On Modified DFP Update for Unconstrained Optimization

Saad Shakir Mahmood, Samira Hassan Shnywer

Department of Mathematics, College of Education, Almustansiryah University, Baghdad, Iraq

Email address: saadshakirmahmood@yahoo.com (S. S. Mahmood), samira.hassan75@yahoo.com (S. H. Shnywer)

To cite this article: Saad Shakir Mahmood, Samira Hassan Shnywer. On Modified DFP Update for Unconstrained Optimization. *American Journal of Applied Mathematics*. Vol. 5, No. 1, 2017, pp. 19-30. doi: 10.11648/j.ajam.20170501.13

Received: December 25, 2016; Accepted: January 9, 2017; Published: February 6, 2017

Abstract: In this paper, we propose a new modify of DFP update with a new extended quasi-Newton condition for unconstrained optimization problem so called \((\alpha - DFP)\) update. This update is based on a new Zhang Xu condition we show that \((\alpha - DFP)\) update preserves the value of determinant of the next Hessian matrix equal to the value of determinant of current Hessian matrix theoretically and practically. Global convergence of the modify is established. Local and super linearly convergence are obtained for the proposed method. Numerical results are given to compare a performance of the modify \((\alpha - DFP)\) method with the standard DFP method on same function is selected.

Keywords: Quasi-Newton Equation, the DFP Updating Formula, Global Convergence and Super Linearly Convergence

1. Introduction

The quasi-Newton methods are very useful and efficient methods for solving the unconstrained minimization problem

\[
\min_{x \in \mathbb{R}^n} f(x)
\]

(1)

Where \(f: \mathbb{R}^n \to \mathbb{R}\) is twice continuously differentiable. Starting from point \(x_0\) and a symmetric positive definite matrix \(B_0\), quasi-Newton method generates sequence \(\{x_k\}\) and \(\{B_k\}\) by the iteration of the form

\[
x_{k+1} = x_k - B_k^{-1}\nabla f(x_k), k = 0, 1, \ldots
\]

(2)

where the update \(B_k \in \mathbb{R}^{n \times n}\) satisfies the following famous quasi-Newton equation or (secant equation):

\[
B_{k+1}s_k = y_k
\]

(3)

with

\[
s_k = \lambda_k d_k, y_k = g_{k+1} - g_k
\]

(4)

where \(\lambda_k\) is the step length and \(d_k\) is the search direction that is obtained by solving the equation:

\[
d_k = -B_k^{-1}\nabla f(x_k)
\]

(5)

in which \(g_k = \nabla f(x_k)\) is the gradient of \(f(x)\) at \(x_k\) and \(B_k\) is an approximation to the Hessian matrix \(G_k = \nabla^2 f(x_k)\). The updating matrix \(B_k\) is required to satisfy the usual quasi-Newton equation (3) with equation (4). So \(B_{k+1}\) is reasonable approximation to \(G_{k+1}\).

The \((\alpha - DFP)\) update consist of iteration of the form (2) where \(d_k\) is the search direction which of the form (5) and the Hessian approximation \(B\) is update by the \((-DFP)\) formula with quasi-Newton equation (3).

\[
B_{k+1} = B - \frac{y_k y_k^T + \alpha y_k^T B_k y_k}{y_k^T y_k} \frac{1}{y_k^T B_k y_k}
\]

(6)

where \(\alpha = \frac{y_k^T B_k y_k}{y_k^T y_k}\)

(7)

The formula is the modifying of DFP update which is satisfy equation (3) and in the next section.

In the following discussion, we shall use \(\|\|\) and \(\|\|_F\) to denote the \(\ell_2\)-norm and the Frobenius norm, respectively. For a symmetric positive definite matrix \(M \in \mathbb{R}^{n \times n}\), we show also use the following weighted norm

\[
\|X\|_M = \|M X M\|_F, \forall X \in \mathbb{R}^{n \times n}
\]

(8)

2. \(\alpha - DFP\) Update

Let \(\delta = \alpha - 1\) and from Zhang Xu condition, we have:

\[
s^* = s + (\alpha - 1)s,
\]

which gives \(s^* = \alpha s\)

(9)

From quasi-Newton equation (3), we get:
\[ B_{k+1}s_k^* = y_k \] which is extended of quasi-Newton equation (10)

Now, substitution (9) in (10), we obtain:

\[ B_{k+1}a s_k = y_k \]  

(11)

Let we consider the determinant of Hessian matrix for the DFP update [7] with replace each \( s_k \) by \( s_k^* \), we get:

\[ \text{Det} (B_{k+1}) = \frac{y_k^T [\text{Det}(B_k)]}{s_k^T B_k s_k} \]  

(12)

And from equation (9), we obtain:

\[ \text{Det} (B_{k+1}) = \frac{y_k^T a s_k [\text{Det}(B_k)]}{s_k^T B_k s_k} \]  

(13)

Which gives:

\[ \text{Det} (B_{k+1}) = \frac{y_k^T s_k [\text{Det}(B_k)]}{s_k^T B_k s_k} \]  

(14)

we suppose that \( \text{Det} (B_{k+1}) = \frac{1}{\text{Det}(B_k)} \).

Then the equation (14) becomes:

\[ \frac{1}{\text{Det}(B_k)} = \frac{y_k^T s_k [\text{Det}(B_k)]}{s_k^T B_k s_k} \]

Now, by multiplying both sides with \( \frac{1}{\text{Det}(B_k)} \), we get:

\[ \frac{1}{\text{Det}(B_k)} = \frac{y_k^T s_k}{s_k^T B_k s_k} \]  

(15)

Hence

\[ \alpha(s_k^T B_k s_k) = \frac{y_k^T s_k}{s_k^T B_k s_k} [\text{Det}(B_k)]^2. \]

Finally:

\[ \alpha = \frac{y_k^T s_k [\text{Det}(B_k)]^2}{s_k^T B_k s_k} \]  

(16)

In addition, by more simplifying we can write \( B_{k+1}(\alpha-DFP) \) as follows:

\[ B_{k+1}(\alpha-DFP) = B_k - \frac{y_k s_k^T [y_k - b_k s_k] y_k}{y_k^T s_k} \]  

\[ + \frac{y_k s_k^T [y_k - b_k s_k] y_k}{y_k^T s_k} \left( 1 + \frac{s_k^T B_k s_k}{y_k^T s_k} \right) \]  

(17)

which is equivalent the following formula:

\[ B_{k+1}(\alpha-DFP) = B_k + \frac{y_k s_k^T [y_k - b_k s_k] y_k}{y_k^T s_k} \]  

\[ - \frac{(y_k - b_k s_k)^T s_k y_k}{y_k^T s_k} \]  

(18)

and also the formula (17) is equivalent the following formula:

\[ B_{k+1}(\alpha-DFP) = \left( I - \frac{s_k^T s_k^*}{y_k^T s_k} \right) B_k \left( I - \frac{s_k^T s_k^*}{y_k^T s_k} \right) + \frac{y_k s_k^T}{y_k^T s_k} \]  

\[ \times \left( I - \frac{s_k^T s_k^*}{y_k^T s_k} \right) \]  

(19)

where \( s_k^* \) is defined by (9). It is clear that any formula is symmetric and satisfies the quasi-Newton equation.

### 3. Convergence Analysis

Now, we study the global convergence of the \( \alpha-DFP \) update.

At the first, we need the following assumptions:

**Assumption (3.1)**

\( (A) : f: R^n \rightarrow R \) is twice continuously differentiable on convex set \( D \subset R^n \)

\( (B) : f(x) \) is uniformly convex, i.e., there exist positive constants \( n \) and \( N \) such that for all \( x \in L(x) = \{ x | f(x) \leq f(x_0) \} \), which is convex, where \( x_0 \) is starting point, we have:

\[ n \|v\|^2 \leq \nabla^2 f(x)v \leq N \|v\|^2, \forall v \in R^n \]  

(20)

The assumption (B) implies that \( \nabla^2 f(x) \) is positive definite on \( (x) \), and that \( f \) has a unique minimizer \( x^* \) in \( L(x) \).

By definition of weighted norm (8) and equation (9), satisfy extended (Q-N) equation then \( W y_k = s_k^* \) and \( y_k = W^{-1} \alpha s_k \) that is then

\[ y_k = \tilde{G}_k a s_k \]  

(21)

where \( \tilde{G}_k = W^{-1} \).

Now by return to the property (20) and from definition of weighted norm, we get:

\[ n \leq \frac{y_k^T a s_k}{\|a s_k\|^2} = \frac{y_k^T \tilde{G}_k a s_k}{\|a s_k\|^2} \leq N \]  

(22)

Where \( \tilde{G}_k \) is the average Hessian, which is defined as:

\[ \tilde{G}_k = \int_0^1 \nabla^2 f(x_k + \theta s_k) \, d\theta \]  

(23)

and

\[ \frac{1}{N} \leq \frac{\|a s_k\|^2}{y_k^T a s_k} \leq \frac{1}{n} \]  

(24)

Since also

\[ \frac{y_k^T a s_k}{\|a s_k\|^2} = \frac{y_k^T \tilde{G}_k a s_k}{\|a s_k\|^2} = \frac{\tilde{G}_k a s_k}{\|a s_k\|^2} \]  

(25)

Assumption (B) of (3.1) means that \( \tilde{G}_k \) is positive definite proven, thus its square root is well defined. There is a symmetric square root \( \tilde{G}_k^{-1/2} \) satisfying

\[ \tilde{G}_k^{-1/2} \tilde{G}_k^{-1/2} = \tilde{G}_k^{-1} \]

If we let \( w_k = \tilde{G}_k^{-1/2} a s_k \), then

\[ \frac{y_k^T a s_k}{\|a s_k\|^2} = \frac{\tilde{G}_k^{-1/2} a s_k}{\|a s_k\|^2} = \frac{\tilde{G}_k^{-1/2} a s_k}{\|a s_k\|^2} \]  

(26)

\[ \frac{y_k^T a s_k}{\|a s_k\|^2} = \frac{\tilde{G}_k^{-1/2} a s_k}{\|a s_k\|^2} \]  

(27)

Substitution equation (26) in equation (27), we get
\[
\frac{\|y_k\|^2}{\alpha_{yk}} = \frac{w_k^T \bar{G} w_k}{w_k^T w_k} \tag{28}
\]

And from (20) \(v^T \nabla^2 f(x^*) v \leq N\|v\|^2\), we know \(\|v\|^2 > 0\). Then we can divide both sides from this, we get:

\[
\frac{v_k^T \bar{G} w_k}{\|v_k\|^2} \leq N
\]

That mean

\[
\frac{w_k^T \bar{G} w_k}{w_k^T w_k} \leq N \tag{29}
\]

Then from equation (28), we get:

\[
\frac{\|y_k\|^2}{\alpha_{yk}} \leq N \tag{30}
\]

In addition, from equations (21) and (9), we get:

\[
\|y_k\| = \|\bar{G} s_k\| \leq \|\bar{G} k\|\|s_k\| \leq \|\bar{G} k\| \|y_k\| \tag{31}
\]

Which gives:

\[
\frac{\|y_k\|}{\|y_k\|} \leq N \tag{32}
\]

And

\[
\frac{\|s_k\|}{\|y_k\|} \leq \frac{1}{n} \tag{33}
\]

Therefore, from the above discussion, we have:

Lemma (3.2)

Let \(f: R^n \rightarrow R\) satisfy Assumption (3.1). Then

\[
\|s_k\| \|y_k\| \leq s_k y_k \|y_k\| \leq s_k y_k \|y_k\|^2 \leq \frac{\|s_k\|}{\|y_k\|} \|y_k\| \leq s_k y_k \|y_k\|^2 \leq \frac{\|s_k\|}{\|y_k\|} \|y_k\| \leq s_k y_k \|y_k\|^2
\]

Are bounded.

Hence, we have

\[
\|s_k\| \|y_k\| \leq s_k y_k \|y_k\| \leq s_k y_k \|y_k\|^2 \leq s_k y_k \|y_k\|^2 \leq s_k y_k \|y_k\|^2
\]

Are bounded.

Lemma (3.3)

Under exact line search, \(\|s_k\|\|^2\) and \(\|y_k\|^2\) are convergent.

Which gives:

\[
\|s_k\|^2\] is convergent and bounded.

From lemma (3.2), we have

\[
\|s_k\|^2\]

is convergent and bounded.

Since \(\|y_k\|^2\) is convergent, we get:

\[
\|s_k\|^2\]

is convergent, then

\[
\sum |s_k|^2 \leq \frac{1}{n} \sum |y_k|^2
\]

Which implies \(\sum |s_k|^2\) is convergent, where \(f(x^*)\) is the minimum of \(f(x)\).

Lemma (3.4)

For all vector \(x\), the inequality

\[
\|g(x)\|^2 \geq n \|f(x) - f(x^*)\|
\]

holds, where \(f(x^*)\) is the minimum of \(f(x)\).

Proof:

Since the function

\[
\phi(\tau) = f(x + \tau(x^* - x)), \quad (0 \leq \tau \leq 1)
\]

is a convex function, we have:

\[
f(x + \tau(x^* - x)) \geq f(x) + \tau(x^* - x)^T g(x)
\]

In particular, set \(\tau = 1\), then we have

\[
f(x + (x^* - x)) \geq f(x) + (x^* - x)^T g(x)
\]

\[
f(x^*) \geq f(x) + (x^* - x)^T g(x)
\]

Which gives

\[
f(x^*) - f(x) \geq (x^* - x)^T g(x)
\]

By multiplying both sides with (-1) we get

\[
f(x) - f(x^*) \leq -(x^* - x)^T g(x)
\]

By Cauchy-Schwarz inequality, we get

\[
f(x) - f(x^*) \leq (x^* - x)^T g(x) \leq \|g(x)\| \|x^* - x\| \tag{34}
\]

From (24) and (9), we have

\[
|\alpha| \|s_k\|^2 \leq \frac{v_k^T s_k}{n} \tag{35}
\]

\[
|\alpha| \|s_k\|^2 \leq \frac{(g(x^*) - g(x))^T s_k}{n} \tag{36}
\]

Hence

\[
|\alpha| \|x^* - x\|^2 \leq \frac{\|g(x)\|^2 \|x^* - x\|}{n}
\]

\[
|\alpha| \|x^* - x\|^2 \leq \frac{\|g(x)\|^2 \|x^* - x\|}{\alpha n}
\]

Let \(\alpha n = n\)

\[
\|x^* - x\|^2 \leq \|g(x)\| \|x^* - x\| / n
\]

\[
\|x^* - x\| \leq \|g(x)\| / n \tag{37}
\]

Substituting (37) into (35) establishes (34).

Theorem (3.5)

Suppose that \(f(x)\) satisfies Assumption (3.1). Then under exact line search the sequence \(\{x_k\}\) generated by \(\alpha - DFP\) method converges to the minimizer \(x^*\) of \(f\).
Proof: Consider $\alpha - DFP$ formula of inverse Hessian approximation
\[ H_{k+1} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{\alpha_k s_k s_k^T}{y_k^T y_k} \]  (38)
and from $\alpha - DFP$ formula (19) of Hessian approximation
\[ B_{k+1} = (I - \frac{y_k y_k^T}{y_k^T y_k}) B_k (I - \frac{s_k s_k^T}{y_k^T y_k}) \]  (39)

Obviously, $B_{k+1} H_{k+1} = I$. By computing the trace of (39), we have
\[ \text{Tr}(B_{k+1}) = \text{Tr}(B_k) - 2 \frac{\alpha_k s_k^T B_k y_k}{y_k^T y_k} + \frac{\alpha_k s_k s_k^T (y_k)^2}{(y_k^T y_k)^2} \]  (40)
The middle two terms can be written as:
\[ -\frac{\alpha_k s_k^T B_k s_k y_k}{y_k^T y_k} + \frac{\alpha_k s_k s_k^T y_k^2}{(y_k^T y_k)^2} \]

From equation (4) and (5), we have
\[ s_k^T y_k = s_k^T (g_k + s_k - g_k) = s_k^T s_k - s_k^T g_k \]

From the property of the $DFP$ method [5] $g_{k+1} = s_k^T g_k = 0$, we obtain:
\[ g_{k+1}^T s_k = \frac{2 \alpha_k g_k^T y_k + \alpha_k^2 s_k^T y_k^2 y_k}{(s_k^T y_k)^2} \]
\[ = \alpha_k \left[ \frac{2 g_k^T y_k + y_k^2 y_k}{s_k^T y_k} \right] \]

From equation (4) and (5) again, we get
\[ g_k^T y_k = \frac{2 g_k^T y_k + y_k^2 y_k}{g_k^T H_k g_k} \]
\[ g_{k+1}^T g_k = \frac{g_{k+1}^T g_k + g_k^T g_k}{g_k^T H_k g_k} = \frac{\|g_k + s_k\|^2 - \|g_k\|^2}{g_k^T H_k g_k} \]  (41)

From the positive definiteness property of $H_{k+1}$, then (38) becomes:
\[ g_{k+1}^T H_{k+1} g_{k+1} = g_k^T \left[ H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} \right] g_{k+1} \]
\[ = g_k^T \left[ H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} \right] g_{k+1} + \sum_{j=0}^{k} \frac{\|y_j\|^2}{\|s_j\|^2} + \frac{\|y_j\|^2}{\alpha_j^2} \]  (42)

By finding the inverse number of the expression, we obtain
\[ g_{k+1}^T H_{k+1} g_{k+1} = \frac{\|g_k + s_k\|^2 - \|g_k\|^2}{\|g_k + s_k\|^2} \]

Using (41) and (42), then (40) becomes:
\[ \text{Tr}(B_{k+1}) = \text{Tr}(B_k) - \frac{\|g_k + s_k\|^2}{\|g_k + s_k\|^2} - \frac{\|g_k\|^2}{\|g_k + s_k\|^2} \]  (43)

By recurrence, we obtain:
\[ \text{Tr}(B_{k+1}) = \text{Tr}(B_0) + \sum_{j=0}^{k} \frac{\|y_j\|^2}{\|s_j\|^2} + \sum_{j=0}^{k} \frac{\|y_j\|^2}{\alpha_j^2} \]  (44)

Therefore, by lemma (3.2), there exists a positive number $N$ which is independent of $k$, such that
\[ \text{Tr}(B_{k+1}) \leq \frac{\|g_k + s_k\|^2}{\|g_k + s_k\|^2} - \frac{\|g_k\|^2}{\|g_k + s_k\|^2} + N \]  (45)

In the left part, we will prove that if the theorem does not hold, then the sum of the last two term in (45) is negative.
Now consider the trace of $H_{k+1}$, from (38), we have
\[ Tr(H_{k+1}) = Tr(H_0) - \sum_{j=0}^{k} \|y_j H_j y_j\|^2 \]

Since \( H_{k+1} \) is positive definite, the right hand side of (46) is positive. By lemma (3.2) there exists \( n > 0 \) which is independent of \( k \) such that

\[ \sum_{j=0}^{k} \|y_j H_j y_j\|^2 \leq \frac{k}{n} \]  

(47)

Note that

\[ (y_j^T H_j y_j)^2 \leq \|H_j y_j\|^2 \|y_j\|^2 \]  

(48)

and

\[ y_j^T H_j y_j > g_{j+1}^T H_j g_{j+1} \]  

(49)

By the positive definiteness of \( H_j \) and exact line search, then by using (49), (48) and (47) in turn, we obtain:

\[ \sum_{j=0}^{k} \frac{g_{j+1}^T H_j g_{j+1}}{\|y_j\|^2} \leq \frac{k}{\|y_j\|^2} \|y_j\|^2 \leq \frac{k}{n} \]  

(50)

By using Cauchy-Schwarz inequality and (50)

\[ \sum_{j=0}^{k} \frac{\|g_{j+1}\|^2}{g_{j+1}^T H_j g_{j+1}} \geq \left( \sum_{j=0}^{k} \frac{\|g_{j+1}\|^2}{\|y_j\|^2} \right)^2 / \sum_{j=0}^{k} \frac{g_{j+1}^T H_j g_{j+1}}{\|y_j\|^2} \]

\[ \geq \frac{n}{k} \left( \sum_{j=0}^{k} \|g_{j+1}\|^2 \right)^2 \]  

(51)

Now suppose that the theorem is not true, that is, there exists \( \epsilon > 0 \) such that for all sufficiently large \( k \),

\[ \|g_k\| \geq \epsilon \]  

(52)

Also, by lemma (3.3), there exists a constant \( \eta > 0 \) such that

\[ \|x^* - x\| \leq (x^* - x)^T (g(x^*) - g(x)) / \eta \]

\[ f(x_k) - f(x_{k+1}) \geq \frac{1}{2} \eta \|s^*_k\|^2 \]

which gives \( \|s^*_k\| \to 0 \) and further \( \|y_k\| \to 0 \). Then, by (51) and (52) we deduce, for any \( k \) sufficiently large, that

\[ \sum_{j=0}^{k} \|g_{j+1}\|^2 \geq Nk \]  

(53)

The above inequality implies that the sum of the last two terms in (45) is negative.

By (53) and (45), we immediately obtain:

\[ Tr(B_{k+1}) < \frac{\|g_{k+1}\|^2}{g_{k+1}^T H_{k+1} g_{k+1}} \]  

(54)

Note that, for a symmetric positive definite matrix, the inverse of trace is the lower bound of the last eigenvalue of inverse of matrix. Then, it follows from (54) that

\[ \frac{g_{k+1}^T H_{k+1} g_{k+1}}{\|g_{k+1}\|^2} < \mu, \]  

(55)

Where \( \mu \) is the lower bound of the last eigenvalue of \( H_{k+1} \). However, from the property of Rayleigh quotient [9], we have

\[ \frac{g_{k+1}^T H_{k+1} g_{k+1}}{\|g_{k+1}\|^2} > \mu, \]  

(56)

which contradicts (55). This contradiction proves that \( \{x_k\} \) converges to \( x^* \) and that our theorem holds.

4. Local Linear Convergence of \( \alpha - DFP \) Method

Now, we shall prove the local linear convergence of \( \alpha - DFP \) method of equivalent formula (18) for \( \alpha > 0 \) under exact line search.

The \( \alpha - DFP \) iteration we consider is:

\[ x_{k+1} = x_k - \alpha B_k^{-1} F(x_k) \]  

(57)

Let \( F: \mathbb{R}^n \to \mathbb{R}^n \) satisfy assumption (A). Then for any \( u, v, \) \( x^* \in D \subset \mathbb{R}^n \) we have

\[ \|F(u) - F(v) - F'(x^*)(u - v)\| \leq \gamma \sigma(u, v) \|u - v\| \]  

(59)

Furthermore, assume that \( F' \) satisfy assumption (C), then

\[ \|F(u) - F(v) - F'(x^*)(u - v)\| \leq \gamma \sigma(u, v) \|u - v\| \]  

(60)

and

\[ \|F(u) - F(v) - F'(x^*)\| \leq \gamma \frac{\|u - x^*\| \|v - x^*\|}{2} \|u - v\|, \]  

(61)

where \( (u, v) = \max\{\|u - x^*\|, \|v - x^*\|\} \). (62)

Lemma (4.3)

Let \( y, s^* \in \mathbb{R}^n, y, s^* \neq 0, \) and \( \rho \in (0, 1) \). If \( \|y - s^*\| \leq \rho \|y\|, \)

\[ \frac{g_{k+1}^T H_{k+1} g_{k+1}}{\|g_{k+1}\|^2} < \mu, \]  

(55)

Where \( \mu \) is the lower bound of the last eigenvalue of \( H_{k+1} \). However, from the property of Rayleigh quotient [9], we have

\[ \frac{g_{k+1}^T H_{k+1} g_{k+1}}{\|g_{k+1}\|^2} > \mu, \]  

(56)

which contradicts (55). This contradiction proves that \( \{x_k\} \) converges to \( x^* \) and that our theorem holds.
then \( y^T s^* \) is positive and

\[
1 - \frac{\|s\|^2}{\|y\|^2} \leq \rho \leq 1 - \left( \frac{y^T s^*}{\|y\| \|s\|} \right)^2 \leq \rho^2 \tag{63}
\]

Conversely, if \( y^T s^* > 0 \) and (63) holds, then

\[
\|y - s\|^2 \leq 3\rho \|y\|^2 \tag{64}
\]

Theorem (4.4) [9]

Let \( F : \mathbb{R}^n \rightarrow \mathbb{R}^n \) satisfy the assumptions (A), (B) and (C) in (4.1), \( U \) an update function such that for all \( (x_k, B_k) \in \text{dom} \ U \)

\[
B_{k+1} \in U(x_k, B_k),
\]

where \( y \) is some constant, or that

\[
\|B_{k+1} - F'(x^*)\| \leq \|B_k - F'(x^*)\| + \frac{\nu}{2}(\|x_{k+1} - x^*\| + \|x_k - x^*\|), \tag{66}
\]

where \( \rho_1 \) and \( \rho_2 \) are some constants, and

\[
\sigma(x_k, x_{k+1}) = \max(\|x_k - x^*\|^2, \|x_{k+1} - x^*\|^2). \tag{68}
\]

Then, there exist constants \( \epsilon \) and \( \delta \), such that, for all \( \|x_k - x^*\| < \epsilon \) and \( \|B_0 - F'(x^*)\| < \delta \), the iteration (57) and (65) is well-defined, and \( \{x_k\} \) converges to \( x^* \) linearly.

To study the local convergence of \( \alpha - DFP \) method, it is required to estimate \( \|B_{k+1} - \nabla^2 f(x^*)\| \).

As show in the following theorem, there is a matrix

\[
P = I - \frac{x_k^T x_k}{y_k^T y_k} \in B_{k+1} - \nabla^2 f(x^*).
\]

Since

\[
\|P\|_2 = \frac{\|x_k\|\|y_k\|}{y_k^T y_k}, \tag{69}
\]

It is a secant of the angle between \( y_k \) and \( s_k \). In general, \( y_k \) and \( s_k \) is not parallel. So \( \|P\|_2 \) may be quite big, and it is not suitable to estimate \( \|B_{k+1} - \nabla^2 f(x^*)\| \) by means of \( \ell_2 \)-norm. However, near \( x^* \),

\[
f(x) \text{ closes a quadratic function, and hence } A^{-1}z y_k \text{ and } A^{-1}z s_k \text{ are approximately parallel, where } A = \nabla^2 f(x^*). \]

In motives us some weighted norm to estimate \( \|B_{k+1} - \nabla^2 f(x^*)\| \). Then, we have

\[
\|E\|_{\alpha-DFP} = \|E\|_{2 \tau^2} = \left\| A^{-1} \nabla^2 f(x^*) \right\|_2. \tag{70}
\]

Below, we first develop the linear convergence of \( \alpha - DFP \).

Theorem (4.5)

Let \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) satisfy Assumption (c) in (4.1). Also let

\[
\mu \gamma \sigma(x_k, x_{k+1}) \leq \frac{1}{3} \tag{71}
\]

In a neighborhood of \( x^* \), where \( \mu = \left\| \nabla^2 f(x^*) \right\| \), \( \sigma(x_k, x_{k+1}) = \max(\|x_k - x^*\|, \|x_{k+1} - x^*\|) \), then, there exist \( \epsilon > 0 \) and \( \delta > 0 \) such that for \( \|x_k - x^*\| < \epsilon \) and \( \|B_0 - \nabla^2 f(x^*)\|_{\alpha - DFP} < \delta \), the iteration (57) and equivalent formula (18) of \( \alpha - DFP \) method is well-defined, and the produced sequence \( \{x_k\} \) converges to \( x^* \) linearly.

Proof:

Based on the lemma (4.4), to prove the linear convergence of \( \alpha - DFP \) method, it is enough to prove

\[
\|B_{k+1} - \nabla^2 f(x^*)\|_{a-DFP} \leq \left\{ 1 + \rho_1 \sigma(x_k, x_{k+1}) \right\} \|B_0 - \nabla^2 f(x^*)\|_{a-DFP} + \rho_2 \sigma(x_k, x_{k+1}) \tag{72}
\]

where \( \rho_1 \) and \( \rho_2 \) are positive constants independent of \( x_k \) and \( x_{k+1} \), \( \sigma \) is defined by (68).

Le \( A = \nabla^2 f(x^*) \) and \( M = [\nabla^2 f(x^*)]^{-\frac{1}{2}} \), which is \( F'(x^*) \) and \( [F'(x^*)]^{-\frac{1}{2}} \) respectively, also are symmetric positive definite matrices.

From (57) and the formula (18) of \( \alpha - DFP \), it follows that

\[
B_{k+1} - A = P^T (B_k - A) + \frac{(y_k - As_k)^T y_k + y_k (y_k - As_k)^T y_k}{y_k^T s_k} \tag{73}
\]

where

\[
P = I - \frac{x_k^T x_k}{y_k^T y_k}. \tag{74}
\]

Thus, from (73), one has

\[
\|B_{k+1} - A\|_{a-DFP} \leq \|P^T (B_k - A)\|_{a-DFP} + \left\| \frac{(y_k - As_k)^T y_k + y_k (y_k - As_k)^T y_k}{y_k^T s_k} \right\|_{a-DFP} \tag{75}\]

Note that \( \|P\|_2 \) is defined by (69).

The first term on the right hand side of (75) can be estimate as:

\[
\|P^T (B_k - A)\|_{a-DFP} \leq \left\| A^2 P A^{-2} \right\|_2 \|B_k - A\|_{a-DFP} \leq \frac{1}{\tau^2} \|B_k - A\|_{a-DFP} \tag{76}
\]

moreover, for the rest two terms on the right hand side of (75) and by (70) we have:

\[
\left\| \frac{(y_k - As_k)^T y_k}{y_k^T s_k} \right\|_{a-DFP} \leq \frac{1}{\tau} \left\| A^2 \right\| \left\| \frac{y_k - As_k}{y_k^T s_k} \right\|_{a-DFP} \tag{77}
\]

...
and

\[ \left\| \frac{y_k(y_k - As_k)}{y_k s_k^T} \right\|_{a-DPP} = \left\| A^{-1}2y_k(y_k - As_k)^TAP^{-1} \right\|_F \]

\[ \left\| \frac{y_k(y_k - As_k)^T}{y_k s_k^T} \right\|_{a-DPP} \leq \frac{1}{z^2} \left\| \frac{1}{A^2s_k} \right\| \]

where

\[ z = \frac{y_k s_k^T}{A^2y_k s_k^T} = \left( \frac{A^2y_k}{A^2s_k} \right)^T \left( \frac{A^2s_k}{A^2s_k} \right) \]

which by lemma (4.3) implies the curvature condition

\[ \left( \frac{A^2y_k}{A^2s_k} \right)^T \left( \frac{A^2s_k}{A^2s_k} \right) = y_k^T s_k^T > 0 \]

Now, we estimate \( \|B_{k+1} - A\|_{a-DPP} \) by using (76), (77) and (78), we have

\[ \|B_{k+1} - A\|_{a-DPP} \leq \frac{1}{z^2} \|B_k - A\|_{a-DPP} + \left( \frac{1}{z^2} + 1 \right) \left( \frac{1}{A^2s_k} \right) \]

\[ \leq [1 + \mu \gamma (x_k, x_{k+1})] \|B_k - A\|_{a-DPP} + \frac{4}{3} + \frac{4}{3} \mu \gamma (x_k, x_{k+1}) \]

Consequently, if \( x_k \) and \( x_{k+1} \) are in the neighborhood of \( x^* \), then

\[ 1 - z^2 \leq [\mu \gamma (x_k, x_{k+1})]^2 < \frac{1}{2} \]

\[ 1 - \frac{1}{2} < \mu \gamma (x_k, x_{k+1}) \]

So, the two terms in (80) satisfy respectively.

\[ \frac{1}{z^2} \|B_k - A\|_{a-DPP} < [1 + \mu \gamma (x_k, x_{k+1})] \|B_k - A\|_{a-DPP} \]

and

\[ \left( \frac{1}{z^2} + 1 \right) \left( \frac{1}{A^2s_k} \right) \leq [1 + \mu \gamma (x_k, x_{k+1})] \|B_k - A\|_{a-DPP} + 3 \mu \gamma (x_k, x_{k+1}) \]

which completes the proof by applying lemma (4.4) with \( \rho_1 = \mu \gamma \) and \( \rho_2 = 3 \mu \gamma \).

5. Super Linear Convergence of \( \alpha - DFP \) Method

Now, we shall prove the super linear convergence of the \( \alpha - DFP \) method. the convergence analysis in this section mainly Dennis and Mor'e [2]. The super linear convergence of the sequence \( \{x_k\} \) generated by the iteration (57) is generally characterized by the following theorem.

Theorem (5.1) [2]

Let \( F: R^n \to R^n \) be a sequence of nonsingular matrices. Suppose for \( x_k \in D \), that the iteration generated by (57) remains in \( D \), \( x_k \neq x^* (\forall k \geq 0) \). Suppose also that \( \{x_k\} \) converges to \( x^* \). Then \( \{x_k\} \) converges to \( x^* \) at super linear rate if and only if

\[ \lim_{k \to \infty} \frac{\|b_k - F(x^*)\|_{s_k^T}}{\|s_k^T\|} = 0 \] (84)

Theorem (5.1) indicates that if \( b_k \) converges to \( F'(x^*) \) along the direction \( s_k^T \), then \( \alpha - DFP \) method converges super linearly. This theorem is very important in analysis of \( \alpha - DFP \). Equation (84) is called the Dennis and Mor'e characterization of super linear convergence.

To apply theorem (5.1), we need a refinement estimate \( \|b_{k+1} - F'(x^*)\|_{s_k^T} \) which is established with the help of the following lemmas.

Lemma (5.2)

Let \( M \in R^{n \times n} \) be a nonsingular symmetric matrix, If, for
\[ \lambda \in \left[0, \frac{1}{2}\right]. \]

The inequality
\[ \|My_k - M^{-1}s_k\| \leq \lambda \|M^{-1}s_k\| \] holds, then for any non-zero matrix \( E \in \mathbb{R}^{n \times n} \), we have:

(a): \( (1 - \lambda)\|M^{-1}s_k\|^2 \leq y_k^T s_k \leq (1 + \lambda)\|M^{-1}s_k\|^2 \)  

(b): \( \|E \left(I - \frac{(M^{-1}s_k)(M^{-1}s_k)^T}{y_k^T s_k}\right)\|_F \leq 1 - \rho \theta^2 \|E\|_F \)  

(c): \( \|E \left[I - \frac{(M^{-1}s_k)(My_k)^T}{y_k^T s_k}\right]\|_F \leq \sqrt{1 - \rho^2} \|E\|_F \) \]  

Where
\[ \rho = \frac{1 - 2\lambda}{(1 - \lambda)^2} \in \left[1, \frac{2}{3}\right], \theta = \frac{\|Em^{-1}s_k\|}{\|E\|_F\|M^{-1}s_k\|} \in [0, 1] \]  

Proof:  
Note that  
\[ y_k^T s_k = (My_k)^T(M^{-1}a_{s_k}) = (My_k - M^{-1}a_{s_k})^T M^{-1}a_{s_k} + \|M^{-1}a_{s_k}\|^2 \]  

Also, it follows from Cauchy-Schwarz inequality and (85) that  
\[ \|My_k - M^{-1}a_{s_k}\|^2 \leq \|My_k - M^{-1}a_{s_k}\|\|M^{-1}a_{s_k}\| \]  

From (90) and (91), we get  
\[ y_k^T s_k \leq \lambda \|M^{-1}a_{s_k}\|^2 + \|M^{-1}a_{s_k}\|^2 = (1 + \lambda)\|M^{-1}a_{s_k}\|^2 \]  

And, in same way, we have  
\[ (1 - \lambda)\|M^{-1}a_{s_k}\|^2 \leq y_k^T s_k^* \]  

Which gives the first result (a).  
Now, we will prove (b) By using the property of the Frobenius norm of a rank-one update.  
\[ \|A + xy^T\|_F^2 = \|A\|_F^2 + 2y^TA^Tx + \|x\|^2\|y\|^2 \]  

To prove (b) we need the following property  
\[ \|E(I - uu^T)\|_F^2 = \|E\|_F^2 - 2v^TEu + \|Eu\|^2\|v\|^2. \]  

In particular,  
Let \( u = \frac{M^{-1}a_{s_k}}{y_k^T a_{s_k}} \) and \( v = M^{-1}a_{s_k} \)  
\[ \left\|E \left[I - \frac{(M^{-1}s_k)(M^{-1}s_k)^T}{y_k^T a_{s_k}}\right]\right\|_F^2 = \|E\|_F^2 - 2\frac{(M^{-1}s_k)^TE^TE(M^{-1}s_k)}{y_k^T a_{s_k}} \]  

\[ + \frac{\|EM^{-1}s_k\|^2}{y_k^T a_{s_k}} \leq \|E\|_F^2 \]  

By using (a) and (89), we get:  
\[ \|M^{-1}a_{s_k}\|^2 - 2y_k^T a_{s_k} \leq \frac{1}{1 - \lambda} y_k^T a_{s_k} - 2 \]  

\[ = \left(1 - \frac{2\lambda}{1 - \lambda}\right) y_k^T a_{s_k} \]  

And therefore  
\[ \left\|E \left[I - \frac{(M^{-1}s_k)(M^{-1}s_k)^T}{y_k^T a_{s_k}}\right]\right\|_F^2 \leq 1 - \rho \theta^2 \|E\|_F \]  

Which shows (b):  
\[ \left\|E \left[I - \frac{(M^{-1}s_k)(M^{-1}s_k)^T}{y_k^T a_{s_k}}\right]\right\|_F \leq \sqrt{1 - \rho^2 \|E\|_F} \]  

Finally, we prove (c) by means of (b). It enough to prove that  
\[ \left\|E \frac{M^{-1}s_k(M^{-1}s_k - My_k)^T}{y_k^T a_{s_k}}\right\|_F \leq \frac{1}{y_k^T a_{s_k}} \]  

(92)  
From (a), we have:  
\[ \frac{1}{y_k^T a_{s_k}} \leq \frac{1}{(1 - \lambda)\|M^{-1}a_{s_k}\|^2} \]  

(93)  
\[ \frac{1}{y_k^T a_{s_k}} \|M^{-1}a_{s_k}\|\|M^{-1}a_{s_k} - My_k\| \leq (1 - \lambda)^{-1}\|M^{-1}a_{s_k}\|\|M^{-1}a_{s_k} - My_k\| \]  

(94)  
which proves (c).  
We have known that if \( f: R^n \to R \) satisfies Assumption (4.1), then (72) holds.
Then under Assumptions of the theorem (4.5), the preceding lemma can be applied with the sitting
\[ A = F'(x') = \nabla^2 f(x'), M = [F'(x')]^{-1} = A^{-1}, \]
\[ \psi_k = \|B_k - \nabla f(x')\|_{a-DFP}, \psi_{k+1} = \|B_{k+1} - \nabla f(x')\|_{a-DFP}, \]
and \( \xi_k = \max\{\rho_1\sigma(x_k, x_{k+1}), \rho_2\sigma(x_k, x_{k+1})\}. \)

Due to the linear convergence of the sequence \( \{x_k\} \) gives in theorem (4.5), we have \( \sum_{k=1}^{\infty} \xi_k < +\infty \), and can sequently by lemma (5.3), there exists a constant \( h \geq 0 \) such that
\[ \lim_{k\to\infty} \|B_k - F'(x')\|_{a-DFP} = h \] (94)

Hence
\[ \lim_{k\to\infty} \|B_k - F'(x')\|_{a-DFP} \]
exists. Lemma (5.3):

Let \( \{\phi_k\} \) and \( \{\delta_k\} \) be sequences of nonnegative numbers satisfying
\[ \phi_{k+1} \leq (1 + \delta_k)\phi_k + \delta_k \] (95)
and
\[ \sum_{k=1}^{\infty} \delta_k < +\infty \] (96)
then \( \{\phi_k\} \) is converges.

These results together then give rise to a refinement estimate of \( \|B_{k+1} - F'(x')\|_{a-DFP} \) as follows:

Lemma (5.4)

Under the assumption of theorem(4.5), there exist positive constants \( \lambda_1, \lambda_2 \) and \( \lambda_3 \), such that \( \forall x_{k+1} \in N(x', \epsilon) \), we have
\[ \|B_{k+1} - \nabla^2 f(x')\|_{a-DFP} \leq [\sqrt{1 - \lambda_1\sigma^2} + \lambda_2\sigma(x_k, x_{k+1})]\|B_k - \nabla^2 f(x')\|_{a-DFP} + \frac{1}{\lambda_3\sigma(x_k, x_{k+1})}, \] (97)
where \( \sigma \) is defined by (68) and
\[ \delta_k = \\frac{[\sigma^2 f(x')^{\frac{1}{2}} (\psi_k - [\sigma^2 f(x')^{\frac{1}{2}}]^{\frac{1}{2}})\sigma_k]}{[B_k - \nabla^2 f(x')]_{a-DFP} [\sigma^2 f(x')^{\frac{1}{2}}]^{\frac{1}{2}} \|\sigma_k\|} \] (98)

Proof:

First m we write = \( \nabla^2 f(x') \). From (75), we have:
\[ \|B_{k+1} - A\|_{a-DFP} \leq \|P^T(B_k - A)\|_{a-DFP} + \| \( y_k - A\alpha s_k \) y_k^T \|_{a-DFP} + \| \( y_k (y_k - A\alpha s_k)^T \|_{a-DFP} \right) \]

\[ R = I - \frac{y_k s_k}{y_k^T s_k}, E_k = A^{-\frac{1}{2}}(B_k - A)A^{-\frac{1}{2}} \] (99)

And
\[ \|P^T(B_k - A)\|_{a-DFP} = \| \left( A^{-\frac{1}{2}} P^T A^{-\frac{1}{2}} \right) \left( A^{-\frac{1}{2}}(B_k - A)A^{-\frac{1}{2}} \right) \| \]
\[ = \| R^T E R \| F \]

Similar to the proof of the theorem (4.5), we known that there exists \( \rho_3 > 0 \) and \( \rho_4 > 0 \) such that
\[ \left\| \left( y_k - A\alpha s_k \right) y_k^T \right\|_{a-DFP} \leq \frac{1}{\lambda_3\sigma(x_k, x_{k+1})} \]
\[ \leq \rho_3\sigma(x_k, x_{k+1}), \]
\[ \left\| \left( y_k (y_k - A\alpha s_k)^T \|_{a-DFP} \leq \frac{1}{\lambda_3\sigma(x_k, x_{k+1})} \]
\[ \leq \rho_4\sigma(x_k, x_{k+1}). \]

If we let \( \lambda_3 = \rho_3 + \rho_4 \), then (75) becomes:
\[ \|B_{k+1} - A\|_{a-DFP} \leq \| R^T E R \| F + \lambda_3\sigma(x_k, x_{k+1})(100) \]

Since
\[ \frac{\left[ \frac{1}{\lambda_3\sigma(x_k, x_{k+1})} \right]}{\|A^{-\frac{1}{2}}(B_k - A)\|_{a-DFP}} \leq \frac{\mu \gamma \sigma(x_k, x_{k+1})}{\frac{1}{3}}, \]

Then, by use of lemma (5.2), and from (99), we get:
\[ \| R^T E R \| F \leq \sqrt{1 - \rho\rho^2} (1 - \lambda)^{-1} \frac{\frac{1}{\lambda_3\sigma(x_k, x_{k+1})}}{\|A^{-\frac{1}{2}}(B_k - A)\|_{a-DFP}} \| R^T E \| F \]

Note that \( \| R^T E \| F = \| E R \| F \), thus, by using lemma (5.2) again, we obtain:
\[ \| R^T E \| F \leq \sqrt{1 - \rho\rho^2} + (1 - \lambda)^{-1} \frac{\frac{1}{\lambda_3\sigma(x_k, x_{k+1})}}{\|A^{-\frac{1}{2}}(B_k - A)\|_{a-DFP}} \sqrt{1 - \rho\rho^2} \]
\[ + (1 - \lambda)^{-1} \frac{\frac{1}{\lambda_3\sigma(x_k, x_{k+1})}}{\|A^{-\frac{1}{2}}(B_k - A)\|_{a-DFP}} \| E \| F. \]
Where \( \theta_k \) is defined by (98). From lemma (5.2) again and by (89), we get
\[
\| R^{T}E\|_{F} \leq \left[ \sqrt{1 - \rho \theta^{2}} + \frac{5}{2}(1 - \lambda)^{-1} \left( \frac{1}{\| A^2 \theta \|} \right) \right] \| E \|_{F}
\]
\[
\leq \left[ \sqrt{1 - \lambda_1 \theta^{2}} + \lambda_2 \gamma \sigma(x_k, x_{k+1}) \right] \| E \|_{F} \tag{101}
\]

Where \( \lambda_1 = \rho, \lambda_2 = \frac{15}{4} \mu \gamma \). Substitution (101) into (100) and from (99), we deduce the desired result (97). The proof is complete.

Finally, using the above four lemmas, we can establish the following super linear convergence theorem of \( \alpha - \text{DFP} \) method.

Theorem (5.5)
Under the assumption of theorem (4.5), \( \alpha - \text{DFP} \) defined by (57) and (58) is convergent super linearly.

Proof:
Since \( (1 - \lambda_1 \theta^{2})^{\frac{1}{2}} \leq \left( 1 - \frac{1}{2} \lambda_1 \theta^{2} \right) \), then (95) can be written as:
\[
\| B_{k+1} - A \|_{\alpha - \text{DFP}} \leq \left( 1 - \frac{1}{2} \lambda_1 \theta^{2} \right) \| B_k - A \|_{\alpha - \text{DFP}}
\]
\[
+ \lambda_2 \gamma \sigma(x_k, x_{k+1}) \| B_k - A \|_{\alpha - \text{DFP}}
\]
\[
+ \lambda_3 \gamma \sigma(x_k, x_{k+1})
\]

\[
\left( \frac{1}{2} \lambda_1 \theta^{2} \right) \| B_k - A \|_{\alpha - \text{DFP}} \leq \| B_k - A \|_{\alpha - \text{DFP}} - \| B_{k+1} - A \|_{\alpha - \text{DFP}}
\]
\[
+ \lambda_2 \gamma \sigma(x_k, x_{k+1}) \| B_k - A \|_{\alpha - \text{DFP}}
\]
\[
+ \lambda_3 \gamma \sigma(x_k, x_{k+1})
\]

Summing the above from \( k = 1 \) to infinity gives:
\[
\frac{1}{2} \lambda_1 \sum_{k=1}^{\infty} \theta_k \| B_k - A \|_{\alpha - \text{DFP}} \leq \sum_{k=1}^{\infty} \| B_k - A \|_{\alpha - \text{DFP}}
\]
\[
- \sum_{k=1}^{\infty} \| B_{k+1} - A \|_{\alpha - \text{DFP}}
\]
\[
+ \lambda_2 \gamma \sum_{k=1}^{\infty} \sigma(x_k, x_{k+1}) \| B_k - A \|_{\alpha - \text{DFP}}
\]
\[
+ \lambda_3 \gamma \sum_{k=1}^{\infty} \sigma(x_k, x_{k+1})
\]

Since, from theorem (4.5), \( \{ x_k \} \) is linearly convergent, then \( \sum_{k=1}^{\infty} \sigma(x_k, x_{k+1}) < \infty \).

Also, since \( \| B_k - A \|_{\alpha - \text{DFP}} \) is bounded, then
\[
\frac{1}{2} \lambda_1 \sum_{k=1}^{\infty} \theta_k \| B_k - A \|_{\alpha - \text{DFP}} < \infty
\]

By (94), the \( \lim_{k \to \infty} \| B_k - A \|_{\alpha - \text{DFP}} \) exists.

Hence, if some subsequence of \( \{ \| B_k - A \|_{\alpha - \text{DFP}} \} \) converges to zero.

The whole sequence converges to zero.

Therefore,
\[
\lim_{k \to \infty} \frac{\| (B_k - A)s_k^* \|}{\| s_k^* \|} = 0,
\]

Which proves the super linear convergence of \( \{ x_k \} \) by theorem (5.1). Otherwise, there exists a positive constant \( \tau \) such that
\[
\| B_k - A \|_{\alpha - \text{DFP}} \geq \tau, \forall k \geq k_0,
\]

then
\[
\frac{1}{2} \lambda_1 \sum_{k=1}^{\infty} \theta_k^2 < \infty
\]

Since \( \lambda_1 = \frac{4}{3} > 0 \), it follows that \( \lim_{k \to \infty} \theta_k = 0 \).

Furthermore, we have:
\[
\lim_{k \to \infty} \frac{\| (B_k - A)s_k^* \|}{\| s_k^* \|} \leq \lim_{k \to \infty} \frac{\| A^2 \| \| A^{-1}(B_k - A)s_k \|}{\| A^2 \|^{-1} \| A^{-1}A^2s_k \|}
\]
\[
= \lim_{k \to \infty} \| A \| \cdot \| B_k - A \| \cdot \| A^{-1}(B_k - A)s_k \| \| B_k - A \| \cdot \| A^{2}s_k \|
\]
\[
= \lim_{k \to \infty} \| A \| \cdot \| B_k - A \| \| \theta_k \|
\]

Where \( \theta_k \) is defined by (98).

Then, by using \( \theta_k \to 0 \), we immediately obtain:
\[
\lim_{k \to \infty} \frac{\| (B_k - A)s_k^* \|}{\| s_k^* \|} = 0.
\]

Hence, \( \{ x_k \} \) is convergent super linearly, we complete the proof.

6. Numerical Results

This section is devoted to numerical experiments. Our purpose was to check whether the modified \( \alpha - \text{DFP} \) algorithm provide improvements on the corresponding standard \( \text{DFP} \) algorithm. The programs were written in \text{MATLAB}. The reason for their selection is that the problems appear to have been used in standard problems in most the literature these functions represent a result of application in the branch of technology and industry.

The test functions are chosen as follows:
1 − f(x) = (1 − x_1)^2 + (1 − x_2)^2. [1]  

2 − A quadratic function. [10]  
\[ f(x) = \sum_{i=1}^{4} (10^{i-1}x_1^4 + x_1^2 + 10^{1-i}x_1^2) \]

3 − Rosen brook's function. [4]  
\[ f(x) = (1 − x_1)^2 + (x_2 − x_1)^2. \]

4 − Rosenbrock's cliff function [8]  
\[ f(x) = 10^{-4}(x_1 − 3)^2 − (x_1 − x_2) + e^{20(x_1−x_2)}. \]

5 − Generalized Edger function. [1]  
\[ f(x) = \sum_{i=1}^{n/2} (x_{2i−1} − 2)^4 + (x_{2i−1} − 2)^2x_{2i} + (x_{2i} + 1)^2. \]

6 − Extended Himmelbla function [1]  
\[ f(x) = \sum_{i=1}^{n/2} (x_{2i−1}^2 + x_{2i} − 11)^2 + (x_{2i−1} + x_{2i}^2 − 7)^2. \]

7 − Rosen rock's function [6]  
\[ f(x) = \sum_{i=3}^{n/2} [100(x_1 − x_i^3)^2 + (1 − x_i)^2]. \]

8 − Trigonometric function [1]  
\[ f(x) = \sum_{i=1}^{n} \left[ n - \sum_{j=1}^{i} \cos x_j + i(1 - \cos x_i) - \sin x_i + e^{x_i} - 1 \right]^2. \]

9 − Extended Rosen rock function [1]  
\[ f(x) = \sum_{i=1}^{n/2} c(x_{2i−1} − x_{2i−2})^2 + (1 − x_{2i−1})^2, \ c = 100 \]

10 − Watson function [6]  
\[ F(x) = \sum_{i=1}^{j} f_i^3(x) \]
\[ f_i(x) = \sum_{j=2}^{3}(j−1)x_{j−1}^2 - (\sum_{j=1}^{3}x_{j}t_{j−1})^2 - 1. \]
\[ t_j = \frac{i}{29} \]

11 − Freudenstein and Roth function [3]  
\[ f(x) = (-13 + x_1 + [(5 - x_2)x_2 - 2]x_2)^2 \]
\[ + (-29 + x_1 + [(x_2 + 1)x_2 - 14]x_2)^2 \]

| Fun. | Starting point | Dim. | DFP Feval | DFP Iter. | α − DFP Feval | α − DFP Iter. | The Best |
|------|----------------|------|------------|----------|----------------|----------------|----------|
| 1    | [0; 0]^T       | 2    | 6.0962e-016 | 2        | 1.7909e-017 | 2              | Same     |
| 1    | [−1; −1]^T     | 2    | 2.7900e-020 | 2        | 2.7900e-020 | 2              | Same     |
| 1    | [0; 1]^T       | 2    | 3.2494e-016 | 2        | 2.9082e-017 | 2              | Same     |
| 2    | [−1; 0; −1; 0]^T| 4    | 1.5076e-007 | 9        | 9.0146e-007 | 7              | α − DFP  |
| 2    | [−1; 0; 0; 0]^T| 4    | 1.8142e-013 | 3        | 7.3965e-011 | 3              | α − DFP  |
| 2    | [−3; 10; ⋯]^T | 4    | 7.9747e-008 | 34       | 7.2012e-006 | 9              | α − DFP  |
| 3    | [0; 0]^T       | 2    | 1.7525e-018 | 20       | 1.6887e-016 | 14             | α − DFP  |
| 3    | [0; −5]^T      | 2    | 1.5613e-016 | 3        | 7.5140e-015 | 3              | α − DFP  |
| 3    | [−3; −3]^T     | 2    | 3.0198e-017 | 14       | 2.404e-017  | 14             | Same     |
| 3    | [−0.5; ⋯]^T   | 4    | 0.2011      | 3        | 0.2011      | 3              | Same     |
| 4    | [0; 0; ⋯]^T   | 12   | 0.2004      | 3        | 0.2004      | 3              | Same     |
| 4    | [0; ⋯]^T      | 12   | 0.2077      | 3        | 0.2077      | 3              | Same     |
| 5    | [−3; 0; ⋯]^T  | 2    | 5.5433e-015 | 8        | 8.8715e-014 | 3              | α − DFP  |
| 5    | [0; 5; ⋯]^T   | 18   | 2.9846e-008 | 9        | 3.6384e-008 | 6              | α − DFP  |
| 5    | [−1; ⋯]^T     | 18   | 1.4714e-010 | 7        | 5.3611e-009 | 6              | α − DFP  |
| 6    | [5; 10]^T     | 2    | 8.1785e-012 | 7        | 3.2157e-011 | 6              | α − DFP  |
| 6    | [0; 0]^T      | 2    | 1.3697e-013 | 8        | 3.1675e-008 | 9              | α − DFP  |
| 6    | [0; −1]^T     | 2    | 1.1910e-012 | 8        | 2.7507e-012 | 8              | α − DFP  |
| 7    | [−1; 1]^T     | 8    | 2.0658e-011 | 6        | 1.9381e-010 | 24             | α − DFP  |
| 7    | [0; 1; ⋯]^T  | 8    | 5.1583e-012 | 6        | 3.1487e-011 | 8              | α − DFP  |
| 7    | [0; ⋯]^T      | 4    | 1.7728e-010 | 2        | 1.7728e-010 | 2              | Same     |
| 8    | [−0.5; ⋯]^T  | 12   | 7.6464e-007 | 13       | 4.6941e-006 | 5              | α − DFP  |
| 8    | [0.5; ⋯]^T   | 12   | 3.7119e-006 | 13       | 3.7598e-006 | 4              | α − DFP  |
| 8    | [2; ⋯]^T     | 12   | 6.8965e-007 | 21       | 1.8229e-006 | 5              | α − DFP  |
| 9    | [1; 1; ⋯]^T  | 3    | 0.0020      | 40       | 0.0118      | 9              | α − DFP  |
| 9    | [−4; ⋯]^T    | 3    | 3.5522      | 31       | 40.2765     | 8              | α − DFP  |
| 10   | [1; 1; 1; ⋯]^T| 4    | 6.0707e-017 | 3        | 1.0838e-016 | 3              | α − DFP  |
| 10   | [1; 0; 1; 0]^T| 4    | 2.1626e-013 | 3        | 3.5351e-008 | 3              | α − DFP  |
| 10   | [0; ⋯]^T     | 10   | 1.6053e-010 | 2        | 5.5796e-010 | 2              | α − DFP  |
| 11   | [5; 10]^T    | 2    | 5.7785e-010 | 12       | 2.2977e-010 | 8              | α − DFP  |
| 11   | [−10; 20]^T | 2    | 2.3185e-013 | 15       | 1.7572e-013 | 9              | α − DFP  |
7. Conclusion

In this thesis, we introduce a new modified of the DFP say $\alpha - DFP$ update, we show that under certain circumstances this update preserve the value of the determinant of hessian matrix and without Quasi-Newton or based on the Zhang Xu condition.

Global convergence of the proposed method establishes under exact line search. The proposed method possesses local linearly convergence and super linearly convergence for unconstrained optimization problem.

Numerical results show that the proposed is efficient for unconstrained optimization problem compared the modified $\alpha - DFP$ method with the standard DFP method on same function is selected, which suggests that a good improvement has been achieved.

References

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

[2] Dennis J. E., More J., (1974), "A characterization of super linear Convergence and its application to quasi-Newton methods Math and Computation 28 (6) 549-60.

[3] F. Freudenstein and B. Roth,(1962), Numerical solution of system of nonlinear equations, Journal of ACM, Vol. 10, No. 4, pp. 550-556.

[4] H. H. Rosen brock, (1960), An automatic method for finding the greatest least value of a function, Computer Journal, Vol. 3, No. 3, pp. 175-184.

[5] J. E. Dennis, Jr. and Robert B. Schnabel, (1996), Numerical Methods for Unconstrained Optimization and Nonlinear Equations, SIAM, Classics in Applied Mathematics.

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

[7] Saad S. Mahmood, (2011), $\alpha - BFGS$ update for Unconstrained Optimization, Journal of college of Education, No. 1, Almustansiriya University.

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

[9] W. Sun and Y. Yuan, (2006), "Optimization Theory and Method: Nonlinear Programing", Vol. 1 of Springer optimization and its Applications, Springer, New York, NY, USA.

[10] Yuan, Y., (1990), On a Modified Algorithm for Unconstrained Optimization, Computing Center, Academia Sinica, Beijing, China.