Lecture 23, 24 Extensions of Least Square.

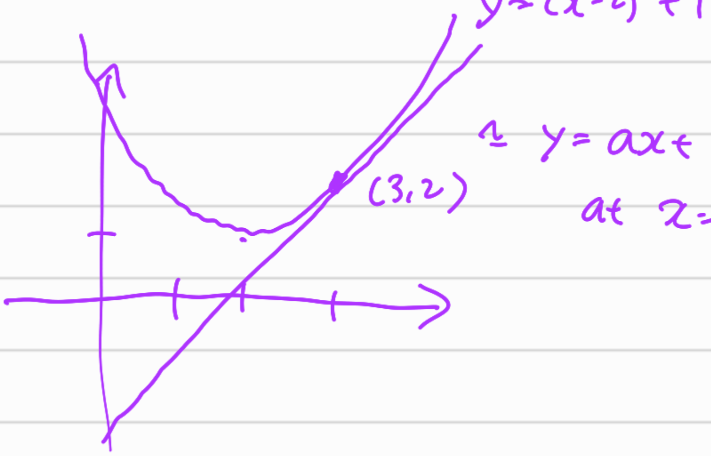
- Regularized LS
- 1) - Nonlinear LS.
- Recursive LS.
- ✓ - Constrained LS (KKT, Laguargian ⇒ quadratic programming)
- Robust LS
- Weighted LS
- Multi-objective LS

⋮

Linearization.

$y = (x-2)^2 + 1$

$= x^2 - 4x + 5$



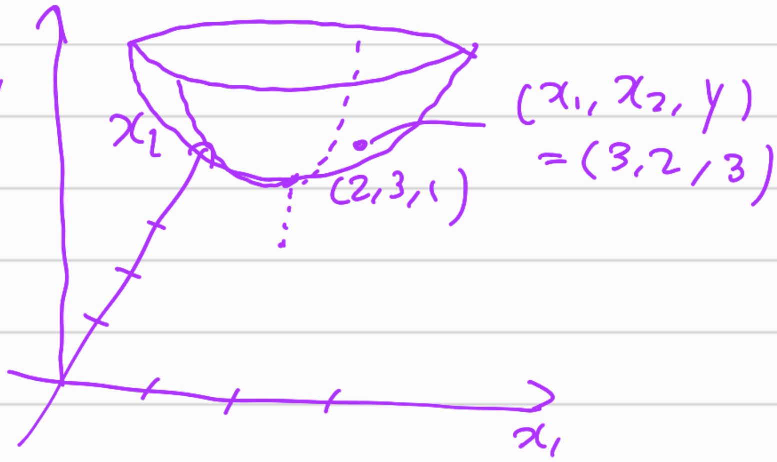
$\simeq y = ax + b$ at $x = 3$

$x = 3, \quad y' = \frac{dy}{dx} = 2x - 4 \Big|_{x=3} = 2$

$y = 2(x-3) + 2$

$= 2x - 4$

$$y = (x_1-2)^2 + (x_2-3)^2 + 1$$



at $(x_1, x_2) = (3,2)$

$$y = 3 + \left.\frac{\partial y}{\partial x_1}\right|_{x_1,x_2=(3,2)} (x_1-3) + \left.\frac{\partial y}{\partial x_2}\right|_{x_1,x_2=(3,2)} (x_2-2)$$

$$= 3 + \left.[2x_1-4]\right|_{(3,2)} (x_1-3) + \left.[2x_2-6]\right|_{(3,2)} (x_2-2)$$

$$= 3 + 2(x_1-3) - 2(x_2-2)$$

$$= 2x_1 - 2x_2 + 1$$

$x \in \mathbb{R}^n$, $f(x) : \mathbb{R}^n \to \mathbb{R}$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad x^{(0)} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{bmatrix}$$

at $x = x^{(0)}$

$$f(x) \simeq f(x^{(0)}) + \left.\frac{\partial f}{\partial x_1}\right|_{x=x^{(0)}} (x_1 - x_1^{(0)})$$

$$+ \left.\frac{\partial f}{\partial x_2}\right|_{x=x^{(0)}} (x_2 - x_2^{(0)})$$

$$+ \cdots$$

$$+ \left.\frac{\partial f}{\partial x_n}\right|_{x=x^{(0)}} (x_n - x_n^{(0)})$$

$$= f(x^{(0)}) + \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \cdots \quad \frac{\partial f}{\partial x_n} \right] \Bigg|_{x = x^{(0)}} \begin{bmatrix} x_1 - x_1^{(0)} \\ x_2 - x_2^{(0)} \\ \vdots \\ x_n - x_n^{(0)} \end{bmatrix}$$

① $$= f(x^{(0)}) + \nabla f(x^{(0)}) (x - x^{(0)})$$

at $x^{(0)}$ linearization $\Uparrow$

$A : \mathbb{R}^n \to \mathbb{R}^m$

LS,

$A : m \times n \quad x : n \times 1$

Linear optimization. $\| Ax - b \|^2 = \| e \|^2$ (e: $m \times 1$) $= \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}$

Nonlinear optimization $\| g(x) - b \|^2 = \| f(x) \|^2 = \| e \|^2 = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}$

$f(x) : \mathbb{R}^n \to \mathbb{R}^m$

$f(x^{(0)}) = \nabla f(x) (x - x^{(0)}) \qquad f(x) : \mathbb{R}^n \to \mathbb{R}^1$

* $f(x^{(0)}) = J(x^{(0)}) (x - x^{(0)}) \qquad f(x) : \mathbb{R}^n \to \mathbb{R}^m$

$$\begin{bmatrix} f_1(x^{(0)}) \\ f_2(x^{(0)}) \\ f_3(x^{(0)}) \\ \vdots \\ f_m(x^{(0)}) \end{bmatrix} \xrightarrow{\nabla} \begin{bmatrix} \nabla f_1(x^{(0)}) \\ \nabla f_2(x^{(0)}) \\ \vdots \\ f_m(x^{(0)}) \end{bmatrix} = J$$

② $$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & & & \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

LS. $\min \|Ax-b\|^2 \Rightarrow \frac{d}{dx}\|Ax-b\|^2 = 0$

$$\frac{d}{dx}\|Ax-b\|^2 = \frac{d}{dx}(Ax-b)^T(Ax-b)$$
$$= \frac{d}{dx}\, x^TA^TAx - 2x^TA^Tb + b^Tb$$
$$= 2A^TAx - 2A^Tb$$
$$= 0$$
$$A^TAx = A^Tb$$

Nonlinear optimization.

$\Rightarrow$ Linearize and least square.

$$\|f(x)\|^2 \simeq \|f(x^{(0)}) + J(x^{(0)})(x-x^{(0)})\|^2$$
$$= \|\underbrace{J(x^{(0)})}_{A}x - \underbrace{(J(x^{(0)})x^{(0)} - f(x^{(0)})}_{b}\|^2$$
$$= \|Ax-b\|^2$$

$$\underbrace{J(x^{(0)})^T J(x^{(0)})}\, x = J(x^{(0)})^T\{J(x^{(0)})x^{(0)} - f(x^{(0)})\}$$

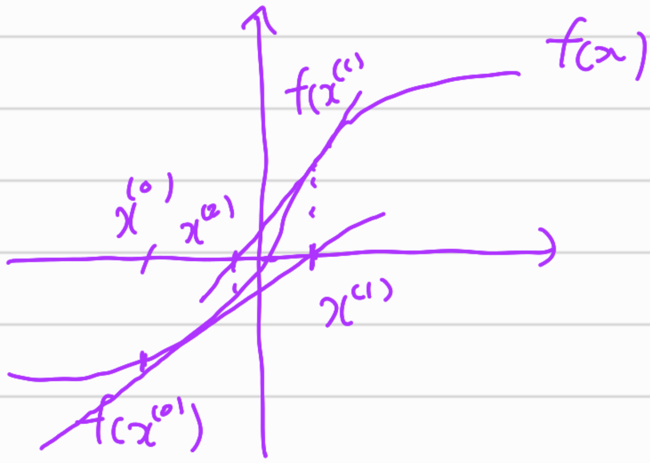
③
$$x = (J(x^{(0)})^TJ(x^{(0)}))^{-1}\{J(x^{(0)})^TJ(x^{(0)})x^{(0)} - J(x^{(0)})^Tf(x^{(0)})\}$$
$$= \underbrace{x^{(0)} - \{J(x^{(0)})^TJ(x^{(0)})\}^{-1}J(x^{(0)})^Tf(x^{(0)})}$$

optimized solution.

$f(x) \rightarrow$ linearize at $x^{(0)}$

$x^{(0)} \rightarrow x^{(1)} \rightarrow$ linearize $\rightarrow x^{(2)} \rightarrow$ linearize $\rightarrow \cdots \rightarrow$ 'converged'

④ Gauss-Newton method.



Find $f(x)=0$

$x^{(0)}$

$$f(x) \simeq f(x^{(0)}) + \nabla f(x^{(0)})(x-x^{(0)}) = 0$$

$x^{(1)}$

$$\rightarrow f(x) \simeq f(x^{(1)}) + \nabla f(x^{(1)})(x-x^{(1)}) = 0$$

$x^{(2)}$

1) linearize

$f(x): \mathbb{R}^1 \to \mathbb{R}^1$

$x^{(0)} \rightarrow$ linearize $\rightarrow$ solution $x^{(1)}$

$\rightarrow$ linearize $\rightarrow$ solution $x^{(2)}$

$\rightarrow$ linearize $\rightarrow$.

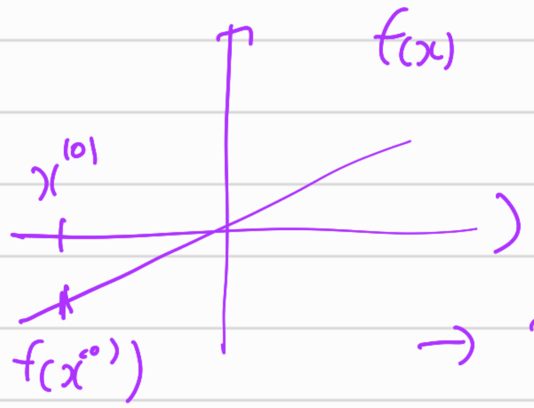
...

at $x^{(k)}$

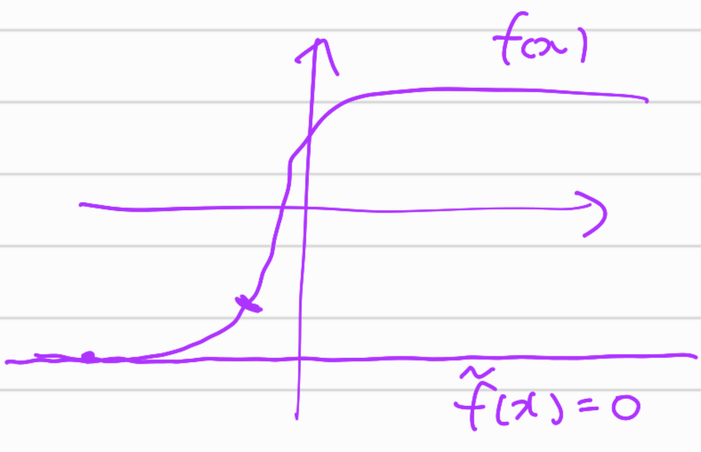
by iteratively linearizing $f(x)$, find a solution $x^{(k+1)}$

$k \gg 1$ $x^{(k)} \rightarrow$ converges to a solution.

$f(x) = \mathbb{R}^n \to \mathbb{R}^m$



$\rightarrow x^{(1)} \Rightarrow$ global solution ✓



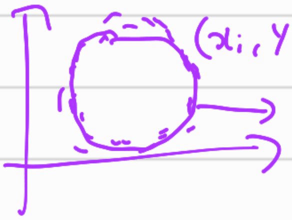
$\Rightarrow$ solution doesnot come out.

G-N doesn't work for the case where linear approximation is not good enough.

Circle fitting. (x_i, y_i)

$$e_i = f(x) = \sqrt{(x_i - a)^2 + (y_i - b)^2} - r$$

find a, b, r



$$x = \begin{bmatrix} a \\ b \\ r \end{bmatrix}$$

$$\nabla f(x) = \left[\frac{\partial f}{\partial a} \quad \frac{\partial f}{\partial b} \quad \frac{\partial f}{\partial r} \right]$$

$$\frac{\partial f}{\partial a} = \frac{\partial}{\partial a} \left\{ \left((x_i - a)^2 + (y_i - b)^2 \right)^{\frac{1}{2}} - r \right\}$$

$$= \frac{1}{2} \left\{ (x_i - a)^2 + (y_i - b)^2 \right\}^{-\frac{1}{2}} (2a - 2x_i)$$

$$= \frac{a - x_i}{\sqrt{(x_i - a)^2 + (y_i - b)^2}}$$

$$\frac{\partial f}{\partial b} = \frac{b - y_i}{\sqrt{(x_i - a)^2 + (y_i - b)^2}}$$

$$\frac{\partial f}{\partial r} = -1$$

$x = \begin{bmatrix} a \\ b \\ r \end{bmatrix}$ pick $x^{(0)} = \begin{bmatrix} a^{(0)} \\ b^{(0)} \\ r^{(0)} \end{bmatrix}$

$$x^{(1)} = x^{(0)} - \left\{ J(x^{(0)})^T J(x^{(0)}) \right\}^{-1} J(x^{(0)})^T f(x^{(0)})$$

$$J(x^{(0)}) = \begin{bmatrix} \nabla f_1(x^{(0)}) \\ \nabla f_2(x^{(0)}) \\ \vdots \\ \nabla f_m(x^{(0)}) \end{bmatrix} =$$

$$\Rightarrow x^{(k)} = \begin{bmatrix} a \\ b \\ r \end{bmatrix}$$

"Levenberg - Marquardt Algorithm" ⑤ ⑥

$$x^{(k+1)} = x^{(k)} - \{ \underbrace{J(x^{(k)})^T J(x^{(k)})}_{} \}^{-1} J(x^{(k)})^T f(x^{(k)})$$

$$x = (A^TA)^{-1}A^Tb$$

regularization $x = (A^TA + \gamma I)^{-1}A^Tb$

$$x^{(k+1)} = x^{(k)} - \{ J^TJ + \gamma I \}^{-1} A^T b \Rightarrow \underline{L\text{-}M}$$

// Linearization $x \simeq x^{(0)}$ valid.

min $\|f(x)\|^2 + \gamma \|x - x^{(0)}\|^2$

$$\simeq \| J(x^{(0)})x - (\underbrace{J(x^{(0)})x^{(0)} - f(x^{(0)})}_{b})\|^2 + \gamma \|x - \underbrace{x^{(0)}}_{c}\|^2$$

$$\frac{d}{dx}(\ ) = 0.$$

$$\frac{d}{dx}(Jx - b)^T(Jx - b) + \gamma(x-c)^T(x-c) = 0$$

$$\frac{d}{dx} x^TJ^TJx - 2b^TJx + b^Tb + \gamma x^Tx - 2\gamma x^Tc + c^Tc$$

$$\Rightarrow (J^TJ + \gamma I)x - J^Tb + c = 0$$

$$x = (J^TJ + \gamma I)^{-1}\{J^Tb + c\}$$

$$= x^{(0)} - (J^TJ + \gamma I)^{-1} J^T f(x^{(0)})$$

$$\|Ax-b\|^2 + \gamma\|x\|^2$$

$$\|J(x-x^{(0)}) - f(x^{(0)})\|^2 + \gamma\|x-x^{(0)}\|^2$$

"<u>L-M</u>"

① Linearization (Gradient)
② Jacobian (for multiple data samples)
③ Linearize and LS
④ Iterate ③ $\Rightarrow$ G-N
⑤ Linearization is only valid $x \simeq x^{(0)}$
⑥ regularization $\Rightarrow$ L-M

Nonlinear Opt.