 \end{pmatrix}$

The initial search direction is $\mathcal{P}_0 = -\mathcal{B}_0^{-1} g_0 = \begin{pmatrix} 0 \\ 0.04 \\ -0.06 \end{pmatrix}$

Now, $x_1 = x_0 + \lambda_0 \mathcal{P}_0$,then $x_1 = \begin{pmatrix} 0.01 \\ 0.04\lambda_0 \\ 0.02 - 0.06\lambda_0 \end{pmatrix}$

Now, we will find $\lambda_0 \ni f(x_1) \leq f(x_0)$

$f(x_1) = f\begin{pmatrix} 0.01 \\ 0.04\lambda_0 \\ 0.02 - 0.06\lambda_0 \end{pmatrix} = 0.0006 - 0.0052\lambda_0 + 0.0152{\lambda_0}^2$

With respect to λ_0 . Since $\frac{\mathrm{d}f}{\mathrm{d}\lambda_0} = 0$ at $\lambda_0 = 0.1710$, we get

$x_1 = \begin{pmatrix} 0.01 \\ 0.0068 \\ 0.0097 \end{pmatrix}$ and $g_1 = \nabla f(x_1) = \begin{pmatrix} 0.0206 \\ 0.0078 \\ 0.0052 \end{pmatrix}$

Now, $\mathcal{S}_0 = x_1 - x_0$ then $\mathcal{S}_0 = \begin{pmatrix} 0 \\ 0.0068 \\ -0.0103 \end{pmatrix}$

$y_0 = g_1 - g_0$, then $y_0 = \begin{pmatrix} 0.0206 \\ 0.0478 \\ -0.0548 \end{pmatrix}$

Update the matrix $\mathcal{B}_1 = \mathcal{B}_0 + \frac{\acute{y}_0 \acute{y}_0^T}{\mathcal{S}_0^T y_0} - \frac{\mathcal{B}_0 \mathcal{S}_0 \mathcal{S}_0^T \mathcal{B}_0}{\mathcal{S}_0^T \mathcal{B}_0 \mathcal{S}_0}$

$\mathcal{S}_0^T y_0 = 0.0008$

$\mathcal{B}_0 \mathcal{S}_0 \mathcal{S}_0^T \mathcal{B}_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0.00004 & -0.00007 \\ 0 & -0.00007 & 0.0001 \end{pmatrix}$

$\mathcal{S}_0^T \mathcal{B}_0 \mathcal{S}_0 = 0.0001$

Now, to find $\acute{y}_0 \acute{y}_0^T$ such that $\acute{y}_0 \acute{y}_0^T = \begin{pmatrix} y_1^2 & \acute{y}_1\acute{y}_2 & \acute{y}_1\acute{y}_3 \\ \acute{y}_1\acute{y}_2 & y_2^2 & \acute{y}_2\acute{y}_3 \\ \acute{y}_1\acute{y}_3 & \acute{y}_2\acute{y}_3 & y_3^2 \end{pmatrix}$

$\acute{y}_1\acute{y}_2 = \frac{\mathcal{S}_0^T y_0}{\mathcal{S}_0^T \mathcal{B}_0 \mathcal{S}_0}\ s_1 s_2 = \frac{0.0008}{0.0001}(\ 0\ ) = 0$

$\acute{y}_1\acute{y}_3 = \frac{\mathcal{S}_0^T y_0}{\mathcal{S}_0^T \mathcal{B}_0 \mathcal{S}_0}\ s_1 s_3 = \frac{0.0008}{0.0001}(\ 0\ ) = 0$

$\acute{y}_2\acute{y}_3 = \frac{\mathcal{S}_0^T y_0}{\mathcal{S}_0^T \mathcal{B}_0 \mathcal{S}_0}\ s_2 s_3 = \frac{0.0008}{0.0001}(-0.00007\ ) = -0.0006$

$y_1^2 = 0.0004\ ,\ y_2^2 = 0.0022$ and $y_3^2 = 0.0030$

Hence $\acute{y}_0 \acute{y}_0^T = \begin{pmatrix} 0.0004 & 0 & 0 \\ 0 & 0.0022 & -0.0006 \\ 0 & -0.0006 & 0.0030 \end{pmatrix}$

Therefore, $\mathcal{B}_1 = \begin{pmatrix} 1.5 & 0 & 0 \\ 0 & 3.3 & 0 \\ 0 & 0 & 3.7 \end{pmatrix}$ and $\|g_1\| = 0.0005 > \epsilon$

Iteration 2 $(\boldsymbol{k = 1})$

The initial search direction is $\mathcal{P}_1 = -\mathcal{B}_1^{-1} g_1 = \begin{pmatrix} -0.0137 \\ -0.0023 \\ -0.0014 \end{pmatrix}$

Now, $x_2 = x_1 + \lambda_1 \mathcal{P}_1$, then $x_2 = \begin{pmatrix} 0.01 - 0.0137\lambda_1 \\ 0.0068 - 0.0023\lambda_1 \\ 0.0097 - 0.0014\lambda_1 \end{pmatrix}$

Now, we will find $\lambda_1\ \ni\ f\ (x_2) \leq f\ (x_1)$

$f(x_2) = f\begin{pmatrix} 0.01 - 0.0137\lambda_1 \\ 0.0068 - 0.0023\lambda_1 \\ 0.0097 - 0.0014\lambda_1 \end{pmatrix}$

$= 2(0.01 - 0.0137\lambda_1)^2 + 2(0.0068 - 0.0023\lambda_1)^2 + 2(0.0097 - 0.0014\lambda_1)^2 \quad - \quad 0.00032592 + 0.00035744\lambda_1 - 0.0000448{\lambda_1}^2$

With respect to λ_1 . Since $\frac{df}{d\lambda_1} = 0$ at $\lambda_1 = 0.4454$, we get

$x_2 = \begin{pmatrix} 0.0038 \\ 0.0057 \\ 0.0090 \end{pmatrix}$ and $g_2 = \nabla f(x_2) = \begin{pmatrix} -0.0028 \\ 0.0048 \\ 0.017 \end{pmatrix}$

Now, $\mathcal{S}_1 = x_2 - x_1$ then $\mathcal{S}_1 = \begin{pmatrix} -0.0062 \\ -0.0011 \\ -0.0007 \end{pmatrix}$

$y_1 = g_2 - g_1$ then $y_1 = \begin{pmatrix} -0.0234 \\ -0.003 \\ 0.0118 \end{pmatrix}$

Update the matrix $\mathcal{B}_2 = \mathcal{B}_1 + \frac{\acute{y}_1 \acute{y}_1^T}{\mathcal{S}_1^T y_1} - \frac{\mathcal{B}_1 \mathcal{S}_1 \mathcal{S}_1^T \mathcal{B}_1}{\mathcal{S}_1^T \mathcal{B}_1 \mathcal{S}_1}$

$\mathcal{S}_1^T y_1 = 0.0001$

$\mathcal{B}_1 \mathcal{S}_1 \mathcal{S}_1^T \mathcal{B}_1 = \begin{pmatrix} 0.00007 & 0.00003 & 0.00002 \\ 0.00003 & 0.00001 & 0.000007 \\ 0.00002 & 0.000007 & 0.000004 \end{pmatrix}$

$\mathcal{S}_1^T \mathcal{B}_1 \mathcal{S}_1 = 0.00005$

Now, to find $\acute{y}_1 \acute{y}_1^T$ such that $\acute{y}_1 \acute{y}_1^T = \begin{pmatrix} y_1^2 & \acute{y}_1 \acute{y}_2 & \acute{y}_1 \acute{y}_3 \\ \acute{y}_1 \acute{y}_2 & y_2^2 & \acute{y}_2 \acute{y}_3 \\ \acute{y}_1 \acute{y}_3 & \acute{y}_2 \acute{y}_3 & y_3^2 \end{pmatrix}$

$\acute{y}_1 \acute{y}_2 = \frac{\mathcal{S}_1^T y_1}{\mathcal{S}_1^T \mathcal{B}_1 \mathcal{S}_1} s_1 s_2 = \frac{0.0001}{0.00005} (0.00003) = 0.00006$

$\acute{y}_1 \acute{y}_3 = \frac{\mathcal{S}_1^T y_1}{\mathcal{S}_1^T \mathcal{B}_1 \mathcal{S}_1} s_1 s_3 = \frac{0.0001}{0.00005} (0.00002) = 0.00004$

$\acute{y}_2 \acute{y}_3 = \frac{\mathcal{S}_1^T y_1}{\mathcal{S}_1^T \mathcal{B}_1 \mathcal{S}_1} s_2 s_3 = \frac{0.0001}{0.00005} (0.00002) = 0.000014$

$y_1^2 = 0.0005$, $y_2^2 = 0.000009$ and $y_3^2 = 0.0001$

Hence $\acute{y}_1 \acute{y}_1^T = \begin{pmatrix} 0.0005 & 0.00006 & 0.00004 \\ 0.00006 & 0.000009 & 0.000014 \\ 0.00004 & 0.000014 & 0.0001 \end{pmatrix}$

Therefore, $\mathcal{B}_2 = \begin{pmatrix} 5.1 & 0 & 0 \\ 0 & 3.2 & 0 \\ 0 & 0 & 4.62 \end{pmatrix}$ and $\|g_2\| = 0.0003 < \epsilon$ stop.

4. Numerical Results

This part devoted to numerical experiments .We present the performance of the $Diagonal - BFGS$ algorithm from the performance of the $BFGS$ algorithm, using starting points and convergence criteria and limits. The computer programs are written in MATLAB version (5.3). The results are presented in Table (4-1).

The test functions are chosen as follows:

$1-$ A quartic function . [9]

$f(x) = \sum_{i=1}^{4} \left(10^{i-1} x_i^4 + x_i^3 + 10^{1-i} x_i^2\right)$

$2-$ Rosen rock's function [7]

$f(x) = \sum_{i=1}^{n/2} [100(x_i - x_i^3)^2 + (1 - x_i)^2]$.

$3-$ Watson function [7]

$F(x) = \sum_{i=1}^{j} f_i^2(x)$

$$f_i(x) = \sum_{j=2}^{3}(j-1)x_j t_j^{j-2} - \left(\sum_{j=1}^{3} x_j t_j^{j-1}\right)^2 - 1.$$

And $t_j = \frac{i}{29}$.

4 – Rosenbroc'k cliff function [8]

$$f(x) = 10^{-4}(x_1 - 3)^2 - (x_1 - x_2) + e^{20(x_1 - x_2)}$$

5 – Trigonometric function [1]

$$f(x) = \sum_{i=1}^{n}\left[n - \sum_{j=1}^{n}\cos x_j + i(1 - \cos x_i) - \sin x_i + e^{xi} - 1\right]^2$$

6 – Generalized Edger function . [1]

$$f(x) = \sum_{i=1}^{n/2}[(x_{2i-1} - 2)^4 + (x_{2i-1} - 2)^2 x_{2i}^2 + (x_{2i} + 1)^2]$$

The table (4-1) gives the Diagonal – BFGS for our selected test functions. At the first, we give the number of iterations then Feval of each function from deferent starting point. Then after that, we give all the results of this analysis

The table (4-2) gives the Hessian matrix.

Table (4-1) : Numerical Results for $Diagonal - BFGS$ update

Fun.	Starting point	Dim.	Feval	Iter.
1	$[-1; 0; 0; 0]^T$	4	-0.1055	3
2	$[0; 1; 1; 1; 1; 1; 1; 1]^T$	8	3.5935e-011	3
3	$[0; 0; 0; 0; 0; 0; 0; 0; 0; 0]^T$	10	5.3614e-010	2
4	$[0.5; 0.5; 0.5; 0.5; 0.5; 0.5;$ $0.5; 0.5; 0.5; 0.5; 0.5; 0.5]^T$	12	0.2004	3
5	$[2; 2; 2; 2; 2; 2; 2; 2; 2; 2; 2; 2]^T$	12	1.5555e-004	4
6	$[-1; -1; -1; -1; -1; -1$ $; -1; -1; -1; -1; -1; -1;$ $-1; -1; -1; -1; -1; -1]^T$	18	4.6834e-007	10
6	$[0; 5; 0; 5; 0; 5; 0; 5; 0;$ $5; 0; 5; 0; 5; 0; 5; 0; 5]^T$	18	2.9758e-007	5

Table (4-2): Numerical results for Hessian matrix for $Diagonal - BFGS$

Fun.	Starting point	Hessian matrix
1	$[-1; 0; 0; 0]^T$	2.3283 0 0 0 0 1.0000 0 0 0 0 1.0000 0 0 0 0 1.0003
2	$[0; 1; 1; 1;^T$ $1; 1; 1; 1]$	1.0e+004 * 1.7168 0 0 0 0 0 0 0 0 0.1520 0 0 0 0 0 0 0 0 0.1520 0 0 0 0 0 0 0 0 0.1520 0 0 0 0 0 0 0 0 0.1520 0 0 0 0 0 0 0 0 0.1520 0 0 0 0 0 0 0 0 0.1520 0 0 0 0 0 0 0 0 0.1520
3	$[0; 0; 0; 0; 0$ $; 0; 0; 0; 0; 0]$	1.0000 0 0 0 0 0 0 0 0 0 0 53.7883 0 0 0 0 0 0 0 0 0 0 3.2597 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 1.0000
4	$[0.5; 0.5; 0.5$ $; 0.5; 0.5; 0.5;$ $0.5; 0.5; 0.5^T$ $; 0.5; 0.5; 0.5]$	81.7495 0 0 0 0 0 0 0 0 0 0 0 0 83.5242 0 0 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 0 0 1.0000 0 0 0 0 0 0 0 0 0 0 0 0 1.0000

5	[2; 2; 2; 2; 2; 2; 2; 2; 2; 2]	$\begin{matrix} 37.4068 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 43.4020 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 49.5533 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 55.2181 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 59.9907 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 63.7766 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 69.2009 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 81.5635 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 99.3627 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 115.8179 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 127.4266 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 79.4963 \end{matrix}$
6	$[-1;-1;-1;-1;-1;-1;-1;-1;-1;-1;-1;-1;-1;-1;-1;-1;-1;-1;-1;-1]^T$	$\begin{matrix} 4.8391 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.6453 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4.8391 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.6453 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4.8391 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.6453 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4.8391 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.6453 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.8391 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.6453 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.8391 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.6453 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.8391 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.6453 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.8391 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.6453 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.8391 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.6453 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.8391 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.6453 \end{matrix}$
6	$[0;5;0;5;0;5;0;5;0;5;0;5;0;5;0;5;0;5;0;5]^T$	$\begin{matrix} 6.8088 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.7214 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6.8090 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.7214 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6.8090 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.7202 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 6.8088 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.7214 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6.8090 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.7214 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6.8090 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.7202 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6.8088 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.7206 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6.8090 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.7206 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6.8090 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.7206 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6.8090 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.7206 \end{matrix}$

5. Conclusion

In this work, we proposed new techniques to solve the unconstrained optimization problem with large dimension where the Hessian matrix becomes sparse matrix. The BFGS update is modified to a diagonal update, so called the diagonal BFGS update. Moreover that it is symmetric and positive definite are given.

Finally, We provided the numerical results show that the proposed is efficient for unconstrained optimization problem of diagonal BFGS update on function is selected, which suggests that a good improvement has been achieved.

6. References

[1] Al-Bayati, A.,(1991)," A new family of self-scaling variable metric Algorithms for unconstrained optimization ", Journal of Educ. and Sci., Iraq, Vol.12, pp.25-54.

[2] Byrd R., Nocedal J. and Yuan Y., (2000) , " Global convergence of a class 0f quasi-Newton methods on convex problem ", SIAM, Vol.12 , Pages 1095-1115.

[3] Mahmood, S. S. (2020). Modified BFGS Update (H-Version) Based on the Determinant Property of Inverse of Hessian Matrix for Unconstrained Optimization. Baghdad Science Journal, 17(3 (Suppl.)), 0994-0994.

[4] Mahmood, S. S., & Muhanah, N. S. (2019). Symmetric and Positive Definite Broyden Update for Unconstrained Optimization. Baghdad Science Journal, 16(3).

[5] Mahmood, S. S., Ibrahim Mansour, A., & Ali Hassan, H. (2012). The Diagonal Update for Unconstrained Optimization. JOURNAL OF EDUCATION AND SCIENCE, 25(3), 68-73.

[6] Mahmood, S. S., Mansour, A. I., & Abdlrazak, B. T. (2011). α BFGS UPDATE FOR UNCONSTRAINED OPTIMIZATION. *Journal of college of Education*, (1).

[7] Oren, S. S, (1973) , " Self-scaling Variable Metric Algorithms without line search for Unconstrained minimization ".Mathematics of computation , 27:873-885.

[8] Todd M. J.,(1984) ,"Quasi-Newton update in abstract spaces ", SIAM Review, 26 367-377 .

[9] Yuan Y., (1991),"A Modified BFGS Algorithm for Unconstrained Optimization", Computing Center, Academia Sinica, Beijing, China .