GRADUATE SEMINAR REPORT

SOME REMARKS ON SECANT METHOD OF APPROXIMATION FOR THE SOLUTIONS OF SYSTEMS OF NON LINEAR EQUATIONS

BY

YASIN ABDELA

ADVISOR

PROF. DR. RER. NAT. HABIL R. DEUMLICH

SCHOOL OF GRADUATE STUDIES
ADDIS ABABA UNIVERSITY

JUNE 2003
ADDIS ABABA

20

Preface

Even though different methods of approximating the solutions of non-linear equations are highly considered. We are not attempted to present an exhaustive study of the approximations.

We are concerned only to consider Broydens method of approximation for the solutions of nonlinear equations.

This seminar report has five chapters which deal with Newton method, secant method (specially Broydens method), local convergence and other related facts.

Before all I am deeply most grateful to my beloved advisor and instructor Prof. Dr. rer.nat.habil. R. Deumlich for his constrictive comments, suggestions, hints and valuable advices. I am realy lucky to have him as my advisor.

I am also grateful to W/o Tobiaw Tefer for her patience and care in typing the manuscript.

Contents

1. Introduction

This s eminar p aper p resents s ecant m ethod o f approximation for the solution of non linear equations. How ever this method is based on Newton's method. Therefore first we will consider some basic facts of Newton's method. The main idea of secant method is to approximate the Jacobian matrix of Newton's method using functional values that we have already calculated.

1.1 Newton's Method for solving non linear equation of one variable.

Let f: $\Re \rightarrow \Re$ be non-linear function. To approximate the zero of f., let x_c be the initial estimate solution, then better estimate x_+ is obtained by considering a tangent line to the graph of f at the point (x_c, f(x_c)). The x-coordinate of the intersection point of the line and the x-axis is x_+.

$$f'(x_c) = \frac{f(x_c) - f(x_+)}{x_c - x_+}$$

$$(x_c - x_+)\, f'(x_c) = f(x_c) = 0, \text{ since } f(x_+) = 0$$

$$x_c - x_+ = \frac{f(x_c)}{f'(x_c)}$$

$$x_+ = x_c - \frac{f(x_c)}{f'(x_c)} \qquad (1.1.1)$$

This method of approximation is called Newton's method. At each iteration we develop a local model of our function f. In this case,

$$M_c(x) = f(x_c) + f'(x_c)\,(x - x_c)$$

Example 1.1.1: Approximate the square root of 5

Solution:

Let x be the square root of 5 then

$$x^2 = 5$$

$$x^2 - 5 = 0$$

we will approximate the zeros of the function $f(x) = x^2 - 5$

$$x_+ = x_c - \frac{f(x_c)}{f'(x_c)} = 3 - \frac{4}{6} = \quad 2.3333334$$

$$x_{++} = x_+ - \frac{f(x_+)}{f'(x_+)} = 2.3333334 - \frac{0.4444447}{4.6666668}$$

$$= 2.2380953$$

$$x_{+++} = x_{++} - \frac{f(x_{++})}{f'(x_{++})} = 2.236069.$$

Continue in this process in order to get a better estimate of the square root of 5.

1.2 Newton's Method of approximation for the solutions of systems of non-linear equations.

Algorithm 1.2.1

Let F: $\Re^n \rightarrow \Re^n$ be continuously differentiable, J(x) is the Jacobian matrix of F and $x_0 \in \Re^n$.

At each iteration k solve,

$$J(x_k)s_k = -F(x_k) \qquad \text{for } k = 0, 1, 2.$$

$$x_{k+1} = x_k + s_k$$

For each k we develop an affine model of F. In particular at x_c the affine model of F is;

$$M_c = F(x_c) + J(x)\,(x-x_c)$$

Example 1.2.1: Let $F(x) = \begin{pmatrix} x_i + x_2 - 3 \\ x_1^2 + x_2^2 - 9 \end{pmatrix}$ which has roots at $(3, 0)^T$ and $(0, 3)^T$.

Let $x_0 = (1, 5)^T$, find the first two iterations of Newton's method.

$$J(x_0) = \begin{pmatrix} 1 & 1 \\ 2 & 10 \end{pmatrix}, \quad -F(x_0) = \begin{pmatrix} -3 \\ -7 \end{pmatrix}$$

$$J(x_0)\, s_0 = -F(x_0)$$

$$\begin{pmatrix} 1 & 1 \\ 2 & 10 \end{pmatrix} s_0 = \begin{pmatrix} -3 \\ -7 \end{pmatrix}$$

$$S_0 = \begin{pmatrix} \frac{-13}{8} \\ \frac{-11}{8} \end{pmatrix}$$

$x_1 = x_0 + s_0 = (-0.625, 3.625)^T$

$J(x_1)\, S_1 = -F(x_1)$

$$\begin{pmatrix} 1 & 1 \\ \frac{-5}{4} & \frac{29}{4} \end{pmatrix} S_1 = \begin{pmatrix} 0 \\ \frac{145}{32} \end{pmatrix}$$

$$S_1 = \begin{pmatrix} \frac{145}{272} \\ -\frac{145}{272} \end{pmatrix}$$

Then $x_2 = x_1 + S_1 \cong (0.092, 3.092)^T$

1.3 Vector and matrix norms

Definition 1.3.1: A function $\|.\|: \Re^n \to \Re$ is said to be a norm on $\Re^n$ if and only if the following properties hold:

1. $\|v\| \geq 0 \ \forall V \in \Re^n$ and $\|v\| = 0$ if and only if
   $V = 0 \in \Re^n$,
2. $\|\alpha v\| = |\alpha|\, \|v\| \qquad \forall v \in \Re^n$ and $\alpha \in \Re$,
3. $\|v + w\| \leq \|v\| + \|w\| \qquad \forall\, v, w \in \Re^n$.

Let $v = (_v, v_2, \ldots v_n)^T \in \Re^n$, then

$$\|v\| = \max_{1 \leq i \leq n} |v_i|$$

$$\|v\|_1 = \sum_{i=1}^{n} |v_i|$$

$$\|v\|_2 = \left(\sum_{i=1}^{n} (v_i)^2 \right)^{1/2}$$

are examples of vector norms.

Let A be a matrix, $A = (a_{11}, a_{21}, \ldots a_{n1}, a_{12}, \ldots, a_n a_n)^T$

Definition 1.3.2: The Frobenius norm of F denoted by

$||A||_F$ is defined as

$$||A||_F = \left(\sum_{j=1}^{n} \sum_{i=1}^{n} \left| a_{ij} \right|^2 \right)^{1/2}$$

Theorem 1.3.1: Let A, B $\in \Re^{n \times n}$ and Let for any norm $||.||$, $||A.B|| \leq ||A||\ ||B||$, let $||I|| = 1$.

If A is non-singular and $||A^{-1}(B-A)|| < 1$. then B is non-singular and

$$||B^{-1}|| = \frac{\left\| A^{-1} \right\|}{1 - \left\| A^{-1}(B-A) \right\|}$$

Proof:

Since $||A^{-1}(B-A)\ || < 1$, we know that $A^{-1}(B - A)$ is invertible. $A^{-1}(B\text{-}A)$ and A are invertible implies B is veritable.

$$||A^{-1}(B-A)|| = ||A^{-1}B - I|| = ||I - A^{-1}B|| < 1$$

$$||\ (I - (I\text{-}A^{-1}B))^{-1}|| \leq \frac{1}{1 - \left\| I - (A^{-1}B) \right\|}$$

$$||AB^{-1}|| \leq \frac{1}{1 - \left\| I - (A^{-1}B) \right\|}$$

$$||A^{-1}.AB^{-1}|| \leq ||A^{-1}||\ ||AB^{-1}|| \leq \frac{1}{1 - \left\| I - (A^{-1}B) \right\|}$$

$$||B^{-1}|| \leq \frac{\left\| A^{-1} \right\|}{1 - \left\| I - (A^{-1}B) \right\|}$$

$$||B^{-1}|| \leq \frac{\left\| A^{-1} \right\|}{1 - \left\| A^{-1}(B-A) \right\|}$$

Definition: 1.3.3: A function g is Lipschitz continuous with constant γ in a set X, written $g \in Lip_\gamma(X)$ if for every x, y $\in$X,

$$|g(x) - g(y)| \leq \gamma\ |x - y|.$$

Lemma 1.3.2: Let F: $\Re^n \to \Re^m$ be continuously differentiable in the open convex set $D \subseteq \Re^n$, let $x \in D$ and J be Lipschitz continuous at x in the neighborhood D. Then for $x + p \in D$,

$$\|F(x+p) - F(x) - J(x)P\| \leq \frac{\gamma}{2}\|p\|^2.$$

Proof

By definition Jacobian matrix of F.

$$F(x+p) - F(x) - J(x)p = [\int_0^1 J(x+tp)p\, dt] - J(x)p.$$

$$= [\int_0^1 J(x+tp) - J(x)]p\, dt.$$

$$\|F(x+p) - F(x) - J(x)p\| \leq \int_0^1 \|J(x+tp) - J(x)\|\,\|p\|dt$$

$$\leq \int_0^1 \gamma\|tp\|\,\|p\|dt.$$

$$= \gamma\|p\|^2 \int_0^1 t\, dt$$

$$= \frac{\gamma}{2}\|p\|^2.$$

Lemma 1.3.3: Let F: $\Re^n \to \Re^n$ be continuously differentiable in the open convex set $D \subseteq \Re^n$, let $x \in D$ and let J be lipschitz continuous with constant γ in the neighborhood D. Then for $v, u \in D$

$$\|F(v) - F(u) - J(x)(v-u)\| \leq \gamma \frac{(\|v-x\| + \|u-x\|)}{2} \|v-u\|.$$

By Lemma 1.3.2, $\|F(v) - F(u) - J(x)(v-u)\| \leq \frac{\gamma}{2}\|v-u\|^2$

$$= \frac{\gamma}{2}\|v-u\|\,\|v-u\|$$

By triangles inequality, $\leq \frac{\gamma}{2}(\|v-x\| + \|u-x\|)\,\|v-u\|$

$$= \gamma \frac{\left(\|v-x\| + \|u-x\|\right)}{2} \|v - u\|.$$

Lemma 1.3.4: Let F, J satisfy the condition of Lemma 1.3.3 and let $J(x)^{-1}$ exists. Then there exists $\varepsilon > 0$, $0 < \alpha < B$ such that,

$$\alpha\| u - v\| \leq \| F(v) - F(u) \| \leq \beta \| v - u \|.$$

for all u, $v \in D$ for which max $\{\|v - x \|, \|u - x \|\} \leq \varepsilon$.

Proof:

$\|F(v) - F(u) \| \leq \| J(x) (v - u) \| + \|F(v) - F(u) - J(x) (v - u) \|$ by triangles in equality

by Lemma 1.3.3, $\leq [J(x) + \frac{\gamma}{2}(\| v - x\| + \| u - x\|)] \|v - u \|.$

$$\leq [\|J(x)\| \mid \gamma\varepsilon] \|v - u\|$$

for $\beta = \|J(x)\| + \gamma\varepsilon$.

$$\|F(v) - F(u)\| \leq \beta\|v - u\| \qquad (1.3.4.1)$$

Similarly

$$\|F(v) - F(u)\| \geq \|J(x) (v - u)\| - \|F(v) - F(u) - J(x) (v-u)\|.$$

By Lemma 1.3.3 $\geq [\left(\frac{1}{\|J(x)^{-1}\|}\right) - \frac{\gamma}{2}(\|v - x\| + \|u-x\|)] \|v - u \|$

$$\geq \left[\frac{1}{\|J(x)^{-1}\|} - \gamma\varepsilon\right]\|v-u\|$$

if $\varepsilon < \frac{1}{\|J(x)^{-1}\|\gamma}$ and lat $\alpha = \left(\frac{1}{\|J(x)^{-1}\|}\right) - \gamma\varepsilon > 0$

$$\alpha \|v - u\| \leq \|F(v) - F(u)\| \qquad (1.3.4.2)$$

from 1.3.4.1 and 1.3.4.2

$$\alpha\|v - u\| \leq \|F(v) - F(u)\| \leq \beta \|v - u\|.$$

2. Secant method

2.1 Secant method for solving non-linear equation in one variable

Let f: $\Re \rightarrow \Re$ and x_c be initial estimate of the zero's of f. There are cases in which f '(x) is not available. For example f(x) may be the result of some computational or experimental procedures. The Newton's method that uses f '(x) to solve f(x) = 0 must be modified to apply only values of f(x). In this case the model of f is a secant line that passes through the point (x_c, f(x_c)) and some othr near point ($x_c + h_c$, f($x_c + h_c$)).

Let the slope of the secant line be a_c,

then $a_c = \dfrac{f(x_c + h_c) - f(x_c)}{h_c}$

Definition 2.1.1: foe a very small h, a_c is called finite difference approximation to f '(x).

Let us modify a_c as follows:

$$a_c = \frac{f(x_-) - f(x_c)}{x_- - x_c}$$

Where x_- is the previous iterate, x_1 is approximated by the finite difference. Substituting a_c in place of f '(x_c) in formula (1.1.1) we will get:

$$x_+ = x_c - \frac{f(x_c)}{a_c}$$

Definition 2.1.2: The Newton's like (quasi-Newton) $x_+ = x_c - \dfrac{f(x_c)}{a_c}$ is said to be secant method

Example 2.1.1: Let $f(x) = x^2 - 1$, $h = 10^{-6}$ and $x_0 = 2$. Find the first three secant iterations.

Solution:

$$a_0 = \frac{f(x_0 + h) - f(x_0)}{h} = \frac{3.0000004 - 3}{10^{-6}} = 4.$$

$$x_1 = x_0 - \frac{f(x_0)}{a_0} = 2 - 0.75 = 1.25$$

$$a_1 = \frac{f(x_0+h)-f(x_0)}{x_0+h-x_1} = \frac{3.0000004-0.5625}{0.7500001} = 3.25$$

$$x_2 = x_1 - \frac{f(x_1)}{a_1} = 1.25 - 0.1730769 = 1.0769231$$

$$a_2 = \frac{f(x_1)-f(x_2)}{x_1-x_2} = \frac{0.5625-0.1597633}{0.1730769} = 2.3269234$$

$$x_3 = x_2 - \frac{f(x_2)}{a_2} = 1.076923 - \frac{0.1597633}{2.3269234} = 1.0082646$$

2.2 Broyden's method (secant method for systems of non-linear equations)

Broyden's method is the most applicable and successful secant approximation of the Jacobian matrix in Newton's method.

Let F: $\Re^n \rightarrow \Re^n$, $A_+ \in \Re^{n \times n}$ and the Jacobian matrix of F is not available, the affine model of F is:

$$M_+(x) = F(x_+) + A_+(x - x_+).$$

And this model should satisfy $M_+(x_c) = F(x_c)$, then

$$F(x_c) = F(x_+) + A_+(x_c - x_+).$$

$$A_+(x_c - x_+) = F(x_c) - F(x_+) \qquad (2.2.1)$$

Definition 2.2.1: The equation $A_+(x_c - x_+) = F(x_c) - F(x_+)$ is called scant equation.

Let $s_c = x_+ - x_c$ and $y_c = F(x_+) - F(x_c)$, then

$$A_+s_c = y_c. \qquad (2.2.2)$$

A_+ is not unique, for n > 1 there is an n (n-1) dimensional affine subspace of matrices satisfying equation (2.2.1). Out of these, A_+ is selected by minimizing the change in the offine model subject to $A_+s_c = y_c$. The change between the new and old affine models at any $x \in \Re^n$ is:-

$$M_+(x) - M_c(x) = F(x_+) + A_+(x-x_+) - F(x_c) - A_c(x-x_c)$$
$$= F(x_+) - F(x_c) - A_+(x_+-x_c) + (A_+-A_c)(x-x_c)$$

$$= (A_+ - A_c)(x - x_c) \qquad (2.2.3)$$

because $F(x_+) - F(x_c) = A_+(x_+ - x_c)$ from equation (2.2.1)

For $x \in \Re^n$, let $x - x_c = \alpha s_c + t$, where $\alpha \in \Re$, $t \in \Re^n$ such that $t^T s_c = 0$. Then equation (2.2.3) will be

$$M_+(x) - M_c(x) = (A_+ - A_c)(x - x_c) = \alpha(A_+ - A_c) s_c + (A_+ - A_c)t$$

It is possible to make $(A_+ - A_c)\, t = 0$, by choosing suitable A_+, for all t orthogonal s_c. That is $A_+ - A_c$ must be a matrix of rank one, in the form us_c^T, $u \in \Re^n$.

we know that

$$A_+ s_c = y_c.$$

$$A_+ s_c - A_c s_c = y_c - A_c s_c$$

$$(A_+ - A_c)s_c = y_c - A_c s_c.$$

$$u s_c^T s_c = y_c - A_s s_c.$$

$$u = \frac{y_c - A_c s_c}{s_c^T s_c} = \frac{A_+ - A_c}{s_c^T}$$

$$A_+ - A_c = \frac{(y_c - A_c s_c)}{s_c^T s_c} s_c^T$$

$$A_+ = A_c + \frac{(y_c - A_c s_c)}{s_c^T s_c} s_c^T \qquad (2.2.2.4)$$

Definition: 2.2.2: The equation $A_+ = A_c + \frac{(y_c - A_c s_c)}{s_c^T s_c}$ is called Broyden's up date or secant update.

Broydon's update is the minimum change to A_c in consistent with $A_+ s_c = y_c$ measured in frobenius norm.

Lemma 2.2.1 : Let $Q(y, s) = \{B \in \Re^{n \times n} / Bs = y\}$, $A \in \Re^{n \times n}$, y, s, v $\in \Re^n$ and $s \neq 0$, $V \neq 0$, then for any matrix norm $\|.\|$, $|||.|||$ such that

$$\|A.B\| \leq \|A\|\; |||B|||, \qquad (2.2.5)$$

$$\left\| \frac{vv^T}{v^T v} \right\| = 1 \tag{2.2.6}$$

The solution to

$$\min_{B \in Q(y,s)} \|B - A\|, \tag{2.2.7}$$

$$\text{is } A_+ = A + \frac{(y - As)s^T}{s^T s}. \tag{2.2.8}$$

when ||.|| is frobenius norm, then (2.2.8) solves(2.2.7) uniquely.

Proof: Let $B \in Q(y,s)$ then

$$\|A_+ - A\| = \left\| \frac{(y - As)s^T}{s^T s} \right\|$$

$$= \left\| \frac{(B - As)ss^T}{s^T s} \right\|, \text{ since } \quad y = Bs$$

$$\leq \| B - A \| \left\| \frac{ss^T}{s^T s} \right\|$$

$$= \| B - A \|, \text{ since } \left\| \frac{ss^T}{s^T s} \right\| = 1.$$

$$\min_{B \in Q(y,s)} \| B - A \| = \|A_+ - A \| .$$

Since the frobeninus norm is strictly convex (2.2.8) is a unique solution of (2.2.7).

Now we have completed our affine model by selecting A_+ as indicated above. The next iterate is the root of the model.

Algorithm 2.2.2: Boydon's method

(Secant method)

Given $F: \Re^n \to \Re^n$, $x_0 \in \Re^n$, $A_0 \in \Re^{n \times n}$.

Solve

$A_k s_k = -F(x_k)$

$x_{k+1} = x_k + s_k$

$y_k = F(x_{k+1}) - F(x_k)$

$$A_{k+1} = A_k + \frac{(y_k - A_k s_k)s_k^T}{s_k^T s_k}$$

for k = 0, 1, 2, …

Example 2.2.1: Let $F(x) = \begin{pmatrix} x_1 + x_2 - 2 \\ x_1^2 + x_2^2 - 4 \end{pmatrix}$ which has roots $(0, 2)^T$ and $(2, 0)^T$ let

$x_0 = (1, 3)$ Find the first two secant iterations.

Solution

For simplicity let $A_0 = J(x_0) = \begin{pmatrix} 1 & 1 \\ 2 & 6 \end{pmatrix}$

$$F(x_0) = \begin{pmatrix} 2 \\ 6 \end{pmatrix}$$

$$A_0S_0 = -F(x_0)$$

$$s_0 = A_0^{-1} F(x_0) = \begin{pmatrix} \frac{-3}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{-1}{4} \end{pmatrix} \begin{pmatrix} 2 \\ 6 \end{pmatrix} = \begin{pmatrix} \frac{-3}{2} \\ \frac{-1}{2} \end{pmatrix}$$

$$x_1 = x_0 + s_0 = \begin{pmatrix} 1 \\ 3 \end{pmatrix} + \begin{pmatrix} \frac{-3}{2} \\ \frac{-1}{2} \end{pmatrix} = \begin{pmatrix} \frac{-1}{2} \\ \frac{5}{2} \end{pmatrix}$$

$$A_1 = A_0 + \frac{(y_0 - A_0 s_0)s_0^T}{s_0^T s_0}$$

$$= \begin{pmatrix} 1 & 1 \\ 2 & 6 \end{pmatrix} + \frac{\left(\begin{pmatrix} -2 \\ \frac{-14}{4} \end{pmatrix} - \begin{pmatrix} -2 \\ -6 \end{pmatrix}\right)\left(\frac{-3}{2}, \frac{-1}{2}\right)}{\left(\frac{-3}{2}, \frac{-1}{2}\right)\begin{pmatrix} \frac{-3}{2} \\ \frac{-1}{2} \end{pmatrix}}$$

$$= \begin{pmatrix} 1 & 1 \\ \frac{1}{2} & \frac{11}{2} \end{pmatrix}$$

$$S_1 = -A_1^{-1}F(x_1) = \begin{pmatrix} \frac{-11}{10} & \frac{1}{5} \\ \frac{1}{10} & \frac{-1}{5} \end{pmatrix}\begin{pmatrix} 0 \\ \frac{10}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$$

$$x_2 = x_1 + s_1 = \begin{pmatrix} \frac{-1}{2} \\ \frac{5}{2} \end{pmatrix} + \begin{pmatrix} \frac{1}{2} \\ \frac{-1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

Example 2.2.2: Let $F(x) = \begin{pmatrix} x_1 + x_2 - 3 \\ x_1^2 + x_2^2 - 9 \end{pmatrix}$ which has roots $(0, 3)^T$ and $(3, 0)^T$, let $x_0 = (1, 5)^T$. Find the first two secant iteration.

Solution

$$\text{Let } A_0 = J(x_0) = \begin{pmatrix} 1 & 1 \\ 2 & 10 \end{pmatrix}$$

$$F(x_0) = \begin{pmatrix} 3 \\ 17 \end{pmatrix}$$

$$A_0 s_0 = -F(x_0)$$

$$S_0 = -A_0^{-1}F(x_0) = \begin{pmatrix} -1.625 \\ -1.375 \end{pmatrix}$$

$$x_1 = x_0 + s_0 = \begin{pmatrix} -0.625 \\ 3.625 \end{pmatrix}$$

$$F(x_1) = \begin{pmatrix} 0 \\ 4.53125 \end{pmatrix}$$

$$A_1 = A_0 + \frac{(y_0 - A_0 s_0)}{s_0^T s_0} = \begin{pmatrix} 1 & 1 \\ 0.375 & 8.625 \end{pmatrix}$$

$$S_1 = -A_1^{-1}F(x_1) \cong \begin{pmatrix} 0.549 \\ -0.549 \end{pmatrix}$$

$$x_2 = x_1 + s_1 \cong \begin{pmatrix} -0.076 \\ 3.076 \end{pmatrix}$$

Local Convergence analysis of Broydens method

3.1 Local q- linearly convergence of the Broyden's itrates (x_k) to x_*

Let x_*, $x_k \in \Re$ and $\lim_{k \to \infty} x_k = x_*$ for k = 0, 1, 2, …,

Def. 3.1.1: The sequence of iterates (x_k) is said to be convergent locally q – linearly to x_*, if there exists c $\in$[0, 1) and $k_o \in N$ such that

$$|x_{k+1} - x_*| \leq c\,|x_k - x_*|,\ \forall k \geq k_0.$$

Lemma 3.1.1: Let x_0, x_k, $x_* \in \Re^n$ $A_0 \in \Re^{n \times n}$, F: $\Re^n \to \Re^n$ be continuously differentiable. Further more let $l_k = x_k - x_*$, x_0 and A_0 are sufficiently close to x_* and $J(x_0)$ respectively and $J(x_*)$ be nonsingular. If e_k converges locally q-linearly to zero, then (x_k) converges locally q-linearly to x_*.

Proof: Suppose e_k converges locally q-linearly to zero

From Broyden's algorithm,

$$A_k s_k = -F(x_k)$$

$$s_k = A_k^{-1} F(x_k)$$

$$x_{k+1} - x_k = A_k^{-1} F(x_k)$$

$$x_{k+1} = x_k - A_k^{-1} F(x_k)$$

Subtract x_* from both sides, it follows,

$$x_{k-1} - x_* = x_k - x_* - A_k^{-1}(F(x_k) - F(x_*))$$

Multiply both sides by A_k we have

$$A_k(x_{k+1} - x_*) = A_k(x_k - x_*) - F(x_k) + F(x_*)$$

i.e. $A_k\, e_{k+1} = Ae_k - F(x_k) + F(x_*)$ 3.1.1.1

Add and subtract $J(x_*)\, e_k$ to the right side of (3.1.1.1)

$$A_k e_{k+1} = A_k e_k - F(x_k) + F(x_*) + J(x_*)e_k - J(x_*)e_k.$$

$$= [-F(x_k) + F(x_*) + J(x_*)e_k] + [A_k - J(x_*)]\, e_k \qquad 3.1.1.2$$

From the hypothesis.

$$e_k \to 0$$

$$x_k - x_* \to 0$$

$$x_k \to x_*$$

$$F(x_k) \to F(x_*)$$

$$-F(x_k) + F(x_*) \to 0 \text{ and } J(x_*)e_k \to 0$$

From (3.1.1.2)

$$A_k e_{k+1} = [-F(x_k) + F(x_*) + J(x_*)e_k\,] + (A_k - J(x_*))\, e_k.$$

$$e_{k+1} = A_k^{-1}[\,-F(x_k) + F(x_*) + J(x_*)e_k] + A_k^{-1}\,(A_k - J(x_*)]e_k.$$

$$\|e_{k+1}\| \le \|A_k^{-1}\|\ [\,\|-F(x_k) + \underset{\to 0}{F}(x_*) + J(x_*)e_k\| + \|\underbrace{A_k - J(x_*)}_{c}\|\,\|e_k\|]$$

$$\|e_{k+1}\,| \le c\|e_k\|.$$

$$\|x_{k+1} - x_*\| \le c\,\|x_k - x_*\|.$$

Then (x_k) converges locally q-linearly to x_*. So the main idea of locally q-linearly convergence of the Broyden's iterates (x_k) will be an analysis of $(A_k - J(x_*))\, e_k$. There for we will prove that (e_k) converges locally q- linearly to zero by showing that $\{\|A_k - J(x_*)\|\}$ remains bounded.

Lemma 3.1.2: Let $D \subseteq \Re^n$ be an open convex set,

$F: \Re^n \to \Re^n$ be a function, $J(x) \in Lip_\gamma(D)$, $x_c, x_* \in D$

with $x_c \ne x_*$, $A_c \in \Re^{n \times n}$ and A_+ is defined by

$$A_+ = A_c + \frac{(y_c - A_c s_c)s_c^T}{s_c^T s_c},$$

then we get

a) $$\|A_+ - J(x_+)\| \le \|A_c - J(x_c)\| + \frac{3\gamma}{2}\|x_2 - x_c\|, \qquad 3.1.2.1$$

where we use either the frobenius norm or the ℓ_2 norm.

b) If J(x) obeys the weaker Lipschitz condition,

$$\|J(x) - J(x_*)\| \le \gamma \| x - x_* \|, \ \forall x \in D, \text{ then,}$$

$$\|A_+ - J(x_*)\| \le \|A_c - J(x_*)\| + \frac{\gamma}{2}(\|x_+ - x\|_2 + \|x_c - x_*\|_2) \qquad 3.1.2.2$$

Proof:

a) Let $J_+ = J(x_+)$, $J_c = J(x_c)$, $y_c = F(x_+) - F(x_c)$

$$A_+ = A_c + \frac{(y_c - A_c s_c)s_c^T}{s_c^T s_c} \text{ be given}$$

subtract J_+ from both sides then

$$A_+ - J_+ = A_c - J_+ + \frac{(y_c - A_c s_c)s_c^T}{s_c^T s_c}$$

Add and subtract $\frac{J_+ - s_c s_c^T}{s_c^T s_c}$ on the right side of the equation, then

$$A_+ - J_+ = A_c - J_+ + \frac{(J_+ - A_c)s_c s_c^T}{s_c^T s_c} + \frac{(y_c - A_c s_c)s_c^T}{s_c^T s_c}$$

i.e. $$A_+ - J_+ = (A_c - J_+)\left(I - \frac{s_c s_c^T}{s_c^T s_c}\right) + \frac{(y_c - A_c s_c)s_c^T}{s_c^T s_c}$$

using the frobenius norm or ℓ_2 norm we have

$$\|A_+ - J_+\| \le \|A_c - J_+\| \left\| I - \frac{s_c s_c^T}{s_c^T s_c} \right\| + \frac{\|y_c - J_+ s_c\|}{\|s_c\|}$$

Since $I - \frac{s_c s_c^T}{s_c^T s_c}$ is an Euclidean matrix orthogonal projection, we have

$$\left\| I - \frac{s_c s_c^T}{s_c^T s_c} \right\| = 1$$

And hence

$$\|A_+ - J_+\| = \|A_c - J_+\| + \frac{\|y_c - J_+ s_c\|}{\|s_c\|} \qquad 3.1.2.3$$

by Lemma 1.3.3 we have,

$$\|y_c - J_+ s_c\| \le \frac{\gamma}{2} (\|x_+ - x_+\| + \|x_c - x_+\|) \|s_c\|$$

$$\Rightarrow \|y_c - J_+ s_c\| \le \frac{\gamma}{2} \|x_c - x_+\| \|s_c\|. \qquad 3.1.2.4$$

Substitute (3.1.2.4) in (3.1.2.3) we get,

$$\|A_+ - J_+\| \le \|A_c - J_+\| + \frac{\gamma}{2} \|x_c - x_+\|$$

$$= \|A_c - J_c + J_c - J_+\| + \frac{\gamma}{2} \|x_c - x_+\|$$

by triangle inequality, $\le \|A_c - J_c\| + \|J_c - J_+\| + \frac{\gamma}{2} \|x_c - x_+\|$

since $J(x) \in Lip_\gamma(D)$, $\le \|A_c - J_c\| + \gamma \|x_c - x_+\| + \frac{\gamma}{2} \|x_c - x_+\|$.

b) Let $J_* = J(x_*)$ and

$A_+ = A_c + (y_c - A_c s_c^T)$ be given

subtract J_* from both solution $A_+ - J_* = A_c - J_* + + \frac{(y_c - A_c s_c) s_c^T}{s_c^T s_c}$

Add and subtract, $\frac{J_* s_c s_c^T}{s_c^T s_c}$ on the right side we get,

$$A_+ - J_* = A_c - J_* + \frac{(J_* s_c - A_c s_c) s_c^T}{s_c^T s_c} + \frac{(y_c - J_* s_c) s_c^T}{s_c^T s_c}$$

$$A_+ - J_* = (A_c - J_*) \left(I - \frac{s_c s_c^T}{s_c^T s_c} \right) + \frac{(y_c - J_* s_c) s_c^T}{s_c^T s_c}$$

Using the frobenius norm or ℓ_2 norm,

$$\|A_+ - J_*\| = \|A_c - J_*\| \left\| I - \frac{s_c s_c^T}{s_c^T s_c} \right\| + \frac{\|y_c - J_* s_c\|}{\|s_c\|}$$

Since $I - \frac{s_c s_c^T}{s_c^T s_c}$ is an Euclidean matrix orthogonal projection

$$\left\| I - \frac{s_c s_c^T}{s_c^T s_c} \right\| = 1$$

And hence

$$\|A_+ - J_*\| = \|A_c - J_*\| + \frac{\|y_c - J_* s_c\|}{\|s_c\|} \qquad 3.1.2.5$$

By Lemma 1.3.3

$$\|_c - J_* s_c\| \leq \frac{\gamma}{2}(\|x_+ - x_*\| + \|x_c - x_*\|) \|s_c\| \qquad 3.1.2.6$$

Substitute (3.1.2.6) in (3.1.2.5) we get,

$$\|A_+ - J_* \| \leq \|A_c - J_*\| + \frac{\gamma}{2}(\|x_+ - x_*\| + \|x_c - x_*\|) \|).$$

Theorem 3.1.3: Let $D \subseteq \Re^n$ an open convex set, $F:\Re^n \to \Re^n$ be continuously differentiable function, $x_* \in \Re^n$ such that $F(x_*) = 0$. Moreover Let r, $\beta \in \Re_+$ such that $N(x_*, r) = \{x \in \Re^n: \|x - x_*\| < r \}$ $\subseteq D$, $J(x) \in Lip_\gamma (N(x_*, r))$ and $J^{-1}(x_*)$ exists with $\|J^{-1}(x_*)\| \leq \beta$. If there exists ε, $\delta > 0$ such that $\|x_0 - x\| \leq \varepsilon$, $\|A_0 - J(x_*)\| \leq \delta$ and (A_k) satisfies (3.1.1.2), then (x_k) converges locally q-linearly to x_*.

Proof: Let $\|.\|$ be the ℓ_2 matrix norm, $e_k = x_k - x_*$ and

$J_* = J(x_*)$, we choose ε and δ such that,

$$6\beta\delta \leq 1 \qquad 3.1.3.1$$

$$3\gamma\varepsilon \leq 2\delta$$

$$\frac{3}{2}\gamma\varepsilon \leq \delta \qquad 3.1.3.2$$

To prove the local q-linearly convergence, it suffices to show that the following two inequalities are true.

$$\|A_k - J_*\| \leq (2 - 2^k)\, \delta \qquad 3.1.3.3$$

$$\|e_{k+1}\| \leq \frac{\|e_k\|}{2} \qquad 3.1.3.4$$

for $k = 0, 1, 2, \ldots,$

Now we prove (3.1.3.3) by induction.

For k = 0 the proof is trivial.

Assume (3.1.3.3) and (3.1.3.4) hold for k = 0, 1, 2, …, i – 1

for k = i,

From (3.1.2.2) we have,

$$\|A_k - J_*\| \le \|A_{k-1} - J_*\| + \frac{\gamma}{2}(\|e_k\| + \|e_k\text{-}1\|)$$

by induction hypothesis,

$$\|A_i - J_*\| \le (2 - 2^{-(i-1)})\ \delta + \frac{\gamma}{2}\left(\frac{\|e_{i-1}\|}{2} + \|e_{i-1}\|\right)$$

$$\|A_j - J_*\| \le (2 - 2^{-(i-1)})\ \delta + \frac{3\gamma}{4}\|e_{i-1}\| \qquad 3.1.3.5$$

From (3.1.3.4) we have,

$$\|e_{i-1}\| \le \frac{\|e_{i-2}\|}{2} \le \frac{\|e_{i-3}\|}{4} \le \ldots \le 2^{-(i-1)}\|e_0\|$$

Since $\|e_0\| < \varepsilon$ we have,

$$\|e_{i-1}\| \le 2^{-(i-1)}\|e_0\| \le 2^{-(i-1)}\varepsilon \qquad 3.1.3.6$$

Substituting (3.1.2.6) in (3.1.2.5) we get,

$$\|A_j - J_*\| \le (2 - 2^{-(i-1)})\ \delta + \frac{3\gamma}{4}\varepsilon 2^{-(i-1)}$$

Since $\frac{3}{2}\gamma\varepsilon < \delta, \le (2 - 2^{-(i-1)}) + 2^{-i})\delta$

$$= (2 - 2^{-1})\ \delta.$$

i.e. $\|A_j - J_*\| \le (2 - 2^{-(i-1)})\ \delta$

$\|A_k - J_*\| \le (2 - 2^{-k})\ \delta.$

To prove (3.1.3.4) we have to show that A_i is invertible, so that the iterations well defined.

We have that $\|J_*^{-1}\| \le \beta$ and $\|A_i - J_*\| \le (2 - 2^{-1})\ \delta$, then

$$\|J_*^{-1}(A_i - J_*)\| \le \|J_*^{-1}\|\ \|A_i - J_*\| \le \beta\ (2 - 2^{-1})\delta,$$

Since $2 - 2^{-1} \leq 2$, $\|J^{-1}(A_i - J_*)\| \leq 2\beta\delta$

From (3.1.2.1) we know $6\beta\delta \leq 1 \Rightarrow \beta\delta \leq \frac{1}{6}$, then

$$\|J^{-1}(A_i - J_*)\| \leq \frac{1}{3}.$$

By Theorem 1.3.1 A_i is non singular and

$$\| A_i^{-1} \| \leq \frac{\|J_*^{-1}\|}{1-\|J_*^{-1}(A_i - J_*)\|} \leq \frac{\beta}{1-\frac{1}{3}} = \frac{3\beta}{2} \qquad 3.1.3.7$$

Thus x_{i+1} is well defined and we have,

$$A_k e_{k+1} = [-F(x_k) + F(x_*) + J_* e_k] + (A_k - J(x_*))e_k,$$

i.e $e_{k+1} = A_k^{-1} [[-F(x_k\_ + F(x_*) + J_* e_k] + (A_k - J(x_*))e_k]$

$$\|e_{k+1}\| \leq \| A_k^{-1} \| [\|-F(x_k) + F(x_*) + J_* e_k \| + \|A_k - J(x_*) \| \|e_k\|] \qquad 3.1.3.8$$

By Lemma 1.3.2,

$$\| -F(x_k) + F(x_*) + J_* e_k \| \leq \frac{\gamma}{2} \|e_k\|^2 \qquad 3.1.3.9$$

Substituting (3.1.3.3), (3.1.3.7) and (3.1.3.9) in (3.1.3.8) we get,

$$\|e_{i+1}\| \leq \frac{3\beta}{2} [\frac{\gamma}{2} \|e_i\| + (2 - 2^{-i}) \delta] \|e_i\| \qquad 3.1.3.10$$

From (3.1.3.2), (3.1.3.4) and from $\|e_0\| \leq \varepsilon$ we have,

$$\frac{\gamma}{2} \|e_i\| \leq 2^{-(i+1)} \gamma_\varepsilon \leq \frac{2^{-i}\delta}{3} \qquad 3.1.3.11$$

Substituting (3.1.3.11) in (3.1.3.10) we get,

$$\|e_{i+1}\| \leq \frac{3\beta}{2} [\frac{1}{3} 2^{-i} + 2\text{-}2^{-i}]\delta \|e_i\|$$

$$= 3\beta \left[1 - \frac{2^{-i}}{2} + \frac{2^{-i}}{6}\right]\delta \|e_i\|$$

$$= e\beta [1+ \left(\frac{1}{2} + \frac{1}{6}\right) 2^{-i}] \delta\|e_i\|$$

$$= 3\beta\ [1 - \frac{-2}{3} 2^{-i}]\delta\ \|e_i\|$$

Since $1 - \frac{2}{3}2^{-i} < 1$, $\leq 3\beta\delta\|e_i\|$

Since $6\beta\delta < 1$, $\leq \frac{\|e_i\|}{2}$.

i.e. $\|e_{i+1}\| \leq \frac{\|e_i\|}{2}$

Since $\|e_{i+1}\| \leq 2^{-(i+1)}\ \|e_0\| \leq 2^{-(i+1)}\varepsilon$ or in general

$\|e_k\| \leq 2^{-k}\|e_0\| \leq 2^{-k}\varepsilon$ for $\varepsilon > 0$ arbitrary implies that (e_k) converges to zero and $\|A_k - J_*\|$ is bounded above i.e $\|A_k - J_*\| \leq (2 - 2^{-k})\delta$.

Therefore (e_k) converges locally q-linearly to zero and hence (x_k) converge locally q-linearly to x_*.

3.2 Local q-super linearly convergence of the Broyden's itrate (x_k) to x_*

Let $x_*, x_k \in \Re$ for $k = 0, 1, 2, \ldots$, and $\lim_{k\to\infty} x_k = x_*$

Def. 3.2.1: The sequence (x_k) is said to converge locally q-super linearly to x_*, if for some sequence (c_k) that converges to zero such that,

$$|x_{k+1} - x_*| - c_k\ |x_k - x_*|.$$

The main idea of this section is to show that the Broyden's iterates (x_k) converges locally q-super linearly to x_* under certain assumptions, i.e. if x_0 and A_0 are sufficiently and $J(x_0)$ respectively and $J(x_*)$ is non-singular, then (x_k) converges locally q-super linearly to x_*. But if we can prove $\lim_{k\to\infty} \frac{\|(A_k - J(x_*))e_k\|}{\|e_k\|} = 0$, then (x_k) converges locally q-super linearly to x_*.

First lat us consider the following lemmas.

Lemma 3.2.1: Let x_*, $x_k \in \Re$ for k = 0, 1, 2, …,

If (x_k) converges locally q-super linearly to x_*, then for any norm $\| . \|$ we have $\lim_{k\to\infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x_*\|} = 1$.

Proof: Suppose (x_k) converges locally q-super linearly to x_*.

Let $s_k = x_{k+1} - x_k$ and $e_k = x_k - x_*$.

clearly, if $\lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|} = 0$, then $\lim_{k\to\infty} \frac{\|s_k\|}{\|e_k\|} = 1$

$$0 \le \left| \lim_{k\to\infty} \frac{\|s_k\|}{\|e_k\|} - 1 \right| = \lim_{k\to\infty} \frac{\|s_k\| - \|e_k\|}{\|e_k\|}$$

$$\le \lim_{k\to\infty} \frac{\|s_k + e_k\|}{\|e_k\|}$$

Since $\|s_k + e_k\| = \|x_{k+1} - x_k + x_k - x_* \| = \|e_{k+1}\|$ we have,

$$\lim_{k\to\infty} \left| \frac{\|s_k\|}{\|e_k\|} - 1 \right| \le \lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|}$$

definition of q-super linearly convergence, = 0

Then

$$\lim_{k\to\infty} \frac{\|s_k\|}{\|e_k\|} = 1 \text{ or } \lim_{k\to\infty} \frac{\|e_{k+1} - x_k\|}{\|x_k - x_*\|} = 1.$$

Lemma 3.2.1 help us to replace e_k by s_k in the sufficient condition for the q-super linearly convergence of the Broyobn's iterates (x_k), i.e.

$$\lim_{k\to\infty} \frac{\|(A - J(x_*))e_k\|}{\|e_k\|} = 0$$

Lemma 3.2.2: Let $s \in .\Re^n$ be non zero, $E \in \Re^{n \times n}$ and

Let $\| . \|$ denote the ℓ_2 norm, then

$$\left\| E\left(I - \frac{ss^T}{s^T s} \right) \right\|_F = \left(\|E\|_F^2 - \left(\frac{\|Es\|}{\|s\|} \right)^2 \right)^{\frac{1}{2}} \qquad 3.2.2.1$$

$$\leq \|E\|_F - \frac{1}{2\|E\|_F}\left(\frac{\|Es\|}{s}\right). \qquad 3.2.2.2$$

Proof: We know that $I - \frac{ss^T}{s^T s}$ is an Euclidean projector and $\frac{ss^T}{s^T s}$ is again a projector.

By Pythagoras theorem,

$$\|E\|_F^2 = \left\|E\frac{ss^T}{s^T s}\right\| + \left\|E\left(I - \frac{ss^T}{s^T s}\right)\right\|_F^2$$

Since $\left\|E\frac{ss^T}{s^T s}\right\| = \frac{\|Es\|}{\|s\|}$, then

$$\|E\|_F^2 = \left(\frac{\|Es\|}{\|s\|}\right)^2 + \left\|E\left(I - \frac{ss^T}{s^T s}\right)\right\|_F^2$$

And hence, $\left\|E\left(I - \frac{ss^T}{s^T s}\right)\right\|_F = \left(\|E\|_F^2 - \left(\frac{\|Es\|}{\|s\|}\right)^2\right)^{1/2}$

Sine for $\alpha > |\beta| \geq 0$, $\left(\alpha^2 - \beta^2\right)^{1/2} \leq \alpha - \frac{\beta^2}{2\alpha}$ we have

$$\left(\|E\|_F^2 - \left(\frac{\|Es\|}{\|s\|}\right)^2\right)^{\frac{1}{2}} \leq \|E\|_F - \frac{1}{2\|E\|_F}\left(\frac{\|Es\|}{\|s\|}\right)^2.$$

Theorem 3.2.3: (Dennis – More 1974)

Let $D \subseteq \Re^n$ be an open convex set, $F:\Re^n \to \Re^n$ be a function $J(x) \in Lip_\gamma(D)$, $x_* \in D$ and $J(x_*)$ be nonsingular.

Furthermore, let (A_k) be a sequence of nonsingular matrices $A_k \in \Re^{n\times n}$, $x_0 \in D$ and (x_k) be sequence of points generated by

$$x_{k+1} = x_k - A_x^{-1} F(x_k), K \in N_0, \qquad 3.2.3.1$$

If
1. $x_k \in D$ for all $k \in N$,
2. $x_k \neq x_*$ for all $k \in N$,
3. $\lim_{k\to\infty} x_k = x_*$,

Then (x_k) converges locally q-super linearly to x_* with respect to some norm and

$F(x_*) = 0$ if and only if $\lim_{k\to\infty} \frac{\|(A_k - J(x_*))s_k\|}{\|s_k\|} = 0$ where $s_k = x_{k+1} - x_k$. 3.2.3.2

Proof:

a) Let $J_* = J(x_*)$, $e_k = x_k - x_*$

suppose $\lim_{k\to\infty} \frac{\|(A_k - J(x_*))\|}{\|s_k\|} = 0$

We have to show that $F(x_*) = 0$ and that (x_k) converges locally q-super linearly to x_*.

We have $x_{k+1} = x_k - A_k^{-1} F(x_k)$

$$x_{k+1} - x_k = - A_k^{-1} F(x_k).$$

$$s_k = - A_k^{-1} F(x_k).$$

$$A_k s_k = - F(x_k)$$

$$A_k s_k + F(x_k) = 0$$

Add and subtract $J_* s_k$ on the left side, then

$$(A_k - J_*) s_k + F(x_k) + J_* s_k = 0$$

Add $- F(x_{k+1})$ on both side we have,

$$-F(x_{k+1}) = (A_k - J_*) s_k + [-F(x_{k+1}) + F(x_k) + J_* s_k] \qquad 3.2.3.3$$

$$\|F(x_{k+1})\| \leq \|(A_k - J_*)s_k\| + \|-F(x_{k+1}) + F(x_k) + J_* s_k\|.$$

$$\frac{\|F(x_{k+1})\|}{\|s_k\|} \leq \frac{\|(A_k - J_*)s_k\|}{\|s_k\|} + \frac{\|-F(x_{k+1}) + F(x_k) + J_* s_k\|}{\|s_k\|}$$

By Lemma 1.3.3 we have,

$$\frac{\|F(x_{k+1})\|}{\|s_k\|} \leq \frac{\gamma}{2} \frac{(\|x_{k+1} - x_*\| + \| x_k - x_*\|)}{\|s_k\|} \|x_{k+1} - x_k\|$$

$$+ \frac{\|(A_k - J_*)s_k\|}{\|s_k\|}$$

$$= \frac{\|(A_k - J_*)s_k\|}{\|s_k\|} + \frac{\gamma}{2} \left(\|e_k\| + \|e_{k+1}\| \right) \qquad 3.2.3.4$$

using our assumption $\lim_{k\to\infty} \frac{\|(A_k - J(x_*))\|}{\|s_k\|} = 0$, $\lim_{k\to\infty} \|e_k\| = 0$ and (3.2.3.4) we get,

$$\lim_{k\to\infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} = 0 \qquad 3.2.3.5$$

Since $\lim_{k\to\infty} \|s_k\| = 0$, this implies

$$F(x_*) = \lim_{k\to\infty} F(x_k) = 0.$$

For x_{k+1}, $x_* \in D$, $\max \{ \|x_{k+1} - x\|, \|x_* - x\| \} \le \varepsilon$, there exist $\alpha > 0$, $k_0 \ge 0$, by Lemma 1.3.4 and $F(x_*) = 0$ we have,

$$\|F(x_{k+1})\| = \|F(x_{k+1}) - F(x_*)\| \ge \alpha \|x_{k+1} - x_*\|, \forall k \ge k_0.$$

$$\|F(x_{k+1}) - F(x_*)\| \ge \alpha \|e_{k+1}\|, \forall k \ge k_0 \qquad 3.2.3.6$$

From (3.2.3.5) and (3.2.3.6) we have,

$$0 = \lim_{k\to\infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} \ge \alpha \frac{\|e_{k+1}\|}{\|s_k\|} \ge \lim_{k\to\infty} \frac{\alpha\|e_{k+1}\|}{\|e_k\| + \|e_{k+1}\|}$$

$$= \lim_{k\to\infty} \frac{\alpha r_k}{1 + r_k} \text{ where } r_k = \frac{\|e_{k+1}\|}{\|e_k\|}$$

i.e $\quad \lim_{k\to\infty} \frac{\alpha r_k}{1 + r_k} = 0$ and hence $\lim_{k\to\infty} r_k = 0$

$$\lim_{k\to\infty} r_k = \lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|} = 0.$$

(x_k) converges locally q-super linearly to x_*.

b) Suppose the sequence (x_k) converges locally q-super linearly to x_*, i.e. $\lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|} = 0$ and $F(x_*) = 0$.

We have to show $\lim_{k\to\infty} \frac{\|(A_k - J(x_*))s_k\|}{\|s_k\|} = 0$

By Lemma 1.3.4 there exist $\beta > 0$, $k_0 \in N$ such that,

$\beta\|e_{k+1}\| \ge \|F(x_{k+1})\|, \qquad \forall_k \ge k_0$

Divide both sides by β $\|e_k\|$, then

$$\frac{\|e_{k+1}\|}{\|e_k\|} \geq \frac{\|F(x_{k+1})\|}{\beta\|e_k\|}$$

$$0 = \lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|} \geq \lim_{k\to\infty} \frac{\|F(x_{k+1})\|}{\beta\|e_k\|} = \frac{1}{\beta} \lim_{k\to\infty} \frac{\|F(x_{k+1})\|}{\|e_k\|}$$

$$= \frac{1}{\beta} \lim_{k\to\infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} \cdot \frac{\|s_k\|}{\|e_k\|} \qquad 3.2.3.7$$

By Lemma 3.2.1 $\dfrac{\|s_k\|}{\|e_k\|} = 1$, then we have

$$0 \geq \frac{1}{\beta} \lim_{k\to\infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} . 1$$

$$= \lim_{k\to\infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} \geq 0, \text{ for } \beta > 0$$

And hence $0 \geq \lim_{k\to\infty} \dfrac{\|F(x_{k+1})\|}{\|s_k\|} \geq 0$ i.e.

$$\lim_{k\to\infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} = 0 \qquad 3.2.3.8$$

From Lemma 1.3.3 and (3.2.3.3),

$$\frac{\|(A_k - J_*)s_k\|}{\|s_k\|} \leq \frac{\|F(x_{k+1})\|}{\|s_k\|} + \frac{\|-F(x_{k+1}) + F(x_k) + J_* s_k\|}{\|s_k\|}$$

$$\leq \frac{\|F(x_{k+1})\|}{\|s_k\|} + \frac{\gamma}{2} \left(\|e_k\| + \|e_{k+1}\| \right)$$

From (3.2.3.8) and $\lim_{k\to\infty} \|e_k\| = 0$

$$0 \leq \lim_{k\to\infty} \frac{\|(A_k - J_*)s_k\|}{\|s_k\|} \leq \lim_{k\to\infty} \frac{\|F(x_{k+1})\|}{\|s_k\|} + \frac{\gamma}{2} \lim_{k\to\infty} \left(\|e_k\| + \|e_{k+1}\| \right) = 0$$

i.e $\quad \lim_{k\to\infty} \dfrac{\|(A_k - J_*)s_k\|}{\|s_k\|} = 0.$

Since $J(x) \in Lip_\gamma$ (D), it is easy to show that Theorem 3.2.3 remains true if (3.2.3.2) is replaced by $\lim_{k\to\infty} \frac{\|(A_k - J_*(x_k))s_k\|}{\|s_k\|} = 0$ 3.2.3.9

Since $s_k = A_k^{-1} F(x_k)$, (3.2.3.9) is equivalent to

$$\lim_{k\to\infty} \frac{\|J(x)(s_k^N - s_k\|}{\|s_k\|} = 0,$$

where $s_k^N = -J(x_k)^{-1} F(x^k)$ which is the Newton step.

Therefore the necessary and sufficient condition for the locally q-linearly convergence of the secant method is that the secant steps converge in magnitude and direction to the Newton steps from the same point.

In Theorem 3.2.3 the sufficient condition for (x_k) converge locally q-super linearly to x_* is

$$\lim_{k\to\infty} \frac{\|(A_k - J_*)s_k\|}{\|s_k\|} = 0$$

Proof: Let $E_k = A_k - J_*$, Then $\lim_{k\to\infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0$ (3.2.3.10)

By (3.1.1.3),

$$\|E_{k+1}\| = \|E_k \left(I - \frac{s_k s_k^T}{s_k^T s_k} \right)\| + \frac{\|(y_c - J_* s_k) s_k^T\|}{s_k^T s_k}$$

3.2.3.11

From the proof of Lemma 3.1.1 we have,

$$\left\| \frac{(y_k - J_* s_k) s_k^T}{s_k^T s_k} \right\| \le \frac{\gamma}{2} (\|e_k\| + \|e_{k+1}\|) \qquad 3.2.3.12$$

From (3.2.3.12) and $\|e_{k+1}\| \le \frac{\|e_k\|}{2}$

$$\left\| \frac{(y_k - J_* s_k) s_k^T}{s_k^T s_k} \right\| \leq \frac{\gamma}{2} \cdot \frac{3}{2} \ \|e_k\| = \frac{3\gamma}{4} \ \|e_k\|$$

By lemma 3.2.2,

$$\|E_{k+1}\| \leq \|E_k\| - \frac{\|E_k s_k\|^2}{2\|E_k\|\|s_k\|^2} + \frac{3\gamma}{4} \|e_k\|$$

$$\frac{\|E_k s_k\|^2}{\|s_k\|^2} \leq 2\|E_k\| \left[\|E_k\| - E_{k+1} + \frac{3\gamma}{4} \| e_k \| \right] \qquad 3.2.3.13$$

In the proof of theorem 3.1.2 we hade, $\|E_k\| \| E_k\| \leq 2\delta\leq, \forall k \geq 0$ and $\sum_{k=0}^{\infty} \|e_k\| < 2\varepsilon$.

Thus (3.2.3.13)

$$\frac{\|E_k s_k\|^2}{\|s_k\|^2} \leq 4\delta \left[\|E_k\| - \| E_{k+1} \| + \frac{3\gamma}{4} \| e_k \| \right] \qquad 3.2.3.14$$

Take the summation of the left and right side of (3.2.3.14) for k = 0, 1, 2, …i,

$$\sum_{k=0}^{i} \frac{\|E_k s_k\|^2}{\|s_k\|^2} \leq \left[\|E_0\| - \| E_{k+1} \| + \frac{3\gamma}{4} \sum_{k=0}^{i} \| e_k \| \right]$$

$$\leq 4\delta \left[\|E_0\| + \frac{3\gamma\varepsilon}{2} \right]$$

$$\leq 4\delta \left[\delta + \frac{3\gamma\varepsilon}{2} \right] \qquad 3.2.3.15$$

Since (3.2.3.15) is true for any $i \geq 0$,

$$\sum_{k=0}^{\infty} \frac{\|E_k s_k\|}{\|s_k\|^2} < \infty$$

$$\Rightarrow \qquad \lim_{k\to\infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0.$$

Example:

$$\text{Let } F(x) = \begin{pmatrix} x_1^2 + x_2^2 - 2 \\ e^{x_1 - 1} + x_2^3 - 2 \end{pmatrix}$$

F(x) has a root $x_* = (1, 1)^T$

The sequence of points generated by Broyden's method and Newton's method from $x_0 = (1.5, 2)^T$ with $A_0 = J(x_0)$ are given below

Broyodens method			**Newton's method**	
1.5	2.0	x_0	1.5	2.0
0.8060692	1.457948	x_1	0.8060692	1.457948
0.7410741	1.277067	x_2	0.8901193	1.145571
0.8022786	1.159900	x_3	0.9915891	1.021054
0.9294701	1.070406	x_4	0.9997085	1.000535
1.00403	1.009609	x_5	0.999999829	1.000000357
1.003084	0.9992213	x_6	0.9999999999992	1.0000000000002
1.000543	0.9996855	x_7	1.0	1.0
0.9999818	1.0000000000389	x_8		
0.9999999885	0.99999999544	x_9		
0.99999999999474	0.9999999999998	x_{10}		
1.0	1.0	x_{11}		

The final approximation to this Jacobian generated by Broyden's method is

$$A_{10} = \begin{pmatrix} 1.999137 & 2.021829 \\ 0.9995643 & 3.11004 \end{pmatrix} \text{ where } J(x_*) \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}$$

In the above example A_{10} has a maximum relative error of 1.1% as approximate to $J(x_*)$.

In general when Broyden's method converges q-super linearly to x_*. One can not assume that the final Jacobian approximation A_k will approximately equal to $J(x_*)$

Example 2 Let $F(x) = \begin{pmatrix} x_1 + x_2 - 4 \\ x_1^2 + x_2^2 - 16 \end{pmatrix}$ has roots $(0, 4)^T$ and $(4, 0)^T$ let $x_0 = (2, 6)$ and $A_0 = J(x_0)$

The sequences of points generated by Broyden's method are given below.

Broyden's method		
x_0	2	6
x_1	-1	5
x_2	0	4

$$A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 11 \end{pmatrix} \text{ but } J(x_1) = \begin{pmatrix} 1 & 1 \\ -2 & 10 \end{pmatrix}$$

In general the final Jacobian approximation may not be approximately equal to $J(x_*)$. But if we could say how fast $\left\{ \frac{\|E_k s_k\|}{\|s_k\|} \right\}$ goes to 0, we could say more about the order of convergence of Broyden's method.

4. Implementation of Quasi-Newton algorithm using Broyden's update

First let us consider some important ideas of matrix factorizations.

For a matrix A the two important factorizations are PLR and the QR. The PLR decomposition yields A = P.L.R or $p^T A$ = L.R for a matrix p which has the same rows and columns as the identity matrix and not in the same order (permutation matrix p) a unit lower triangular L, upper triangular R. The decomposition is obtained by Gaussian elimination with partial pivoting. The QR decomposition yields A = Q.R, for orthogonal matrix Q, upper triangular matrix R. The decomposition is obtained by transforming A to R by multiplying A with a series of n-1 orthogonal matrices Q_i. Each Qi is zeros out of the elements of the i^{th} column of $Q_{i-1},\dots, Q_1$, A below the main diagonal, while leaving the first i-1 columns unchanged. Q_i is called Householder transformation and is often of the form $Q_i = I - u_i u_i^T$ where $(u_i)j = 0$ for i = 1, …, i – 1 and the remaining elements of u_i are selected so that Q_i is orthogonal induces the desired zeros in column i.

Now the main idea of this section is to consider two issues in using a secant approximation instead of analytic or finite difference Jacobian, these are:

1. The details involved in implementing the approximation.
2. The changes to be made if they are needed to the rest of the algorithm.

One of the most known problems of secant updates is how to get the first (initial) approximation A_0. But in most cases the finite difference approximation method is used to approximate A_0.

If Broyden's up date is implemented directly, then we want to say something about its implementation,

Let $t = y - As$ and $\delta = s^T s$,

$$A_{ij} = A_{ij} + \frac{t_i s_i^T}{\delta} \quad i = 1, \dots n, j = 1, \dots, n$$

Recall that the first thing our quasi = Newton algorithm will do with A_+ is determined its QR factorization in order to calculate the step - $A_+^{-1}F(x_+)$. Since the Broyden's update exactly fits the algebraic structure of equation

$A_+ = A_c + uv^T$, A_c is nonsingular matrix, the more efficient updating strategy is to update the factorization Q_cR_c of A_c in to the factorization Q_+R_+ of A_+ at every iteration after the first.

We will refer to these two possible implementations of Broyden's update as factored and unfactored forms. The factored form leads to more efficient algorithm and should be used in any production implementation. The un factored form is simpler to code and may be preferred for preliminary implementations.

The diagonal scaling of the independent variables causes Broyolens update to become,

$$A_+ = A_c + \frac{(y_c - A_c s_c)(D_x^2 s_c)^T}{s_c^T D_x^2 s_c} \tag{4.1.1}$$

And this is the update actually implemented. The update is independent of diagonal scaling of the dependent variables. Under some circumstances we decide not to change a row of A_c. Notice that if $(y_c - A_c s_c)_i = 0$ the (4.1.1) automatically causes row i of A_+ equal to row i of A_c. This makes sense because Broyden's update makes the smallest change to row i of A_c consistent with $(A_+)_i S_c = (y_c)i$ and if $(y_c - A_c s_c)_i = 0$, no change is necessary. Our algorithm also leave row i of A_c unchanged if $|(y_c - A_c s_c)i\ |$ is smaller than the computational uncertainty i n $(y_c)_i$, measured by $\eta(|\ f_i(x_c)\ | + |(f_i\ (x_+)|\ )$ where η is the relative error in function values.

There is another implementation of Broyden's method that also results in arithmetic work per iteration. It uses the Sherman-Morrison Woodbury formula given below:

<u>Lemma 4.1.1</u>: (Sherman-Morrison Woodbury) Let u, $v \in \Re^n$ and assume that $A \in \Re^{n \times n}$ is nonsingular. Then $A + uv^T$ is nonsingular if and only if

$$1 + v^T A^{-1} u \triangleq \delta \neq 0,$$

Further more

$$(A + uv^T)^{-1} = A^{-1} - \frac{1}{\delta} A^{-1} uv^T A^{-1}.$$

Proof: 1 Let $A + uv^T$ be nonsingular, i.e $(A + uv^T)^{-1}$ exists, let $B = A+uv^T$

$$(A + uv^T) B = I$$
$$AB + uv^T B = I$$
$$AB = I - uv^T B$$
$$B(A + uv^T) = I$$
$$BA + Buv^T = I$$
$$BA = I - Buv^T,$$

From $AB = I - uv^T B$

$$B = A^{-1} - A^{-1}uv^T B.$$

Multiple both sides by v^T from the left,

$$v^T B = v^T A^{-1} - v^T A^{-1} uv^T B$$

Multiply both sides by u from the right,

$$v^T Bu = v^T A^{-1}u - (v^T A^{-1}u)(v^T Bu)$$
$$v^T A^{-1}u = v^T Bu + (v^T A^{-1}u)(v^T Bu).$$
$$v^T A^{-1}u = v^T Bu + (1+ v^T A^{-1}u) \qquad *$$

obviously $1 + v^T A^{-1}u \neq 0$, i.e. $\delta \neq 0$. If $1 + v^T A^{-1}u = 0$,

i.e. $v^T A^{-1}u = -1 \neq 0$.

And from * we get $0 = -1$ contradiction

Therefore $1 + v^T A^{-1}u \triangleq \delta \neq 0$

2. Suppose $\delta \neq 0$, A is nonsingular,

so, $A^{-1} - \frac{1}{\delta} A^{-1}uv^T A^{-1}$ exists

$$(A + uvT) . (A^{-1} - \frac{1}{\delta} A^{-1}uv^T A^{-1})$$

$$= AA^{-1} - \frac{1}{\delta} AA^{-1}uv^T A^{-1} - \frac{1}{\delta} uv^T A^{-1}uv^T A^{-1} + uv^T A^{-1}$$

$$= I + [-\frac{1}{\delta}(1 + uv^T A^{-1}) + 1] uv^T A^{-1}$$

Since $1 + uv^T A^{-1} \triangleq \delta$, $= I + [-\frac{1}{\delta} . \delta + 1] = I$

And hence, $(A + UV^T)^{-1} = A^{-1} - \frac{1}{\delta} A^{-1}UV^TA^{-1}$

If A_c^{-1} is known, then by applying lemma 4.1.1 the Broyden's update can expressed as:

$$\text{If } A_+^{-1} = A_c^{-1} + \frac{\left(s_c - A_c^{-1} y_c\right) s_c^T A_c^{-1}}{s_c^T A_c^{-1} y_c} \tag{4.1.1.1}$$

The secant direction $-A_+F(x_+)$, then can be calculated by a matrix vector multiplication requiring an additional operations. The notion of sequencing QR factorization algorithms using Broyden's update were implemented by initially calculating A_0^{-1} and using (4.1.1.1). This implementation usually works fine in practice but its disadvantage is it makes it hard to detect ill-conditioning A_+.

The other question we answer in this section is "what changes need to be made to the rest of our quasi-Newton algorithm for non linear equations when we use secant up dates?". From the way we have resented the general quasi-Newton frame work we would like the answer to be "None" but there are two exceptions. When the factored form of the update is used, the QR factorization is ommitted and several implementation details are adjusted.

The other change is more significant, it is no longer guaranteed that the quasi-Newton direction $s_c = -A_c^{-1} F(x_c)$ is a decent direction for

$$f(x) = \frac{1}{2} \|F(x)\|_2^2 \text{ And } \nabla f(x_c) = \frac{d}{dx} \sum_{i=1}^{n} \frac{1}{2} (f_i(x_c))^2 = \sum_{i=1}^{n} \nabla f_i(x_c).\ f(x_c)$$

$$= J(x_c)^T F(x_c).$$

so that, $s^T \nabla f(x_c) = -F(x_c)^T A_c^{-1} J(x_c) F(x_c)$.

Which is non positive if $A_c = J(x_c)$ but not necessarily other wise. There is no way that we can check whether s_c is descent direction for $f(x)$ in a secant algorithm. Since we can not calculate $\nabla f(x_c)$ with out $J(x_c)$. This means that s_c may be an uphill direction for $f(x)$, and our global step may fail to produce any satisfactory point. Luckily this does not happen in practice since A_c is an approximation to $J(x_c)$. If it does, there is an easy fix we reset A_c, using a finite difference approximation to $J(x_c)$ and continue the algorithm.

5. Other Secant updates for non-liner equations

So far we have considered one secant update, i.e. Broyden's up date in approximating the Jacobian matrix. In this section we mention some other secant updates.

One of the most known alternative was suggested by Broyden. Calculating A_k^{-1} rather than A_k is the suggestion of Broyden.

$$A_+^{-1} = A_c^{-1} + \frac{\left(s_c - A_c^{-1}\right)y_c^T}{y_c^T y_c} \qquad (5.1.1)$$

is the solution of $\min_{B\in(y_c,s_c)} \|B^{-1} - A^{-1}\|_F$ (5.1.2)

From Lemma 4.1.1, (5.1.1) is equivalent to

$$A_+ = A_c + \frac{(y_c - A_c s_c)y_c^T}{y_c^T A_c s_c} \qquad (5.1.3)$$

which is clearly a different update from Broyden's up date.

Update (5.1.1) has attractive property that it produces values of A_+^{-1} directly, with out problems of a zero denominator. More over it shares all the important theoretical properties of Broyden's updates i.e. this method of approximating $J^{-1}(x)$ is of bounded deterioration as an approximation to $J^{-1}(x_*)$.

The sequence of iterates (x_k) converges q-super linearly to x_* assuming that x_0 is sufficiently close to x_*, $J(x_*)$ is nonsingular and A_0 is sufficiently close to $J(x_o)$. However this method is less successful than the same method using Broyden's up dates in practice.

Def. 4.1.1: The update (4.1.1) is called Broyden's bad update.

It is an interesting that Broyden's good update comes from minimizing the change in the affine model subject to satisfying the secant equation while the Broyden's bad update is related to minimizing the change in the solution to the model perhaps the fact that we use information about the function and not about its solution in forming the secant model is the reason why the former approach is more desirable.

There are other secant updates for nonlinear equations. For example $v_k \in \Re^n$ and $v_k s_k \neq 0$

$$A_{k+1} = A_k = \frac{(y_c - A_k s_k) V_k^T}{V_k^T s_k} \qquad (5.1.4)$$

is well defined and satisfies the secant equation. Broyden's good and bund updates are just the choices $V_k = s_k$ and $V_k = A_k^T y_k$ respectively.

Branes (1965) and Gay and Schnabel (1978) have proposed an update of the form (5.1.4) where V_k is the Euclidean projection of s_k orthogonal to $s_{k-1}, \ldots s_{k-m}$, $0 \leq m < n$. This enables A_+ to satisfy $m + 1$ secant equations $A_{k+1} s_i = y_i$ $i = k - m, \ldots, k$ meaning that the affine model interpolates $F(x_i)$, $i = k - m \ldots, k - 1$, as well as $F(x_k)$ and $F(x_{k+1})$.

In implementations where $m < n$ is choosen so that $s_{k-m}, \ldots, s_k$ are very linearly independent. This update has slightly out performed Broyden's update, but e xperience w th it is still limited. There for we still recommend Broyden's update as the secant approximation for solving systems of nonlinear equations.

References

1. Dennis, J.E, Jr. schnable. Robeert B.: Numerical Methods for unconstrained optimization and nonlinear equations, 1983.

2. Deumlich, R.: Optimization and Theory of approximation
   Text book for lecture Addis Ababa 1997.

3. Deumlich, R.: Optimization Theory I
   Text book for lecture Addis Ababa 1996

4. Deumlich, R.: Functional Analysis I
   Text book for lecture Addis Ababa 1997