Broyden's method Summary

Newton's method for systems:

$$\boxed{\vec{p}_{n+1} = \vec{p}_n - J^{-1}(p_n)\vec{F}(\vec{p}_n)} \qquad (1.1)$$

$O(n^2)$ funct. evaluations to form $J(\vec{p}_n)$.

$O(n^3)$ operations to solve system:

$$\boxed{J(\vec{p}_n)\vec{y}_{n+1} = -\vec{F}(\vec{p}_n)} \qquad (1.2)$$

where $$\boxed{\vec{y}_{n+1} = \vec{p}_{n+1} - \vec{p}_n.} \qquad (1.3)$$

Broyden's idea is to construct a matrix A_k A_k at every iteration step k, such that it only requires

$O(n)$ funct. evaluations. to construct it.

$O(n^2)$ operations. to solve system:

$$\boxed{\vec{X}^{(k+1)} = \vec{X}^{(k)} - A_k^{-1}\vec{F}(X^{(k)})} \qquad (1.4)$$

Broyden's method is inspired in Secant method

<u>Secant method</u> replace $f'(p_n)$ in Newt formula

$$p_{n+1} = p_n - \frac{f(p_n)}{f'(p_n)}$$

by $\quad f'(p_n) \rightarrow \dfrac{f(p_n) - f(p_{n-1})}{p_n - p_{n-1}} \quad$ (avoid Computation of derivatives)

Broyden's method replaces $J(\vec{p}_n)$ in (1.1) by

Only $O(n)$ functional evaluation.

<u>Step k:</u>

$$\boxed{A_k = A_{k-1} + \frac{\vec{F}(\vec{x}^{(k)}) - \vec{F}(\vec{x}^{(k-1)}) - A_{k-1}(\vec{x}^{(k)} - \vec{x}^{(k-1)})}{\|\vec{x}^{(k)} - \vec{x}^{(k-1)}\|^2}(\vec{x}^{(k)} - \vec{x}^{(k-1)})^T} \tag{2.1}$$

Motivation for definition of A_k in following pages.

First goal of having only $O(n)$ functional evaluations is accomplished by A_k.

QN4 - QN5

It can be proved that the inverse of A_k is given by (pages QN$_6$-QN$_7$)

$$A_k^{-1} = A_{k-1}^{-1} - \frac{\left[A_{k-1}^{-1}\left(\vec{F}(\vec{x}^{(k)}) - \vec{F}(\vec{x}^{(k-1)})\right) - (\vec{x}^{k} - \vec{x}^{k-1})\right](\vec{x}^{k} - \vec{x}^{k-1})^T A_{k-1}^{-1}}{(\vec{x}^{k} - \vec{x}^{k-1})^T A_{k-1}^{-1}\left(\vec{F}(\vec{x}_k) - \vec{F}(\vec{x}^{k-1})\right)}$$

Only n^2 operations: $\text{matrix}_{n\times n} * \text{vector}_{n\times 1}$

Implementation:

Broyden's method requires two initial guess (similar to secant method).

Step 1: $\vec{x}^0$ initial guess

Compute $J(\vec{x}^0)$. and $F(\vec{x}^0)$

Step 2: Compute $\vec{x}^1$ using Newton's method

But, first compute the inverse

$$J^{-1}(\vec{x}^0)$$

and then obtain

$$\vec{x}^{(1)} = \vec{x}^0 - J^{-1}(\vec{x}_0)F(\vec{x}^0)$$

Step 3:

Use $\vec{x}^0$, $\vec{x}^{(1)}$ and $J^{-1}(\vec{x}^0)$ to define A_1^{-1}

In fact, if

$$\vec{s}^1 = \vec{x}^1 - \vec{x}^0$$

$$\vec{y}^1 = F(\vec{x}^1) - F(\vec{x}^0) \longrightarrow n \text{ functional Evaluations.}$$

$$\vec{p}^1 = J^{-1}(\vec{x}^0)\,\vec{y} \longrightarrow n^2 \text{ operations}$$

Then,

$$\boxed{A_1^{-1} = J^{-1}(\vec{x}^0) + \frac{(\vec{s}^1 - \vec{p}^1) * \vec{s}^{1T} * J^{-1}(\vec{x}^0)}{\vec{s}^{1T} * \vec{p}^1}}$$

Step 4:

Compute $\vec{x}^2$

$$\boxed{\vec{x}^2 = \vec{x}^1 - \overbrace{A_1^{-1}\vec{F}(\vec{x}^1)}^{n^2 \text{ operations.}}}$$

STEP 5:

Repeat the previous steps. and obtain A_k^{-1} from A_{k-1}^{-1} and then $\vec{x}^{k+1}$

$$\boxed{A_k^{-1} = A_{k-1}^{-1} + \frac{(\vec{s}^k - \vec{p}^k) * (\vec{s}^k)^T A_{k-1}^{-1}}{(\vec{s}^k)^T \vec{p}^k}}$$

and $\boxed{\vec{x}^{k+1} = \vec{x}^k - A_k^{-1}\vec{F}(\vec{x}^k)}$

where $\vec{s}^k = \vec{x}^k - \vec{x}^{k-1}$, $\vec{y}^k = F(\vec{x}^k) - F(\vec{x}^{k-1})$

and $\vec{p}^k = A_{k-1}^{-1}\vec{y}^k$.

Exercise 10.1 (3)

Find the roots of

$$x_1^2 - 10x_1 + x_2^2 + 8 = 0$$

$$x_1 x_2^2 + x_1 - 10x_2 + 8 = 0$$

Initial guess: $\vec{x}^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Step 1: Computation of

$$J\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{bmatrix} -9 & 1 \\ 5/4 & -19/2 \end{bmatrix}$$

$$\vec{F}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{bmatrix} 7/2 \\ 29/8 \end{bmatrix}$$

Step 2:

$$J^{-1}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{bmatrix} -38/337 & -4/337 \\ -5/337 & -36/337 \end{bmatrix}$$

$$\vec{X}^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - J^{-1}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \vec{F}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.937685 \\ 0.939169 \end{pmatrix}$$

Step 3: Definition of A_1^{-1}

$$\vec{s}^1 = \vec{x}^1 - \vec{x}^0 = \begin{pmatrix} 295/674 \\ 148/337 \end{pmatrix}, \quad \vec{p}^1 = J^{-1}(\vec{x}^0)\left(F(\vec{x}^1) - F(\vec{x}^0)\right)$$

or $p^1 = \begin{pmatrix} 0.3899 \\ 0.3936 \end{pmatrix}$

Then,

$$A_1^{-1} = J^{-1}(1) + \frac{(\vec{s}^1 - \vec{p})(\vec{s}^1)^T J^{-1}(1)}{(\vec{s}^1)^T \vec{p}}$$

or

$$A_1^{-1} = \begin{bmatrix} -0.12 & -0.02 \\ -0.02 & -0.11 \end{bmatrix}, \quad \vec{F}\begin{pmatrix} 0.938 \\ 0.9391 \end{pmatrix} =$$

Step 4:

$$\hat{X}^{(2)} = \begin{pmatrix} 0.9912 \\ 0.9902 \end{pmatrix} - A_1^{-1}\begin{pmatrix} 0.3844 \\ 0.3731 \end{pmatrix} =$$

or

$$\hat{X}^{(2)} = \begin{pmatrix} 0.991154 \\ 0.990152 \end{pmatrix}$$

10.3 Quasi-Newton Methods.

They are inspired in the "Secant Method" which is used instead of Newton to reduce number of operations due to computation of $f'(x)$.

Newton Iteration:

$$p_n = g(p_{n-1}) = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \quad n \geq 1$$

p_0 initial guess.

Secant Method:

$$p_n = g(p_{n-1}) = p_{n-1} - \frac{f(p_{n-1})}{\frac{f(p_{n-1}) - f(p_{n-2})}{p_{n-1} - p_{n-2}}}, \quad n \geq 1$$

p_0, p_1 initial guess

$f'(p_{n-1})$ is replaced by $\dfrac{f(p_{n-1}) - f(p_{n-2})}{p_{n-1} - p_{n-2}}$

Recall that

$$f'(p_{n-1}) = \lim_{x \to p_{n-1}} \frac{f(x) - f(p_{n-1})}{x - p_{n-1}} \overset{\text{for } x \text{ close to } p_{n-1}}{\simeq} \frac{f(x) - f(p_{n-1})}{x - p_{n-1}}$$

If $x = p_{n-2}$, then

$$f'(p_{n-1}) \simeq \frac{f(p_{n-2}) - f(p_{n-1})}{p_{n-2} - p_{n-1}} = \frac{f(p_{n-1}) - f(p_{n-2})}{p_{n-1} - p_{n-2}}$$

QN2

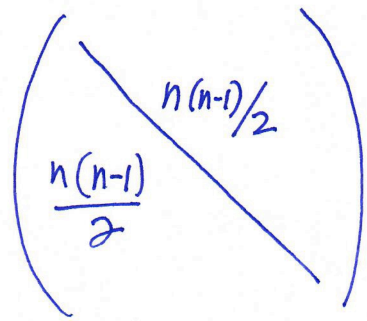
– Newton has quadratic convergence or $\lim_{n\to\infty} \frac{|p_n - p|}{|p_{n-1} - p|^2} = \alpha$

– Secant method is order ≈ 1.62 or it has superlinear convergence

$$\lim_{n\to\infty} \frac{|p_n - p|}{|p_{n-1} - p|^{1.62}} = \lambda.$$

<u>Counting operations</u>. <u>Computational cost.</u>

i) Gauss Elimination : $A_{n\times n}$



To eliminate each entrie below diagonal (in total $\frac{n(n-1)}{2}$) takes n operations.

Then G-E requires $\frac{n(n-1)}{2} \cdot n = O(n^3)$. operations.

ii) Computation of the Jacobian at each Newt. iteration requires n^2 funct. evaluation (1 funct. evaluation/per entrie)

iii) Evaluation of the vector $\vec{F}(\vec{x})$ in Newt. iteration for nonlinear systems

$$\begin{cases} - \boxed{J(\vec{x}^k)\,\vec{Y}^{(k)} = -\vec{F}(\vec{x}^k)} & (2.1) \\ - \boxed{\vec{X}^{(k+1)} = \vec{X}^k + \vec{Y}^k} & (2.2) \end{cases}$$

requires n function evaluations.

$$\boxed{\vec{X}^{(k+1)} = \vec{X}^k - J(\vec{x}^k)^{-1}\vec{F}(\vec{x}^k)} \qquad (2.3)$$

So

$$1 \text{ Newt. iteration} \simeq \begin{cases} \mathcal{O}(n^2+n) & \text{Evaluations} \\ n^3 & \text{operations} \end{cases}$$

Most attractive feature of Newt. method : Quadratic Convergence.

Broyden's Method

Jacobian matrix is approximated by a matrix A_K at every iteration step : $\boxed{\vec{x}^{(k+1)} = \vec{x}^{(k)} - A_k^{-1} \vec{F}(\vec{x}^k)}$

i) Computation of A_K only requires $\mathcal{O}(n)$ evaluations, instead of $\mathcal{O}(n^2)$ of Newton method per iteration.

ii) Computation of the inverse of A_k requires $\mathcal{O}(n^2)$ operations, instead of $\mathcal{O}(n^3)$ operations required by Newton method to solve linear system at each iteration.

Summarizing,

	Newt	Broyden
Constructing matrix	$\mathcal{O}(n^2)$	$\mathcal{O}(n)$
# opers. to solve linear system	$\mathcal{O}(n^3)$	$\mathcal{O}(n^2)$
order of convergence	order 2	Superlinear

$$\lim_{k\to\infty} \frac{\|\vec{x}^k - \vec{p}\|}{\|\vec{x}^{k-1} - \vec{p}\|} = 0$$

but $$\lim_{k\to\infty} \frac{\|\vec{x}^k - \vec{p}\|}{\|\vec{x}^{k-1} - \vec{p}\|^2} = \infty$$

Broyden method replaces $J(\vec{x}^k)$ by a matrix A_K

$$\boxed{\vec{X}^{(k+1)} = \vec{X}^{(k)} - A_k^{-1}\vec{F}(\vec{x}^k)} \qquad (4.1)$$

Construction of A_K at each iteration and ⓑ $\boxed{\vec{X}^{(1)} = \vec{X}^{(0)} - J(\vec{x}^0)^{-1}\vec{F}(\vec{x}^0)}$

Process starts with ⓐ $\boxed{A_0 = J(\vec{X}^0)}$. Then, to obtain A_1 two conditions are required

$$\boxed{A_1(\vec{X}^{(1)} - \vec{X}^{(0)}) = F(\vec{X}^1) - F(\vec{X}^0)} \qquad (4.2)$$

Not enough

and

$$\boxed{A_1\vec{z} = J(\vec{X}^0)\vec{z},} \quad \text{when } \vec{z}\cdot(\vec{X}^1 - \vec{X}^0) = 0 \qquad (4.3)$$

It can be shown that (4.2) and (4.3) defines A_1 uniquely.

$$A_1 \equiv J(\vec{x}^0) + \frac{[F(\vec{x}^1) - F(\vec{x}^0) - J(\vec{x}_0)(\vec{x}^1 - \vec{x}^0)](\vec{x}^1 - \vec{x}^0)^T}{\|\vec{X}^{(1)} - \vec{X}^{(0)}\|_2^2}$$

Check conds. (4.2) and (4.3).

A_1 is now used to compute $\vec{X}^2$.

$$\vec{X}^{(2)} = \vec{X}^{(1)} - A_1^{-1}\vec{X}^{(1)}$$

At step k : $\vec{X}^{(k+1)}$ is computed

Assuming $\vec{X}^{(k)}$ and A_{k-1} are known

$$A_k = A_{k-1} + \underbrace{\frac{[F(\vec{x}^k) - \vec{F}(\vec{x}^{k-1}) - A_{k-1}(\vec{x}^k - \vec{x}^{k-1})](\vec{x}^k - \vec{x}^{k-1})^T}{\|\vec{x}^k - \vec{x}^{k-1}\|_2^2}}$$

and

↳ n function evaluations. only.

$$\boxed{\vec{X}^{k+1} = \vec{X}^{(k)} - A_k^{-1}\vec{F}(\vec{x}^k)}$$

It still requires $O(n^3)$ operations for matrix inversion of A_k^{-1}

Not Enough!

SHerman-Morrison Formula

1) A nonsingular
2) $\vec{x}, \vec{y}$ two vectors such that $\vec{y}^t A^{-1} \vec{x} \neq -1$

Then, $A + \vec{x}\vec{y}^T$ is nonsingular and

$$(A + \vec{x}\vec{y}^T)^{-1} = A^{-1} - \frac{A^{-1}\vec{x}\vec{y}^T A^{-1}}{1 + \vec{y}^T A^{-1}\vec{x}}$$

Thm. 10.8 (Sherman Morrison Formula)

If ① A is nonsingular.

② $\vec{u}, \vec{w}$ two vectors such that $\vec{w}^T A^{-1} \vec{u} \neq -1$

Then,

$A + uw^t$ is nonsingular and

$$\left(A + uw^t\right)^{-1} = A^{-1} - \frac{A^{-1}\vec{u}\vec{w}^t A^{-1}}{1 + \vec{w}^t A^{-1}\vec{u}} \qquad (6.0)$$

Application to obtain Broyden's method

At step k, Broyden's method matrix is given by

$$(6.1) \qquad A_k = A_{k-1} + \frac{\left[\vec{F}(\vec{x}^k) - \vec{F}(\vec{x}^{k-1}) - A_{k-1}(\vec{x}^k - \vec{x}^{k-1})\right](\vec{x}^k - \vec{x}^{k-1})^T}{\|\vec{x}^k - \vec{x}^{k-1}\|_2^2}$$

If $\boxed{\vec{u} = \frac{\left[\vec{F}(\vec{x}^k) - \vec{F}(\vec{x}^{k-1}) - A_{k-1}(\vec{x}^k - \vec{x}^{k-1})\right]}{\|\vec{x}^k - \vec{x}^{k-1}\|_2^2}}$

and $\boxed{\vec{w} = (\vec{x}^k - \vec{x}^{k-1})}$

Then (6.1) reduces to $\boxed{A_k = A_{k-1} + \vec{u}\vec{w}^T}$

QN7

Applying S-M formula (6.0)

$$A_k^{-1} = \left(A_{k-1} + \vec{u}\vec{w}^T\right) = A_{k-1}^{-1} - \frac{A_{k-1}^{-1}\,\vec{u}\,\vec{w}^T A_{k-1}^{-1}}{1 + w^T A_{k-1}^{-1}\vec{u}}$$

Now,

$$A_{k-1}^{-1}\cdot\vec{u}\vec{w}^T A_{k-1}^{-1} = A_{k-1}^{-1}\left(\frac{\left[\vec{F}(\vec{x}^k)-\vec{F}(\vec{x}^{k-1}) - A_{k-1}(\vec{x}^k-\vec{x}^{k-1})\right]}{\|\vec{x}^k-\vec{x}^{k-1}\|_2^2}\,(\vec{x}^k-\vec{x}^{k-1})^T\right)A_{k-1}^{-1}$$

or

$$\boxed{A_{k-1}^{-1}\vec{u}\vec{w}^T A_{k-1}^{-1} = \frac{\left[A_{k-1}^{-1}\left(\vec{F}(\vec{x}^k)-\vec{F}(\vec{x}^{k-1})\right) - (\vec{x}^k-\vec{x}^{k-1})\right](\vec{x}^k-\vec{x}^{k-1})^T A_{k-1}^{-1}}{\|\vec{x}^k-\vec{x}^{k-1}\|_2^2}}$$

and $$1+w^T A_{k-1}^{-1}\vec{u} = 1 + (\vec{x}^k-\vec{x}^{k-1})^T A_{k-1}^{-1}\left(\frac{\vec{F}(\vec{x}^k)-\vec{F}(\vec{x}^{k-1}) - A_{k-1}(\vec{x}^k-\vec{x}^{k-1})}{\|\vec{x}^k-\vec{x}^{k-1}\|_2^2}\right)$$

or

$$\boxed{1+\vec{w}^T A_{k-1}^{-1}\vec{u} = \frac{\cancel{\|\vec{x}^k-\vec{x}^{k-1}\|_2^2} + (\vec{x}^k-\vec{x}^{k-1})^T A_{k-1}^{-1}\left(\vec{F}(\vec{x}^k)-\vec{F}(\vec{x}^{k-1})\right) - \cancel{\|\vec{x}^k-\vec{x}^{k-1}\|_2^2}}{\|\vec{x}^k-\vec{x}^{k-1}\|_2^2}}$$

Therefore,

$$\boxed{A_k^{-1} = A_{k-1}^{-1} - \frac{\left[A_{k-1}^{-1}\left(\vec{F}(\vec{x}^k)-\vec{F}(\vec{x}^{k-1})\right) - (\vec{x}^k-\vec{x}^{k-1})\right](\vec{x}^k-\vec{x}^{k-1})^T A_{k-1}^{-1}}{(\vec{x}^k-\vec{x}^{k-1})^T A_{k-1}^{-1}\left(\vec{F}(\vec{x}_k)-\vec{F}(\vec{x}^{k-1})\right)}}$$

only n^2 Operations

$(\text{matrix})_{n\times n}$ times a $(\text{Vector})_n$