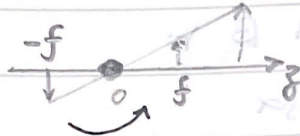


## Pose Estimation via Vision (15) notes

- pinhole model
- triangulation (2D-3D)
- ICP problem (3D-3D)
- optimization

### pinhole model



$$\frac{Z}{f} = \frac{X}{x'} = \frac{Y}{y'}$$

$$\therefore \begin{cases} x' = f \frac{X}{Z} \\ y' = f \frac{Y}{Z} \end{cases}$$

from  $x', y'$  to  $u, v$   
enlarge  $\alpha, \beta$   
translate  $c_x, c_y$

$$\begin{cases} u = \alpha x' + c_x \\ v = \beta y' + c_y \end{cases}$$

$$\Rightarrow \begin{cases} u = \alpha f \frac{X}{Z} + c_x \\ v = \beta f \frac{Y}{Z} + c_y \end{cases}$$

$$\Rightarrow \begin{cases} u = f_x \frac{X}{Z} + c_x \\ v = f_y \frac{Y}{Z} + c_y \end{cases}$$

let  $s = \frac{1}{Z}$

$$\Rightarrow s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

$$\Rightarrow s u = K P_0$$

$$= K T_{cw} P_w$$

$$= K (R_{cw} P_w + t)$$

$$s u = f_x X + c_x Z$$

$$u = f_x \frac{X}{Z} + c_x$$

$s u =$

$s u$

$$\frac{(u - c_x)}{f_x} \cdot s$$

$\frac{s}{f_x}$

$$\frac{376 - 635}{630} \cdot 4.3 = -1.76$$

$$\frac{432 - 635}{630} \cdot 5.56 = -1.79$$

$$\frac{209 - 369}{629} \cdot 4.3 = -1.09$$

$$\frac{212 - 369}{629} \cdot 5.56 = -1.38$$

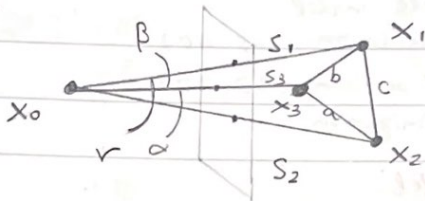
$$\frac{212 - 369}{629} \cdot 5.56 = -1.38$$

$$(376, 209)$$

$$(432, 212)$$

(376.209)

(432, 212)

Triangulation (2D-3D)P3Pknown  $X_1, X_2, X_3, \alpha, \beta, r$ 

$$a^2 = S_2^2 + S_3^2 - 2S_2S_3 \cos \alpha$$

$$b^2 = S_1^2 + S_3^2 - 2S_1S_3 \cos \beta$$

$$c^2 = S_1^2 + S_2^2 - 2S_1S_2 \cos r$$

$$\textcircled{1} a^2 = S_2^2 + S_3^2 - 2S_2S_3 \cos \alpha$$

$$\text{set } u = \frac{S_2}{S_1}, \quad v = \frac{S_3}{S_1}$$

$$a^2 = S_1^2(u^2 + v^2 - 2uv \cos \alpha)$$

$$\therefore S_1^2 = \frac{a^2}{u^2 + v^2 - 2uv \cos \alpha}$$

$$\textcircled{2} b^2 = S_1^2 + S_3^2 - 2S_1S_3 \cos \beta$$

$$\text{set } u = \frac{S_2}{S_1}, \quad v = \frac{S_3}{S_1}$$

$$b^2 = S_1^2(1 + v^2 - 2v \cos \beta)$$

$$\therefore S_1^2 = \frac{b^2}{1 + v^2 - 2v \cos \beta}$$

$$\textcircled{3} c^2 = S_1^2 + S_2^2 - 2S_1S_2 \cos r$$

$$\text{set } u = \frac{S_2}{S_1}, \quad v = \frac{S_3}{S_1}$$

$$c^2 = S_1^2(1 + u^2 - 2u \cos r)$$

$$\therefore S_1^2 = \frac{c^2}{1 + u^2 - 2u \cos r}$$

$$\therefore S_1^2 = \frac{a^2}{u^2 + v^2 - 2uv \cos \alpha} = \frac{b^2}{1 + v^2 - 2v \cos \beta} = \frac{c^2}{1 + u^2 - 2u \cos r}$$

parameterize  $u$  as  $v$ 

$$\Rightarrow A_4 v^4 + A_3 v^3 + A_2 v^2 + A_1 v + A_0 = 0$$

get 4 (up to)  $v$  solutionsget  $v \rightarrow$  solve  $u$  $\rightarrow$  solve  $S_1, S_2, S_3$ 

now we have two pairs of 3D points  
 $\rightarrow$  solve 3D-3D problem

EPnP- we need reference points  $P$  to do geometry

- any reference points  $P$   
 can be expressed by control points  
 $C_j \times 4$

$$\Rightarrow P_i^w = \sum_{j=1}^4 \alpha_{ij} C_j^w \quad w/ \sum_{j=1}^4 \alpha_{ij} = 1$$

known

unknown

side note 1:

- how to get  $C_j^w$ 

side note 2:

- why 4 points,  
 not 3, 2!

now try to get  $\alpha$  (intermediate media)

$$P_i^c = R_{cw} P_i^w + t$$

$$= R_{cw} \left( \sum_{j=1}^4 \alpha_{ij} C_j^w \right) + t$$

$$\text{as } \sum_{j=1}^4 \alpha_{ij} = 1$$

$$t = \sum_{j=1}^4 \alpha_{ij} t$$

$$\therefore P_i^c = R_{cw} \left( \sum_{j=1}^4 \alpha_{ij} C_j^w \right) + \sum_{j=1}^4 \alpha_{ij} t$$

$$= \sum_{j=1}^4 \left[ \alpha_{ij} (R_{cw} C_j^w + t) \right]$$

$$= \sum_{j=1}^4 \alpha_{ij} C_j^c$$

Both  $\{C\}$  &  $\{W\}$  share same  $\alpha$ 

$$P_i^w = \sum_{j=1}^4 \alpha_{ij} C_j^w$$

known

known

unknown

get unknown  $\alpha_{ij} \rightarrow$  known

$$P_i^c = \sum_{j=1}^4 \alpha_{ij} (R_{cw} C_j^w + t) = \sum_{j=1}^4 \alpha_{ij} C_j^c$$

unknown

known

our objective is to get  $P_i^c$ 

side note 3:

how to get

 $\alpha_{ij}$ ?



$$P_i^c = \sum_{j=1}^4 \alpha_{ij} (Row C_j^w + t) = \sum_{j=1}^4 \alpha_{ij} C_j^c$$

$$\Rightarrow w_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & 0 & u_c \\ 0 & f_v & v_c \\ 0 & 0 & 1 \end{bmatrix} \sum_{j=1}^4 \alpha_{ij} \begin{bmatrix} x_j^c \\ y_j^c \\ z_j^c \end{bmatrix}$$

From Last Row:

$$w_i = \sum_{j=1}^4 \alpha_{ij} z_j^c$$

$$\text{we get } \begin{cases} w_i u_i = f_u \sum_{j=1}^4 \alpha_{ij} x_j^c + u_c \sum_{j=1}^4 \alpha_{ij} z_j^c \\ w_i v_i = f_v \sum_{j=1}^4 \alpha_{ij} y_j^c + v_c \sum_{j=1}^4 \alpha_{ij} z_j^c \end{cases}$$

$$\Rightarrow \sum_{j=1}^4 [\alpha_{ij} f_u x_j^c + \alpha_{ij} (u_c - u_i) z_j^c] = 0$$

$$\sum_{j=1}^4 [\alpha_{ij} f_v y_j^c + \alpha_{ij} (v_c - v_i) z_j^c] = 0$$

we hv  $\alpha$  - K matrix,  $u_i, v_i$

get  $x_j^c, y_j^c, z_j^c \Rightarrow 4$  control pts  
3 dimension  
12 unknown

$$Mx = 0$$

$$x \in \mathbb{R}^{12}$$

$$M \in \mathbb{R}^{2n \times 12} \quad \text{no. of points } i=0, 1, \dots, n$$

solve  $x$  get  $C_j^c (j=1, 2, 3, 4)$

we then hv  $C_j^w, C_j^c \rightarrow 3D-3D$  (ICP problem)

we then also hv  $P_i^c = \sum_{j=1}^4 \alpha_{ij} C_j^c \rightarrow P_i^w P_i^c \rightarrow \text{ICP}$

side note 1:

how to get  $C_j^w$ ?

① as long as  $C_j^w$  is invertible

② as per Lepetit et al.: get centre of weight as  $C_j^w (j=1)$

$$C_{j=1}^w = \frac{1}{n} \sum_{i=1}^n P_i^w$$

$$\text{let } A = \begin{bmatrix} P_1^{wT} & \dots & C_1^{wT} \\ \vdots & & \vdots \\ P_n^{wT} & & C_1^{wT} \end{bmatrix}$$

$A^T A \xrightarrow{\text{get}} \lambda_{c,i} \text{ value } i=1, 2, 3$   
 $V_{c,i} \text{ vector } i=1, 2, 3$

$$C_j^w = C_{j=1}^w + \lambda_{c,j-1}^{\frac{1}{2}} V_{c,j-1} \quad j=2, 3, 4$$

side note 2:

why 4 - not others?

$$P_i^w = \begin{bmatrix} x_i^w \\ y_i^w \\ z_i^w \end{bmatrix} = [C_1^w \ C_2^w \ C_3^w] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

$$\sum_{i=1}^3 \alpha_{ij} = 1$$

4 equations w/ 3 unknown

overdefined.

$\therefore j=4$

side note 3:

how to get  $\alpha_{ij}$ ?

4 equations 4 unknowns

$$P_i^w = \sum_{j=1}^4 \alpha_{ij} C_j^w \quad w/ \sum_{j=1}^4 \alpha_{ij} = 1$$

solve linear system

remark 3D-2D  $4\alpha_j \rightarrow \text{varies}$   
 $4C_j \rightarrow \text{const}$

### ICP Problem

known data association

$$\begin{aligned} Y &= \{y_1, \dots, y_n\} \\ X &= \{x_1, \dots, x_n\} \end{aligned} \quad \} \quad C = \{i, j\}$$

$$\sum_{(i,j) \in C} \|y_i - Ax_j - t\|^2 \rightarrow \min$$

matched points pairs:

$$x_n, y_n$$

$$\bar{x}_n = Rx_n + t$$

$$\sum \|y_n - \bar{x}_n\|^2 p_n \rightarrow \min$$

exist & let's derive direct solution

$$y_0 = \frac{\sum y_n p_n}{\sum p_n} \quad x_0 = \frac{\sum x_n p_n}{\sum p_n}$$

$$H = \sum (y_n - y_0)(x_n - x_0)^T p_n$$

$$\text{sval}(H) = UDV^T$$

$$R = UV^T$$

$$t = y_0 - Rx_0$$

why is this a good solution?  $\rightarrow$

$$\bar{x}_n = R x_n + t$$

$$\bar{x}_n - y_0 = R x_n + t - y_0$$

$$\bar{x}_n - y_0 = R (x_n + R^T t - R^T y_0)$$

$$\bar{x}_n - y_0 = R (x_n - x_0) \quad \text{set as unknown variable } x_0$$

### ICP (SVD) (cont'd)

$$\sum \|y_n - \bar{x}_n\|^2 p_n \rightarrow \min$$

$$\Rightarrow \sum \|y_n - y_0 - \underbrace{R x_n - t + y_0}_{\bar{x}_n}\|^2 p_n \rightarrow \min$$

$$\Rightarrow \sum \|y_n - y_0 - (\bar{x}_n - y_0)\|^2 p_n \rightarrow \min$$

$$\Rightarrow \sum \|y_n - y_0 - R(x_n - x_0)\|^2 p_n \rightarrow \min$$

$$\therefore R^* x_0^* = \arg \min_{R, x_0} \sum \|y_n - y_0 - R(x_n - x_0)\|^2 p_n$$

$$\Phi(x_0, R) = \sum [(y_n - y_0) - R(x_n - x_0)]^T [(y_n - y_0) - R(x_n - x_0)] p_n$$

↓

$$\begin{aligned} \Phi(x_0, R) &= \sum (y_n - y_0)^T (y_n - y_0) p_n \quad \text{no } x_0, R \\ &\quad + \sum (x_n - x_0)^T (x_n - x_0) p_n \quad \text{no } R \\ &\quad - 2 \sum (y_n - y_0)^T R (x_n - x_0) p_n \end{aligned}$$

① w.r.t  $x_0$

$$\frac{\partial \Phi(x_0, R)}{\partial x_0} = -2 \sum (x_n - x_0) p_n + 2 \sum R^T (y_n - y_0) p_n$$

$$\text{set } \frac{\partial \Phi(x_0, R)}{\partial x_0} = 0$$

$$\therefore 0 = -2 \sum (x_n - x_0) p_n + 2 \sum R^T (y_n - y_0) p_n$$

$$\Rightarrow \sum (x_n - x_0) p_n = 0$$

$$\sum x_n p_n - \sum x_0 p_n = 0$$

$$\Rightarrow x_0 = \frac{\sum x_n p_n}{\sum p_n}$$

optimal value for

$x_0$  is the weighted mean of points  $x_n$

② w.r.t  $R$

$$R^* = \arg \min_R -2 \sum (y_n - y_0)^T R (x_n - x_0) p_n$$

$$= \arg \max_R 2 \sum (y_n - y_0)^T R (x_n - x_0) p_n$$

$$\text{let } b_n = y_n - y_0$$

$$a_n = x_n - x_0$$

$$R^* = \arg \max_R \sum b_n^T R a_n p_n$$

$$R^* = \arg \max_R \text{tr}(RH)$$

$$H = \sum (a_n b_n^T) p_n$$

find  $R$  max  $\text{tr}(RH)$

$$\text{SVD}(H) = U D V^T$$

$$U^T U = I$$

$$V^T V = I$$

$$D = \text{diag}(d_i)$$

$$R = V U^T$$

$$H = U D V^T$$

$$\text{tr}(RH) = \text{tr}(V U^T U D V^T)$$

$$= \text{tr}(V D V^T)$$

$$= \text{tr}(V D^T D V^T) \quad \text{set } A = V D^T$$

$$= \text{tr}(A A^T)$$

$$\therefore \text{tr}(RH) = \text{tr}(A A^T)$$

$$\text{tr}(A A^T) \geq \text{tr}(R' A A^T)$$

Schwarz inequality

$$\text{tr}(RH) = \text{tr}(A A^T) \geq \text{tr}(R' A A^T)$$

$$\text{tr}(R' R H)$$

any other rotation matrix

$$\therefore \text{choice } R = V U^T$$

↓  
optimal to maximize the trace

$R'R$  is also another  $\text{SO}(3)$

$$\text{let } R'R = R''$$

$$\text{tr}(R'' H) \text{ will always } < \text{tr}(RH)$$

$$\text{if } R = V U^T$$

i.e. if  $RH$  can be written as  $AA^T$

$\text{tr}(RH)$  will be max.



## What's SVD

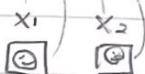
- for data deduction
- data-driven generalization of Fourier Transform (FFT)
- "tailored" to specific problem
- solve  $Ax = b$

for non-square  $A$

- regression
- PCA
- correlation

$$\bar{X} = \begin{bmatrix} | & | & | & \dots & | \\ x_1 & x_2 & x_3 & \dots & x_m \\ | & | & | & \dots & | \end{bmatrix}_{n \times m}$$

$$x_k \in \mathbb{R}^n$$



$$\bar{X} = \begin{bmatrix} | & | & | & \dots & | \\ x_1 & x_2 & x_3 & \dots & x_m \\ | & | & | & \dots & | \end{bmatrix}_{n \times m} = U \Sigma V^T$$

eigen importances mixtures of U's to make X's

$$U \Sigma V^T = \begin{bmatrix} | & | & | & \dots & | \\ u_1 & u_2 & u_3 & \dots & u_n \\ | & | & | & \dots & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \sigma_3 & & \\ & & & \ddots & \\ & & & & \sigma_m \\ & & & & & 0 \end{bmatrix} \begin{bmatrix} | & | & | & \dots & | \\ v_1 & v_2 & v_3 & \dots & v_m \\ | & | & | & \dots & | \end{bmatrix}^T$$

eigen faces order by importance eigen time series eigen mixtures left singular vectors singular values right singular vectors

-  $U, V$  unitary

normal length orthogonal

$$U^T U = U U^T = I_{n \times n}$$

$$V V^T = V^T V = I_{m \times m}$$

$$\Sigma \text{ diagonal } \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$$

$$\bar{X} = \begin{bmatrix} | & | & | & \dots & | \\ x_1 & x_2 & x_3 & \dots & x_m \\ | & | & | & \dots & | \end{bmatrix} = U \Sigma V^T = \begin{bmatrix} | & | & | & \dots & | \\ u_1 & u_2 & u_3 & \dots & u_n \\ | & | & | & \dots & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \sigma_3 & & \\ & & & \ddots & \\ & & & & \sigma_m \\ & & & & & 0 \end{bmatrix} \begin{bmatrix} | & | & | & \dots & | \\ v_1 & v_2 & v_3 & \dots & v_m \\ | & | & | & \dots & | \end{bmatrix}^T$$

$n \gg m$ , usually

$$= u_1 \sigma_1 v_1^T + u_2 \sigma_2 v_2^T + \dots + u_m \sigma_m v_m^T + 0$$

$$= \hat{U} \hat{\Sigma} \hat{V}^T \text{ (economy SVD)}$$

$$[U, \Sigma, V] = \text{svol}(X, \text{'econ'})$$

$$= \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_m u_m v_m^T + 0$$

$$= \sigma_1 \begin{bmatrix} | & | & | & \dots & | \\ u_1 & u_2 & u_3 & \dots & u_m \\ | & | & | & \dots & | \end{bmatrix} \begin{bmatrix} | & | & | & \dots & | \\ v_1 & v_2 & v_3 & \dots & v_m \\ | & | & | & \dots & | \end{bmatrix}^T + \dots$$

$$= \tilde{U} \tilde{\Sigma} \tilde{V}^T \text{ (Eckart-Yung [1936] Theory)}$$

truncate at rank  $r$

$$\arg \min_{\tilde{X} \text{ s.t. } \text{rank}(\tilde{X})=r} \|\bar{X} - \tilde{X}\|_F = \tilde{U} \tilde{\Sigma} \tilde{V}^T$$

SVD + correlation

$$\bar{X}^T \bar{X} = \frac{1}{m} \begin{bmatrix} x_1^T & x_2^T & \dots & x_m^T \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \dots & x_m \end{bmatrix} = \frac{1}{m} \begin{bmatrix} x_1^T x_1 & x_1^T x_2 & \dots & x_1^T x_m \\ x_2^T x_1 & x_2^T x_2 & \dots & x_2^T x_m \\ \vdots & \vdots & \ddots & \vdots \\ x_m^T x_1 & x_m^T x_2 & \dots & x_m^T x_m \end{bmatrix}$$

correlation matrix

$$= \begin{bmatrix} x_1^T x_1 & x_1^T x_2 & \dots & x_1^T x_m \\ x_2^T x_1 & x_2^T x_2 & \dots & x_2^T x_m \\ \vdots & \vdots & \ddots & \vdots \\ x_m^T x_1 & x_m^T x_2 & \dots & x_m^T x_m \end{bmatrix}$$

$$x_i^T x_j = \langle x_i, x_j \rangle$$

$$\bar{X}^T \bar{X} = V \hat{\Sigma} U^T U \hat{\Sigma} V^T$$

$$\bar{X} = \tilde{U} \tilde{\Sigma} \tilde{V}^T$$

$$\bar{X}^T = \tilde{V} \tilde{\Sigma}^T \tilde{U}^T$$

$$= V \hat{\Sigma}^2 V^T \text{ (eigen decomposition)}$$

$$\bar{X}^T \bar{X} V = V \hat{\Sigma}^2 \rightarrow \text{eigenvalues}$$

$$\bar{X} \bar{X}^T = \tilde{U} \tilde{\Sigma} \tilde{V}^T \tilde{V} \tilde{\Sigma} \tilde{U}^T = \tilde{U} \tilde{\Sigma}^2 \tilde{U}^T$$

$$\bar{X} \bar{X}^T \tilde{U} = \tilde{U} \hat{\Sigma}^2 \rightarrow \text{eigenvalues}$$

ICP unknown data association

corresponding points  $y_n, x_n \quad n=1, \dots, N$   
weights  $p_n \quad n=1, \dots, N$   
find  $R, t$  s.t.

$$\bar{x}_n = R x_n + t \quad n=1, \dots, N$$

so that

$$\sum \|y_n - \bar{x}_n\|^2 p_n \rightarrow \min$$

Correspondences unknown

guess  $\rightarrow$  point locations or point correspondences  
iteration

ICP (Vanilla)

$$\bar{x}_n = x_n$$

$$e = \infty$$

while  $e > \text{thres}$

$$C = \text{correspondences}(\{y_n, \bar{x}_n\})$$

$$(t, R) = \text{compute } T(C)$$

$$\bar{x}_n = R(x_n - x_0) + y_0$$

$$e = E(t, R) = \frac{1}{2} (R^T y_0 - R^T t, R)$$

return  $\{\bar{x}_n\}$



if easy  $\rightarrow$  analytical form  
 $\rightarrow$  through  $\frac{dF}{dx} = 0$   
 $\rightarrow$  else

### Non-linear Least-Square Problem

$$\min_x F(x) = \frac{1}{2} \|f(x)\|_2^2 \quad f(x): \mathbb{R}^n \rightarrow \mathbb{R}$$

objective: get  $X$  s.t.  $F(x)$  has min  
 optimization algo

while true

at  $i = k$

search  $\Delta X_k$

get  $\|f(X_k + \Delta X_k)\|_2^2$

if  $\Delta X_k < b$

return

else

continue

$$X_{k+1} = X_k + \Delta X_k$$

How to get  $\Delta X_k$ ?

A: different tricks

Here introduce

### Gauss-Newton Method

$$f(x) = \frac{1}{2} \sum_{j=1}^m r_j(x)^2 \quad \begin{cases} r_j(x) = \phi(x; \tau_j) - y_j \\ j = 1, 2, 3, \dots, m \\ r(x) = [r_1(x), r_2(x), \dots, r_m(x)]^T \end{cases}$$

$$= \frac{1}{2} \sum_{j=1}^m \|r(x)\|_2^2$$

$x$  is the objective parameter  
 (e.g. pose)  $x \in \mathbb{R}^n$

$r$  is the residual result calculated  
 from  $x$  (e.g. reprojection  $u, v$ )  
 $m = 1, 2$

math preliminaries

$$\nabla f(x_1, x_2) = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right) \quad J(x) = \left[ \frac{\partial r_j}{\partial x_i} \right]_{i=1, \dots, n}^{j=1, \dots, m}$$

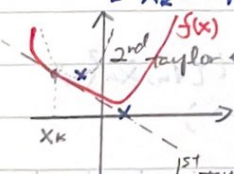
$$H(f(x_1, x_2)) = \left( \frac{\partial^2 f}{\partial x_i^2}, \frac{\partial^2 f}{\partial x_i \partial x_j}, \frac{\partial^2 f}{\partial x_j^2} \right) = \begin{bmatrix} \nabla r_1(x)^T \\ \nabla r_2(x)^T \\ \vdots \\ \nabla r_m(x)^T \end{bmatrix}$$

### recall Newton's Method

$$X_{k+1} = X_k + \Delta X_k$$

$$\Rightarrow X_{k+1} = X_k - [J^T J(X_k)]^{-1} J^T f(X_k)$$

$$= X_k - H^{-1} J_r(X_k)^T r(X_k)$$



2<sup>nd</sup> Taylor expansion  $f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2} f''(x_k)(x - x_k)^2$

1<sup>st</sup> Taylor expansion  $f(x) = f(x_k) + f'(x_k)(x - x_k)$

not a good value for  $\Delta X_k$ , as it could over-shoot exaggeratingly

$$1^{st} \text{ order: } f(x_k) + f'(x_k)(x - x_k) = 0$$

$$2^{nd} \text{ order: } f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2} f''(x_k)(x - x_k)^2 = 0$$

$$g(x)$$

solve  $x$  value,

and set it as next  $X_k$ .

Newton's method opt the second 2<sup>nd</sup> order.

$$\frac{d}{dx} g(x) = f'(x_k) + f''(x_k)(x - x_k)$$

$$x^* = \underset{x}{\operatorname{argmin}} g(x)$$

$$\therefore 0 = \frac{d}{dx} g(x) = f'(x_k) + f''(x_k)(x - x_k)$$

$$\Delta x = \frac{-f'(x_k)}{f''(x_k)}$$

$$\therefore X_{k+1} = X_k - [J^T J(X_k)]^{-1} J^T f(X_k)$$

$$= X_k - H^{-1} J_r(X_k)^T r(X_k)$$

yet: still greedy

Hessian matrix not easy to solve

### Gauss-Newton Method

$$w/ \nabla f = J_r^T r$$

$$\nabla^2 f = J_r^T J_r + \sum_{j=1}^m r_j \nabla^2 r_j$$

$$= J_r^T J_r + Q$$

$$\nabla^2 f \approx J_r^T J_r$$

substitute  $H$  as  $J^T J$

$$\therefore X_{k+1} = X_k - [J_r(X_k)^T J_r(X_k)]^{-1} J_r(X_k)^T r(X_k)$$

$$= X_k - \frac{J_r(X_k)^T r(X_k)}{\Delta x}$$

$$\Delta x = -[J_r(X_k)^T J_r(X_k)]^{-1} J_r(X_k)^T r(X_k)$$

$$\nabla^2 f(x) = \nabla^2 \left( \frac{1}{2} \sum_{j=1}^m r_j(x)^2 \right)$$

$$= \sum_{j=1}^m \nabla r_j(x) \nabla r_j(x)^T + \sum_{j=1}^m r_j(x) \nabla^2 r_j(x)$$

$$= J(x)^T J(x) + Q(x)$$

much better. at least we hv a minimum value (quadratic)