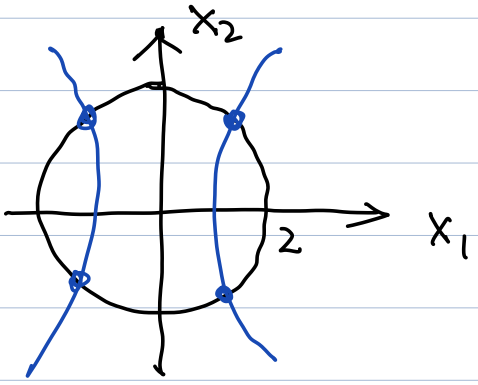
Nonlinear Eqns

e.g. 1.

$$x_1^2 + x_2^2 - 4 = 0$$

$$4x_1^2 - x_2^2 - 4 = 0$$



General form:

$x_1, x_2, \cdots, x_n$: unknowns

$$f_1(x_1, x_2, \cdots, x_n) = 0$$

$$f_2(x_1, x_2, \cdots, x_n) = 0$$

- - - -

$$f_n(x_1, x_2, \cdots, x_n) = 0$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \vec{f} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

$$\Downarrow$$

$$\underline{\vec{f}(\vec{x}) = 0}$$

Newton's method: $\vec{x}^{(0)}$: initial guess

$$\vec{f}(\vec{x}) = \vec{f}(\vec{x}^{(0)}) + \underline{D\vec{f}(\vec{x}^{(0)})}(\vec{x} - \vec{x}^{(0)}) + \text{h.o.t.}$$

$$D\vec{f} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

$\downarrow$ Jacobian matrix

$$\vec{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$\vec{f}(\vec{x}^{(0)}) + D\vec{f}(\vec{x}^{(0)})(\vec{x} - \vec{x}^{(0)}) = 0$$

$$\vec{x}^{(1)} = \vec{x}^{(0)} - D\vec{f}(\vec{x}^{(0)})^{-1}\vec{f}(\vec{x}^{(0)})$$

$$\cdots$$

$$\boxed{\vec{x}^{(n)} = \vec{x}^{(n-1)} - \underbrace{D\vec{f}(\vec{x}^{(n-1)})^{-1}\vec{f}(\vec{x}^{(n-1)})}}$$

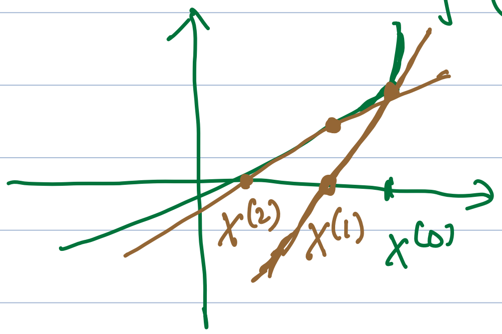
Newton's method

$A^{-1}\vec{b} = \vec{y}$

$A\vec{y} = \vec{b}$

one-dim case:

$$x^{(n)} = x^{(n-1)} - \frac{f(x^{(n-1)})}{f'(x^{(n-1)})}$$



Nonlinear Eqns

$\vec{f}(\vec{x}) = 0$

$\phi = f_1(\vec{x})^2 + f_2(\vec{x})^2 + \cdots + f_n(\vec{x})^2$

$\min \phi(\vec{x})$

$\Rightarrow \vec{f}(\vec{x}) = 0$

Optimization

$\min\limits_{\vec{x}\in\mathbb{R}^n} \phi(\vec{x})$

$\phi: \mathbb{R}^n \to \mathbb{R}$

$\nabla\phi(\vec{x}) = 0$

nonlinear system

$\frac{\partial\phi}{\partial x_1}(\vec{x}) = 0$

$\frac{\partial\phi}{\partial x_2}(\vec{x}) = 0$

$\cdots$

$$\frac{\partial \phi}{\partial x_n}(\vec{x}) = 0$$

Quasi-Newton's methods (Broyden's method)

Secant method:

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{\underset{A_k =}{\underbrace{\boxed{\frac{f(x^{(k)}) - f(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}}}}} \approx f'(x^{(k)})$$

$$A_k (x^{(k)} - x^{(k-1)}) = f(x^{(k)}) - f(x^{(k-1)})$$

$$x^{(k+1)} = x^{(k)} - A_k^{-1} f(x^{(k)})$$

Generalization to $\vec{f}(\vec{x}) = 0$

$$A_k \underbrace{(\vec{x}^{(k)} - \vec{x}^{(k-1)})}_{\vec{s}_k} = \underbrace{\vec{f}(\vec{x}^{(k)}) - \vec{f}(\vec{x}^{(k-1)})}_{\vec{y}_k}$$

If the relation holds, A_k is said to be compatible.

Given A_{k-1}, define

$$\underline{A_k} = \underset{\underset{B}{\downarrow}}{A_{k-1}} + \boxed{\frac{\overset{\vec{u}}{\boxed{(\vec{y}_k - A_{k-1} \vec{s}_k)}} \overset{\vec{v}}{\boxed{\vec{s}_k^{\,T}}}}{\| \vec{s}_k \|_2^2}}$$

e.g.:

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} (2, 3) = \begin{pmatrix} 2 & 3 \\ -2 & -3 \end{pmatrix} \qquad \underline{\text{rank} = 1}$$

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} - \underline{A_k^{-1}}\, \vec{f}(\vec{x}^{(k)})$$

- A_k is compatible

$$A_k \vec{s}_k \approx \cancel{A_{k-1}\vec{s}_k} + \frac{(\vec{y}_k - \cancel{A_{k-1}\vec{s}_k})\,\cancel{\vec{s}_k^T\vec{s}_k}}{\cancel{\|\vec{s}_k\|_2^2}}$$

$$\vec{v} = \begin{pmatrix} \vec{v}_1 \\ \vdots \\ \vec{v}_n \end{pmatrix} \qquad \vec{v}^T\vec{v} = v_1^2 + v_2^2 + \cdots + v_n^2$$
$$= \|\vec{v}\|_2^2$$

$$\Rightarrow \quad A_k \vec{s}_k = \vec{y}_k$$

- Sherman - Morrison formula

$$(\underset{\text{invertible}}{\underset{\downarrow}{B}} + \underset{\text{rank-1}}{\underset{\downarrow}{\vec{u}\vec{v}^T}})^{-1} = B^{-1} - \frac{B^{-1}\vec{u}\vec{v}^T B^{-1}}{1 + \vec{v}^T B^{-1}\vec{u}}$$

It holds if $\vec{v}^T B^{-1}\vec{u} \neq -1$

$$(B + \vec{u}\vec{v}^T)\left(B^{-1} - \frac{B^{-1}\vec{u}\vec{v}^T B^{-1}}{1 + \vec{v}^T B^{-1}\vec{u}}\right)$$

$$= I - \frac{u\vec{v}^T B^{-1}}{1 + \vec{v}^T B^{-1}\vec{u}} + \vec{u}\vec{v}^T B^{-1} \swarrow$$

$$- \frac{\vec{u}\,\boxed{\vec{v}^T B^{-1}\vec{u}}\,\vec{v}^T B^{-1}}{1 + \vec{v}^T B^{-1}\vec{u}} \longrightarrow \text{a number}$$

$$= I - \frac{\cancel{(1 + \vec{v}^T B^{-1}\vec{u})}\,\underline{\vec{u}\vec{v}^T B^{-1}}}{\cancel{1 + \vec{v}^T B^{-1}\vec{u}}} + \underline{\vec{u}\vec{v}^T B^{-1}}$$

$$= I$$

Initially:

$$\begin{cases} B_0 = I \\ B_0 = \operatorname{diag} Df(\vec{x}^{(0)}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & & & \\ & \frac{\partial f_2}{\partial x_2} & & \\ & & \ddots & \\ & & & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \\ B_0 = Df(\vec{x}^{(0)}) \Leftarrow \text{Gaussian-Jordan} \end{cases}$$

Quasi-Newton's method

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} - A_k^{-1} \vec{f}(\vec{x}^{(k)}), \qquad A_0 \text{ given}$$

$$A_k = \underbrace{A_{k-1}}_{B} + \underbrace{\boxed{\frac{(\vec{y}_k - A_{k-1}\vec{s}_k)}{\|\vec{s}_k\|_2^2}}}_{\vec{u}} \; \underbrace{\vec{s}_k^{\,T}}_{\vec{v}}$$

SM ⇒

$$A_k^{-1} = B^{-1} - \frac{B^{-1}\vec{u}\,\vec{v}^T B^{-1}}{1 + \vec{v}^T B^{-1}\vec{u}} \qquad B^{-1} = \underline{A_{k-1}^{-1}}$$

Let $\vec{w} = B^{-1}\vec{u} = \dfrac{B^{-1}\vec{y}_k - \vec{s}_k}{\|\vec{s}_k\|^2}$

$$\Rightarrow A_k^{-1} = B^{-1} - \frac{\vec{w}\,\vec{v}^T B^{-1}}{1 + \vec{v}^T \vec{w}}$$

} only involve A_{k-1}^{-1}

Convergence properties:

$$\vec{x}^{(k)} = \vec{x}^{(k-1)} - D\vec{f}(x^{(k-1)})^{-1}\vec{f}(x^{(k-1)})$$

Let $\vec{p}$ be the solution. $\vec{f}(\vec{p}) = 0$

Theorem: Assume that:

(1) $D\vec{f}(\vec{p})^{-1}$ non-singular.

(2) $\|D\vec{f}(\vec{x}) - D\vec{f}(\vec{y})\| \leq L\,\|\vec{x} - \vec{y}\|$, for all $\vec{x}$ & $\vec{y}$

(3) $\vec{x}^{(0)}$ is "close" to $\vec{p}$

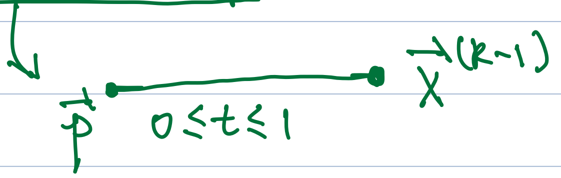
Then $\vec{x}^{(k)} \to \vec{p}$ and

$$\underbrace{\|\vec{x}^{(k)} - \vec{p}\|}_{\vec{e}_k} \underset{\text{quadratic}}{\leq} C\,\underbrace{\|\vec{x}^{(k-1)} - \vec{p}\|}_{\vec{e}_{k-1}}^{2}$$

$$\underset{-p}{\vec{x}^{(k)}} = \underset{-p}{\vec{x}^{(k-1)}} - D\vec{f}(x^{(k-1)})^{-1}\vec{f}(x^{(k-1)})$$

$$\Rightarrow \vec{e}_k = \vec{e}_{k-1} - D\vec{f}(\vec{x}^{(k-1)})^{-1}\left(\vec{f}(\vec{x}^{(k-1)}) - \vec{f}(\vec{p})\right)$$

Let $\phi(t) = \vec{f}(\underline{\vec{p} + t(\vec{x}^{(k-1)} - \vec{p})})$



$$\phi(1) - \phi(0) = \int_0^1 \phi'(t)\,dt$$

||

$$\vec{f}(\vec{x}^{(k-1)}) - \overset{\Downarrow}{\vec{f}(\vec{p})} = \int_0^1 D\vec{f}(\vec{p} + t\vec{e}_{k-1})\,dt\ \vec{e}_{k-1}$$

$$\vec{e}_k = D\vec{f}(\vec{x}^{(k-1)})^{-1}\left\{ D\vec{f}(\vec{x}^{(k-1)}) - \int_0^1 D\vec{f}(\vec{p} + t\vec{e}_{k-1})\,dt \right\} \vec{e}_{k-1}$$

$$= \underline{D\vec{f}(\vec{x}^{(k-1)})^{-1}}\left\{ \int_0^1 D\vec{f}(\vec{x}^{(k-1)}) - D\vec{f}(\vec{p} + t\vec{e}_{k-1})\,dt \right\} \vec{e}_{k-1}$$

$$\|\vec{e}_k\| \lesssim C L \underline{\int_0^1 (1-t)\,dt}\ \underline{\|\vec{e}_{k-1}\|\ \|\vec{e}_{k-1}\|}$$

$$\|AB\| \leq \|A\|\,\|B\|$$
$$\|A\vec{x}\| \leq \|A\|\,\|\vec{x}\|$$

In general, ↗ modified-Newton

$$\|\vec{e}_k\| \leq c\,\|\vec{e}_{k-1}\|, \ 0<c<1, \ \text{linear}$$

$$\|\vec{e}_k\| \leq c\,\|e_{k-1}\|^{\alpha}, \ 1<\alpha<2, \ \text{superlinear}$$

↓ example: Quasi-Newton