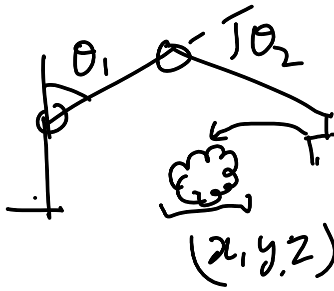
Forward and $\boxed{\text{inverse kinematics}}$

What should the joint ~~angles~~ state/conf. of the robot be so that the end-effector reaches a desired pose?



[How to move the end-effector to a desired pose [position + orientation]]

← end-effector

gripper
or suction cup

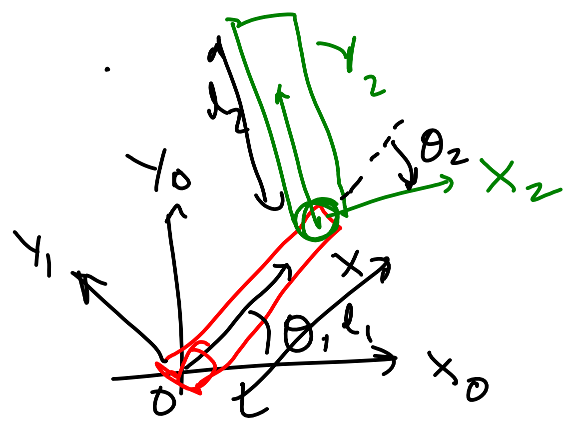
Forward kine.

If my joint ~~angles~~ state/conf. are given what would the pose of end-effector be?

Forward kinematics

$$ {}^{0}T_{2} = {}^{0}T_{1}(\theta_1, l_1)\,{}^{1}T_{2}(\theta_2, l_2) $$

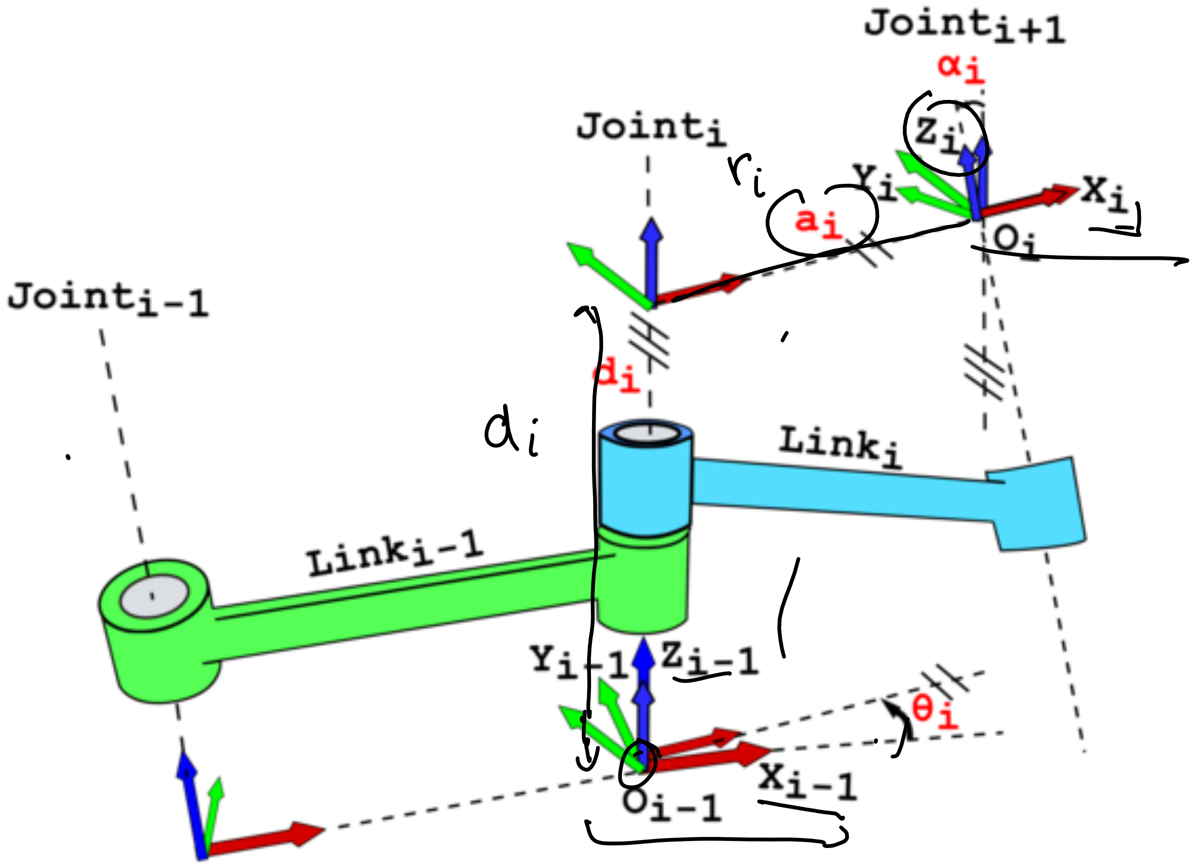
in terms of $\underbrace{\theta_1 \text{ and } \theta_2}_{\text{Given}}$



(Denavit Hartenberg Parameters / Convention)

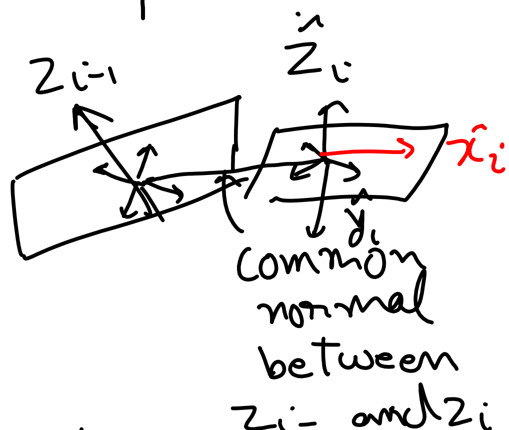
Denavit Hartenberg parameters

https://www.youtube.com/watch?v=rA9tm0gTln8



① $\hat{z}_{i-1}, \hat{z}_i$ aligned along the axis of rotation

② Choose $\hat{x}_i$ along the common normal between $\hat{z}_{i-1}, \hat{z}_i$



③ $\hat{y}_i = \hat{z}_i \times \hat{x}_i$

a) θ_i = Rotation along z_{i-1} (to align x_{i-1} with x_i)

b) d_i = translation along z_{i-1} (to align the origins)

c) α_i = Rotation along x_i (to align z_{i-1} with z_i)

d) r_i/a_i = translation along x_i (to align the origin)

(a) and (b) can be swapped
(c) and (d)

But Tranformation along z goes first
followed by " " x

$$\underset{\uparrow \atop \text{target}}{{}^{i-1}} T_{\underset{\uparrow \atop \text{source}}{i}} = {}^{i-1}T_{zi} \; {}^{i-1}T_{xi}$$

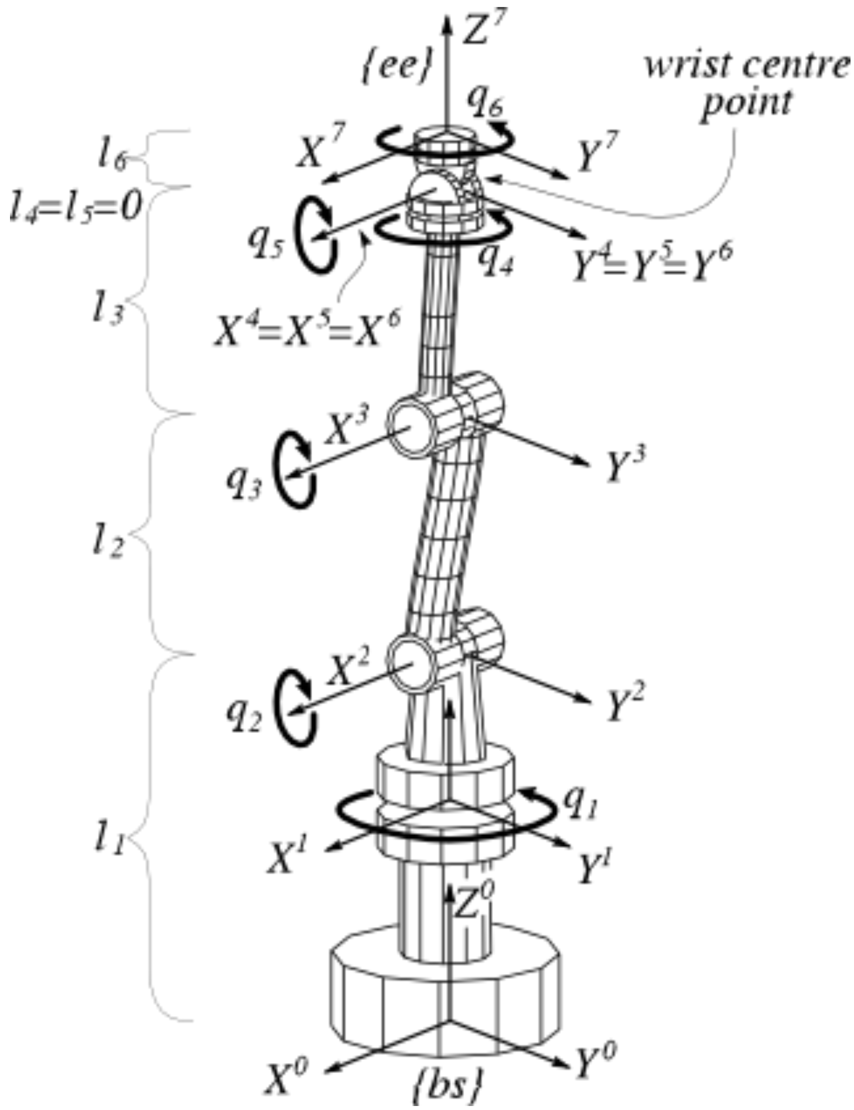
Tranformations are applied
right to left

$$ {}^{i-1}T_{xi} = \left[\begin{array}{ccc|c} 1 & 0 & 0 & r_i \\ 0 & \cos\alpha_i & -\sin\alpha_i & 0 \\ 0 & \sin\alpha_i & \cos\alpha_i & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]_{4\times 1} $$

3x3 3x1

$$ {}^{i-1}T_{zi} = \begin{bmatrix} \cos\theta_i & -\sin\theta_i & 0 & 0 \\ \sin\theta_i & \cos\theta_i & 0 & 0 \\ 0 & 0 & 1 & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix}_{4\times 1} $$

θ_i, d_i



Z7
{ee}
wrist centre
point
q6
l6
X7
Y7
l4=l5=0
q5
q4
Y4=Y5=Y6
l3
X4=X5=X6
X3
q3
Y3
l2
X2
q2
Y2
q1
l1
X1
Y1
Z0
X0
{bs}
Y0

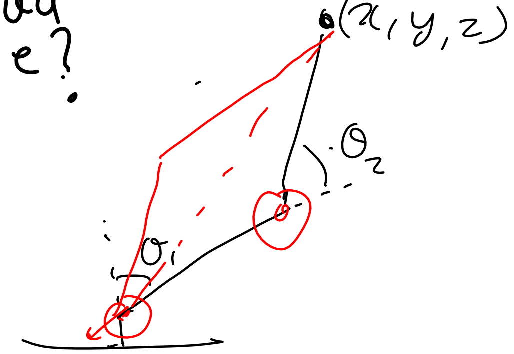
Numerical solutions to IK problems: Jacobian inverse technique

Forward and inverse kinematics

Inverse kinematics
↳ closed form solution
↳ Numerical / Iterative solutions

only polynomials of degree ≤ 5 have closed form solutions

What should the joint ~~angles~~ state/conf. of the robot be so that the end-effector reaches a desired pose?



$\cancel{\theta_1 = x_1^2 + y_2 + \cos^{-1}\left(\frac{z}{y}\right)}$

$\cos(\theta_1)$ $\quad$ $\underbrace{\sin(\theta_1)}_{x_1}$

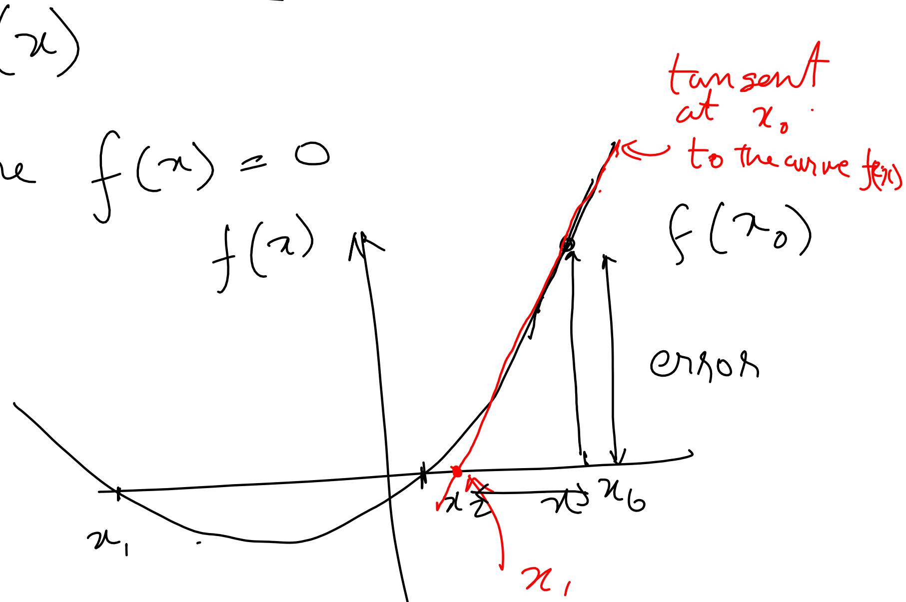
$\sqrt{1 - x_1^2}$

Newton-Raphson method (Gradient descent) (optimization solution)

Suppose a function $y = f(x)$
to find

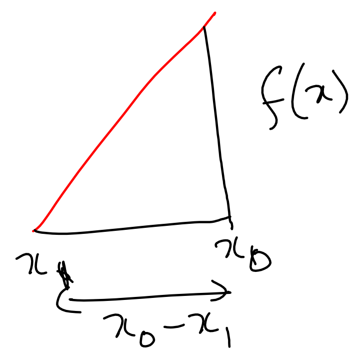
we want ~~all~~ any x where $f(x) = 0$

① Initial guess
x_0
↑
iteration



② Improve the initial guess

$$f'(x)\Big|_{x_0} = \frac{f(x_0)}{x_0 - x_1}$$



$$\Rightarrow \quad x_0 - x_1 = \left[f'(x_0)\right]^{-1} f(x_0)$$

$$x_1 = x_0 - \left[f'(x_0)\right]^{-1} f(x_0)$$

② Repeat

$$x_2 = x_1 - \left[f'(x_1)\right]^{-1} f(x_1)$$

$$x_n = x_{n-1} - \left[f'(x_{n-1})\right]^{-1} f(x_{n-1})$$

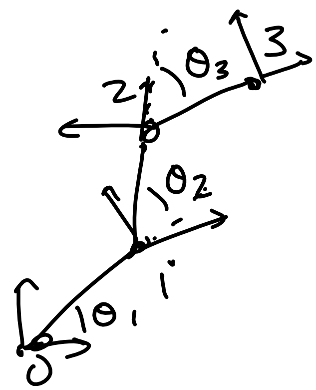
---

find the square root of 2

$$f(z) = z^2 - 2 = 0$$

$$z_n = z_{n-1} - \frac{(z_{n-1}^2 - 2)}{2 z_{n-1}}$$

Forward Kinematics



$$\underbrace{{}^0T_3}_{\text{Given}} = {}^0T_1(\theta_1)\; {}^1T_2(\theta_2)\; {}^2T_3(\theta_3)$$

Find

$$\underbrace{({}^0T_3)^{-1}\, {}^0T_1(\theta_1)\; {}^1T_2(\theta_2)\; {}^2T_3(\theta_3) - I_{4\times4}} = 0_{4\times4}$$

$$\underbrace{F_{4\times4}(\theta_1, \theta_2, \theta_3)} = 0_{4\times4}$$

$$\underline{f}_{16\times1}\left(\underbrace{\begin{bmatrix}\theta_1\\ \theta_2\\ \theta_3\end{bmatrix}_{3\times1}}_{\text{vector}}\right) = \begin{bmatrix}0\\ \\ 0\end{bmatrix}_{16\times1}$$

$f(\underline{\theta})$ is a vector function

$$\underline{f}_{16\times1}(\underline{\theta}) = \begin{bmatrix} f_1(\underline{\theta}) \\ f_2(\underline{\theta}) \\ \vdots \\ f_{16}(\underline{\theta}) \end{bmatrix}$$

vector-valued vector function

What's a Jacobian matrix

is ~~a~~ many derivative of vector-valued vector function

$$
\begin{array}{cc}
 & \text{inputs} \\
J\left\{\underline{f}_{16\times 1}\left(\underline{\theta}_{3\times 1}\right)\right\} = \ \text{outputs} &
\begin{bmatrix}
\frac{\partial f_1}{\partial \theta_1} & \frac{\partial f_1}{\partial \theta_2} & \cdots & \frac{\partial f_1}{\partial \theta_3} \\
\frac{\partial f_2}{\partial \theta_1} & \ddots & & \\
\vdots & & \ddots & \\
\frac{\partial f_{16}}{\partial \theta_1} & & & \ddots
\end{bmatrix}
\end{array}
$$

Hand-waviness

scalar Newton Raphson

generalizes to (vector Newton - Raphson)

Gauss Newton method

(1) Initial guess $\underline{\theta}_0$

Psuedo-inverse dagger

(2) $\underline{\theta}_1 = \underline{\theta}_0 - J\left[\underline{f}_{16\times 1}\left(\underline{\theta}_{3\times 1}\right)\right]^{\dagger} \underline{f}_{16\times 1}\left(\underline{\theta}_{3\times 1}\right)$

$$x_1 = x_0 - [f'(x_0)]^{-1} f(x_0)$$

(3) $$\underline{\theta}_n = \underline{\theta}_{n-1} - J[\underline{f}(\underline{\theta}_{n-1})]^{+} \underline{f}(\underline{\theta}_{n-1})$$

(1) What is a Pseudo inverse?

→ (2) How can we compute the Jacobian?

(3) What is the relationship (similarities and differences) b/w Newton-Raphson, Gauss Newton, Gradient descents

---

Problem 4 of Midterm helps in computing derivative of rotation matrices

$$K^3 = -K$$

$$R(\theta, \hat{k}) = I_{3\times3} + \sin\theta\, K + (1-\cos\theta) K^2$$

$$\frac{\partial}{\partial\theta} R(\theta, \hat{k}) = 0 + \cos\theta\, \underset{\sim}{K} + (0+\sin\theta) K^2$$

$$= -\cos\theta\, K^3 + \sin\theta\, K^2$$

$$= K(-\cos\theta\, K^2 + \sin\theta\, K)$$

$$= K(I - I - \cos\theta\, K^2 + \sin\theta\, K)$$

$$= K\left(I - \cos\theta\, K^2 + \sin\theta\, K\right) - K$$

$$= K\left(I + K^2 - K^2 - \cos\theta\, K^2 + \sin\theta\, K\right) - K$$

$$= K\left(I + (1-\cos\theta)K^2 + \sin\theta\, K\right) - K - K^3$$

$$= K\,R(\theta,\hat{k}) - \cancel{K} + \cancel{K}$$

$$\boxed{\frac{\partial}{\partial\theta} R(\theta,\hat{k}) = K\,R(\theta,\hat{k})} \qquad \boxed{\frac{d}{dx} f(x) = a f(x)}$$

$$\downarrow \uparrow$$

scalar matrix

$$R(\theta,\hat{k}) = \exp(\theta K) = \frac{I}{1} + \frac{\theta K^1}{1!} + \frac{\theta^2 K^2}{2!} + \frac{\theta^3 K^3}{3!} + \dots \infty$$

matrix exponention

$$\frac{\partial}{\partial\theta} R(\theta,\hat{k}) = 0 + \frac{K^1}{1!} + \frac{2\theta K^2}{2!} + \frac{3\theta^2 K^3}{3!} + \dots$$

$$= K\left(\frac{I}{1!} + \frac{\theta K}{1!} + \frac{\theta^2 K^2}{2!} + \dots\right)$$

$$T = \begin{bmatrix} R(\theta,\hat{k})_{3\times3} & \underline{t}_{3\times1} \\ \underline{0}^T & 1 \end{bmatrix}$$

$$\frac{\partial T}{\partial\theta} = \begin{bmatrix} K\,R(\theta,\hat{k}) & 0 \\ 0 & 0 \end{bmatrix} \qquad K = [\underline{\hat{k}}]_\times$$

Jacobian of forward kinematics

$${}^{0}T_{3} = \underbrace{{}^{0}T_{1}(\theta_1)\ {}^{1}T_{2}(\theta_2)\ {}^{2}T_{3}(\theta_3)}$$

$$\boxed{\frac{\partial\, {}^{0}T_{3}}{\partial \theta_1} = \left[\frac{\partial\, {}^{0}T_{1}(\theta_1)}{\partial \theta_1}\right] {}^{1}T_{2}(\theta_2)\ {}^{2}T_{3}(\theta_3)}$$

$$\frac{\partial}{\partial \theta} A(\theta) B(\theta) = \left[\frac{\partial A(\theta)}{\partial \theta}\right] B(\theta) + A(\theta)\left[\frac{\partial B(\theta)}{\partial \theta}\right]$$

scalar

$$C(\theta) = A(\theta) B(\theta)$$

$$\begin{bmatrix} c_{11} & \cdots \\ \vdots & \end{bmatrix} = \begin{bmatrix} a_{11}(\theta) & \cdots & a_{1n}(\theta) \\ & & \end{bmatrix} \begin{bmatrix} b_{11}(\theta) & \\ \vdots & \\ b_{n1}(\theta) & \end{bmatrix}$$

$$c_{11} = a_{11}(\theta) b_{11}(\theta) + a_{12}(\theta) b_{21}(\theta) + \ldots + a_{1n}(\theta) b_{n1}(\theta)$$

$$\frac{\partial c_{11}}{\partial \theta} = \left\{\left[\frac{\partial a_{11}(\theta)}{\partial \theta}\right] b_{11}(\theta) + a_{11}(\theta)\left[\frac{\partial}{\partial \theta} b_{11}(\theta)\right]\right\}$$

$$+ \quad \ldots$$

$$\left\{\frac{\partial ({}^{0}T_3)}{\partial \theta_2}\right\}_{4\times4} = {}^{0}T_1(\theta_1)\left[\frac{\partial}{\partial \theta_2}\,{}^{1}T_2(\theta_2)\right]{}^{2}T_3(\theta_3)$$

$$\mathrm{vec}({}^{0}T_3)_{4\times4} = [\quad]_{16\times1} = \text{row-wise vectorization operator}$$

$$J_{\underline{\theta}}\left[\mathrm{vec}({}^{0}T_3)\right] = \begin{bmatrix} | & | & | \\ \frac{\partial\,\mathrm{vec}({}^{0}T_3)}{\partial\theta_1} & \frac{\partial\,\mathrm{vec}({}^{0}T_3)}{\partial\theta_2} & \frac{\partial\,\mathrm{vec}({}^{0}T_3)}{\partial\theta_3} \\ | & | & | \end{bmatrix}_{16\times3}$$

12×3

$$\frac{\partial}{\partial\theta_1}\,{}^{0}T_1(\theta_1) = \left[\frac{\partial T_z(\theta,d)}{\partial\theta_1}\right]\left[T_x(r,\alpha)\right]$$

$$= \begin{bmatrix} K_z R_z(\theta) & \underline{0}'_{3\times} \\ \underline{0}^{T}_{1\times3} & 0' \end{bmatrix} T_x(r,\alpha)$$

$$ {}^{0}T_3 = \begin{bmatrix} & \\ 0^{T}_{1\times3} & 1 \end{bmatrix}$$

we know how to compute Jacobian ✓

$$\underline{\theta}_n = \underline{\theta}_{n-1} - J\left[\underline{f}(\underline{\theta}_{n-1})\right]^{\dagger} \underbrace{\underline{f}(\theta_{n-1})}_{\text{error}}$$

$$\underline{f}(\underline{\theta}_{n-1}) = \mathrm{vec}\left({}^{0}\underline{T}_3(\underline{\theta}_{n-1}) - \boxed{{}^{0}T_3^{*}}\right)$$

Desired end-effector pose

equivalent in terms of roots

$$\underline{f}(\underline{\theta}_{n-1}) = \mathrm{vec}\left(\left[{}^{0}T_3^{*}\right]^{-1} {}^{0}T_3(\underline{\theta}_{n-1}) - I_{4\times4}\right)$$

$$J\left[\underline{f}(\underline{\theta}_{n-1})\right] = J\left[\mathrm{vec}\left[{}^{0}T_3(\underline{\theta}_{n-1})\right]\right]$$

$$\underline{\theta}_n = \underline{\theta}_{n-1} - J\left[\mathrm{vec}\left({}^{0}T_3(\theta_{n-1})\right)\right]^{\dagger} \mathrm{vec}\left({}^{0}\underline{T}_3(\underline{\theta}_{n-1}) - \boxed{{}^{0}T_3^{*}}\right)$$

Pseudo Inverse

systems of eqns in multiple variable

① $\boxed{\begin{matrix} 2x + 3y = 9 \\ 5x + 7y = 9 \end{matrix}}$ $\begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \end{bmatrix}$ 1 solution

$A\underline{x} = \underline{b}$ $\underline{x} = A^{-1}\underline{b}$

No. of eqns = No of unknowns

② No of eqns > No of unknowns] no solution

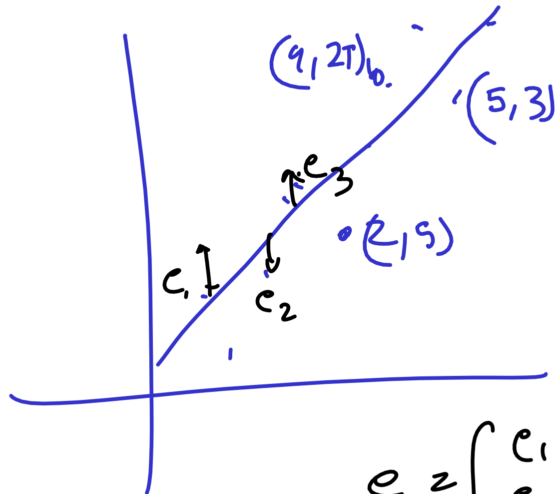
③ " < "] multiple solution

(2) No of eqns > No of unknowns

$$\begin{bmatrix} 2x + 3y \\ 5x + 7y \\ 6x + 13y \end{bmatrix} \stackrel{?}{=} \begin{bmatrix} 4 \\ 9 \\ 13 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} 2 & 3 \\ 5 & 7 \\ 6 & 13 \end{bmatrix}}_{A_{3\times 2}} \begin{bmatrix} x \\ y \end{bmatrix} \stackrel{?}{=} \begin{bmatrix} 4 \\ 9 \\ 13 \end{bmatrix}$$

$$\underline{x}_{2\times 1} = \underline{b}_{3\times 1}$$



$$\underline{e} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \end{bmatrix}$$

$$\left.\begin{aligned} 5 &= 2m + c \\ 27 &= 9m + c \\ 3 &= 5m + c \end{aligned}\right\} \longrightarrow \text{Least square linear regression}$$

$$A_{m\times n}\, \underline{x}_{n\times 1} \approx \underline{b}_{m\times 1}$$

\# of eqns, \# of unknowns

$m > n$

Linear Least square

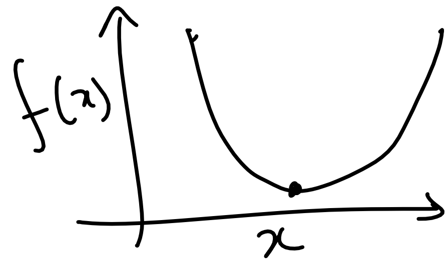
$$\begin{bmatrix} \quad \end{bmatrix}_{m\times n}$$

tall matrix

$$\boxed{\min_{\underline{x}} \|\underbrace{A\underline{x} - \underline{b}}_{\underline{e}}\|^2}$$

$$|\underline{e}|^2 = e_1^2 + e_2^2 + e_3^2 + \ldots$$

Background



① $\frac{d}{dx} f(x) = 0$

$\downarrow$ vector

Scalar valued vector function $f(\underline{x})$

$$J_{\underline{x}}[f(\underline{x})] = \underbrace{\left[\frac{\partial}{\partial x_1} f(\underline{x}), \; \cdots\cdots \; \frac{\partial}{\partial x_n} f(\underline{x}) \right]}_{\text{Gradient}} = \underline{0}^T$$

② $\|\underline{c}\| = \sqrt{\underline{c}^T \underline{c}}$

③ Chain rule with Jacobians

$$\frac{d}{dx} f(g(x)) = \frac{d}{dg} f(g) \, \frac{d}{dx} g(x)$$

$$f(\underline{y}) : \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$\underline{g}(\underline{x}) : \mathbb{R}^m \longrightarrow \mathbb{R}^n$$

$$J_{\underline{x}}[f(\underline{g}(\underline{x}))] = \underbrace{J_{\underline{g}}[f(\underline{g})]}_{1 \times n} \, \underbrace{J_{\underline{x}}[\underline{g}(\underline{x})]}_{n \times m}$$

Example $n = 2$, $m = 1$

$$J_x[f(\underline{g}(x))] = \left[\frac{\partial}{\partial x} f(\underline{g}(x))\right]_{1\times 1}$$

$$\rightarrow \left[\frac{\partial}{\partial x} f\left(\begin{bmatrix} g_1(x) \\ g_2(x) \end{bmatrix}\right)\right] = \frac{\partial}{\partial x} f(g_1(x), g_2(x))$$

$$= \frac{\partial}{\partial g_1} f(\underline{g}(x)) \frac{\partial g_1(x)}{\partial x} + \frac{\partial}{\partial g_2} f(\underline{g}(x)) \frac{\partial g_2(x)}{\partial x}$$

$$= \underbrace{\begin{bmatrix} \frac{\partial}{\partial g_1} f(\underline{g}(x)) & \frac{\partial}{\partial g_2} f(\underline{g}(x)) \end{bmatrix}}_{J_{\underline{g}}[f(\underline{g})]} \underbrace{\begin{bmatrix} \frac{\partial g_1(x)}{\partial x} \\ \frac{\partial g_2(x)}{\partial x} \end{bmatrix}}_{J_x[\underline{g}(x)]}$$

(4) $$J_{\underline{x}}[\underbrace{\underline{a}^T \underline{x}}_{f(\underline{x})}] = \left[\frac{\partial(\underline{a}^T \underline{x})}{\partial x_1} \cdots\cdots \frac{\partial \underline{a}^T \underline{x}}{\partial x_n}\right]_{1\times n} \qquad \underline{x} \in \mathbb{R}^{n\times 1}, \quad \underline{a} \in \mathbb{R}^{n\times 1}$$

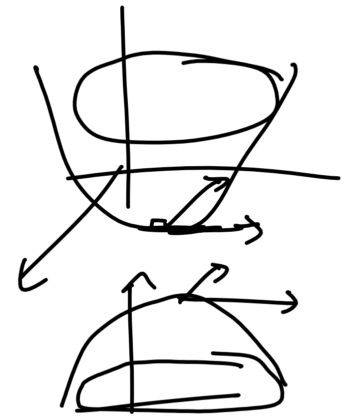
$$\underline{a}^T \underline{x} = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

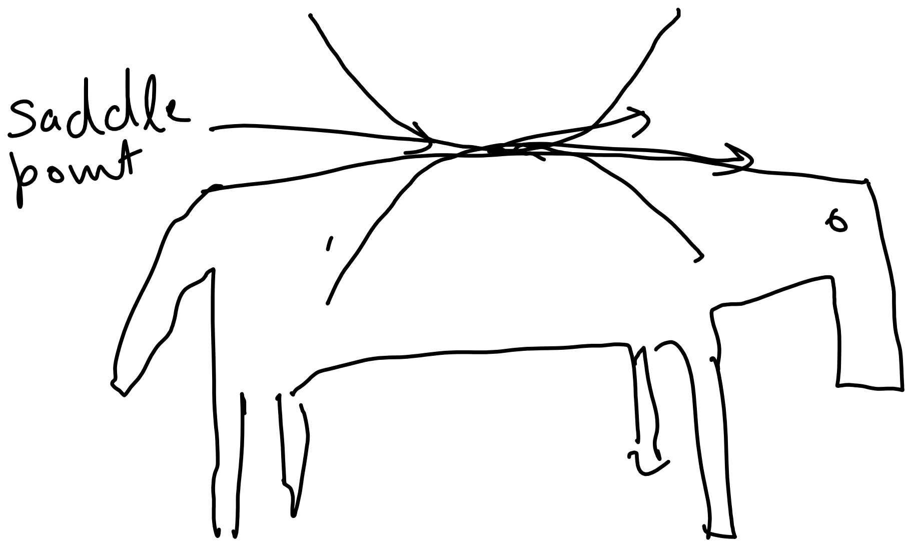
$$J_{\underline{x}}[\underline{a}^T \underline{x}] = [a_1, a_2 \;\cdots\; a_n] = \underline{a}^T$$

$$J_{\underline{x}}[A\underline{x}] = J_{\underline{x}}\begin{bmatrix} \underline{a}_1^T \underline{x} \\ \vdots \\ \underline{a}_m^T \cdot \underline{x} \end{bmatrix} = \begin{bmatrix} \underline{a}_1^T \\ \vdots \\ \underline{a}_m^T \end{bmatrix} = A \qquad A = \begin{bmatrix} - \underline{a}_1^T - \\ - \underline{a}_2^T - \\ \vdots \\ - \underline{a}_m^T - \end{bmatrix}$$

1. The minimum / maximum / saddle points of a function $f(\underline{x})$ occur when

$$J_{\underline{x}}[f(\underline{x})] = \underline{0}^T$$





2. $\|\underline{e}\| = \sqrt{\underline{e}^T \underline{e}}$

3. $J_{\underline{x}} f(\underline{g}(\underline{x})) = J_{\underline{g}}[f(\underline{g})]\, J_{\underline{x}}[\underline{g}(\underline{x})]$

4. $J_{\underline{x}}[A\underline{x}] = A$

$$\min_{\underline{x}} \|A\underline{x} - b\|^2$$

$$\min_{\underline{x}} (A\underline{x} - b)^T (A\underline{x} - b)$$

Using (2)

5. $J_{\underline{x}}[\underline{x}^T A \underline{x}] = \underline{x}^T (A + A^T)$

Quadratic form

Proof left as an exercise

$$\underline{x}^T A \underline{x} = \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= a_{11} x^2 + a_{12} xy + a_{21} xy + a_{22} y^2$$

if

A is symmetric matrix, $A^T = A$

$$J_x [\underline{x}^T A \underline{x}] = x^T (A^T + A) = 2\underline{x}^T A$$

. Special case

$$J_{\underline{x}} [\underline{x}^T \underline{x}] = J_{\underline{x}} [\underline{x}^T \underline{I} \underline{x}] = 2 \underline{x}^T$$

—

$$\min_{\underline{x}} \underbrace{(A\underline{x} - b)}_{\underline{e}(x)}^T (A\underline{x} - b)$$

$$\min_{\underline{x}} [\underline{e}(\underline{x})^T \underline{e}(\underline{x})]$$ where $e(\underline{x}) = A\underline{x} - \underline{b}$

$$J_{\underline{x}} [\underline{e}(x)^T \underline{e}(\underline{x})] = \underline{0}^T$$

$$J_{\underline{e}} [\underline{e}(\underline{x})^T \underline{e}(\underline{x})] \, J_{\underline{x}} [\underline{e}(\underline{x})] = \underline{0}^T$$

$$2\underline{e}(\underline{x})^T J_{\underline{x}} [A\underline{x} - \underline{b}] = \underline{0}^T$$

$$2\underline{e}(\underline{x})^T \left(J_{\underline{x}} [A\underline{x}] - J_{\underline{x}} [\underline{b}] \right) = \underline{0}^T$$

$$2\,\underline{e}(\underline{x})^T\left[A - O_{m\times n}\right] = \underline{0}^T$$

$$2(A\underline{x} - \underline{b})^T A = \underline{0}^T$$

$$\Rightarrow 2\left((A\underline{x})^T - \underline{b}^T\right) A = \underline{0}^T$$

$$\Rightarrow 2\left(\underline{x}^T A^T - \underline{b}^T\right) A = \underline{0}^T$$

$$\Rightarrow 2\left(\underline{x}^T A^T A - b^T A\right) = \underline{0}^T$$

$$\Rightarrow x^T \underbrace{A^T A} = b^T A$$

$$\Rightarrow \underline{x}^T = b^T A (A^T A)^{-1}$$

$$\Rightarrow \boxed{\underline{x} = \underbrace{(A^T A)^{-1} A^T}_{A^{\dagger}} \underline{b}}$$

for $m > n$

$A \in \mathbb{R}^{m\times n}$

$b \in \mathbb{R}^{m\times 1}$

$x \in \mathbb{R}^{n\times 1}$

$m > n$

$5 = 2m + c$

$10 = 4m + 2c$

when $m < n$

m ↑ # of eqns

n ↑ # of unknowns

$$A_{m\times n}\,x_{n\times 1} = b_{m\times 1}$$

$$\left[\min_{\underline{x}} \ \|\underline{x}\|^2 \quad \text{s.t. } A\underline{x} = \underline{b}\right]$$

$$\underline{x} = \underbrace{A^T (AA^T)^{-1}}_{A^+} \underline{b}$$

$\|x\|_1 = L_1\text{-norm of } \underline{x} = |x_1| + |x_2| + |x_3| + \dots |x_n|$

$\|x\|_2 = L_2 \text{ norm of } \underline{x} = \sqrt{x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2}$

$\|x\|_p = L_p \text{ norm of } \underline{x} = \left(x_1^p + x_2^p + \dots + x_n^p\right)^{1/p}$

for $A \in \mathbb{R}^{m \times n}$

$$A^+ = \begin{cases} A^{-1} & \text{exact} & \text{if } m = n \\ (A^TA)^{-1}A^T & & \text{if } m > n \\ A^T(AA^T)^{-1} & & \text{if } m < n \end{cases}$$

least square sol
min $\|Ax - b\|$

least norm solution
min $\|x\|^2$
s.t $A\underline{x} = \underline{b}$

SVD : Singular Value Decomposition

$$A = U \Sigma V^T$$

orthonormal

Diagonal matrix

$$\begin{bmatrix} \lambda_1 & & & & & \\ & \lambda_2 & & & 0 & \\ & & \ddots & & & \\ & & & \lambda_r & & \\ & 0 & & & 0 & \\ & & & & & \ddots \ 0 \end{bmatrix}$$

$$\Sigma^+ = \begin{bmatrix} 1/\lambda_1 & & & & & \\ & 1/\lambda_2 & & & 0 & \\ & & \ddots & & & \\ & & & 1/\lambda_r & & \\ & 0 & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

$$A^+ = V \Sigma^+ U^T$$

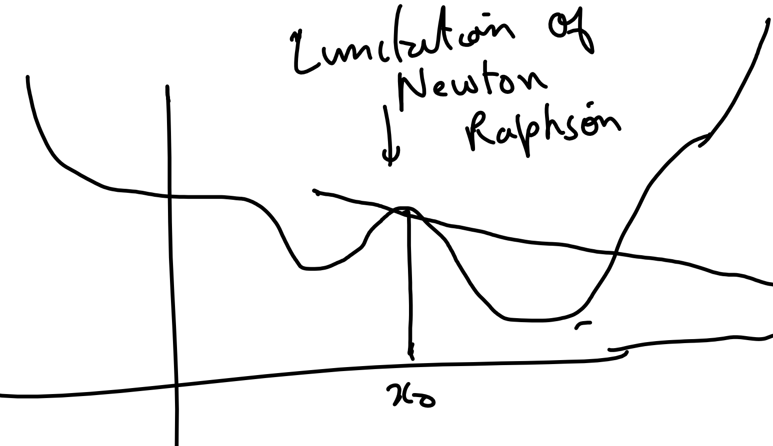
(3) What is the relationship (similarities and differences) b/w Newton-Raphson, Gauss Newton, Gradient descents

Newton - Raphson

Finds roots of function $f(x)$

$$f'(x_0) = \frac{f(x_0) - 0}{x_0 - x_1}$$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$



It may not converge

Can easily create overflow conditions

Comparing Newton Raphson (NR) with Gauss Newton (GN), Gradient Descent (GD)

↳ NR is root finding algo.

↳ GN/GD is minimizing algo.

Find roots of $f(x)$ $\equiv$ Minimize $(f(x))^2$

Minimize $f(x)$ $\equiv$ Find roots of $f'(x)$

Gauss Newton algo.

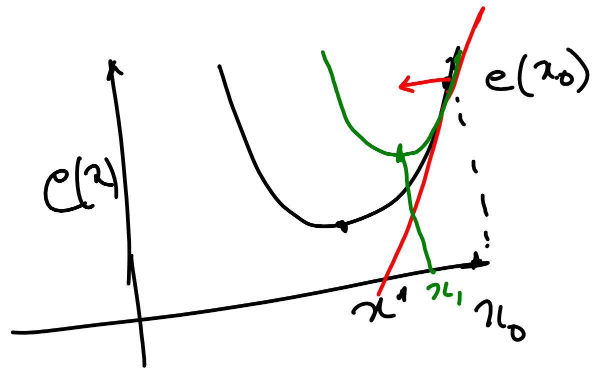
Used to minimize L2 norm of a residual vector

$$\underline{e}(\underline{x}) : \mathbb{R}^m \longrightarrow \mathbb{R}^n$$

$$\min_{\underline{x}} \|\underline{e}(\underline{x})\|^2$$

Quadratic approx.

$$e(x) = \boxed{e(x_0) + \frac{(x-x_0)}{1!} e'(x_0) + \frac{(x-x_0)^2}{2!} e''(x_0)} + \ldots \infty$$



$$\underbrace{x_1 = x_0 - \frac{de(x_0)}{dx}\alpha}_{\text{Gradient descent}}$$

$$x_1 = \min_x \underbrace{\left[e(x_0) + (x-x_0)\, e'(x_0) + \frac{(x-x_0)^2}{2} e''(x_0) \right]}_{\frac{\partial}{\partial x}}$$

$$e(x) = [f(x)]^2$$

$$e'(x) = 2f(x)$$

$$e''(x) = 2f'(x)$$

$$e'(x_0) + \frac{2(x-x_0)}{2} e''(x_0) = 0$$

$$x_1 = x_0 - \frac{e'(x_0)}{e''(x_0)} \quad \Bigg| \quad x_1 = x_0 - \frac{f(x)}{f'(x)} \quad \Bigg| \quad x_1 = x_0 - \alpha e'(x)$$

Bastardized GN in scalar form | NR | GD

Jacobian matrix is first derivative in vector form $\underline{f}(\underline{x})$

Hessian matrix is second derivative $f(\underline{x})$

scalar valued error function $e(\underline{x})$

Its Taylor series for the first two terms

$$e(\underline{x}) = e(\underline{x}_0) + \frac{(\underline{x} - \underline{x}_0)^T}{1!} J_{\underline{x}}[e(\underline{x}_0)]$$

$$+ \boxed{\frac{1}{2!}(\underline{x} - \underline{x}_0)^T H_{\underline{x}}[e(\underline{x}_0)]_{n \times n}(\underline{x} - \underline{x}_0)}$$

$$+ \cdots \infty$$

$$\underline{x}^T A \underline{x}$$
$$= x^2 a_{11} + y^2 a_{22} + xy\, a_{12} + yx\, a_{21}$$

$$\Rightarrow \quad J_x^T[e(\underline{x}_0)] + \frac{2}{2}(\underline{x} - \underline{x}_0)^T H_x[e(\underline{x}_0)] = 0$$

$$\Rightarrow \quad \underline{x}_1 = \underline{x}_0 - \underbrace{\left[H_x[e(x_0)]\right]^{-1}}_{\text{second-order derivative}} J_x[e(x_0)]$$

$$\begin{cases} H_x[e(x_0)] \approx J_x[e(x_0)]^T J_x[e(x_0)] \\ \Rightarrow \underline{x}_1 = x_0 - \left(J_x[e(x_0)]^T J_x[e(\underline{x}_0)]\right)^{-1} J_x[e(\underline{x}_0)] \end{cases}$$

GN

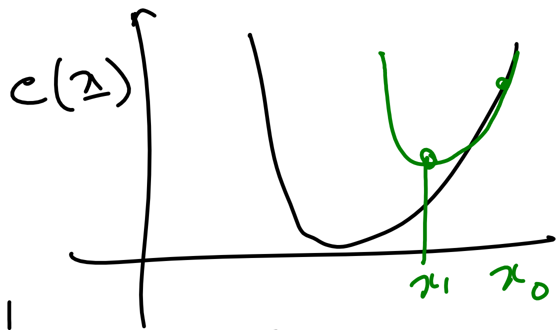
Gauss Newton

$$e(\underline{x}) = \|\underline{r}(\underline{x})\|^2 = \underline{r}(x)^T \underline{r}(\underline{x})$$

residual

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

$$e(\underline{x}) : \mathbb{R}^m \to \mathbb{R}$$



$$\underline{x}_1 = \underline{x}_0 - \underbrace{\left[H_x[e(\underline{x}_0)]\right]^{-1}}_{\text{second-order derivation}} J_x[e(\underline{x}_0)]$$

$$J_x[e(\underline{x}_0)] = J_{\underline{x}}\left[\underline{r}(\underline{x}_0)^T r(\underline{x}_0)\right]$$

$$= \underline{2\underline{r}(\underline{x}_0)^T J_x[r(\underline{x}_0)]}$$

$$J_x[\underline{x}^T A \underline{x}] = x^T(A+A^T)$$

$$H_x[e(\underline{x}_0)] = \begin{bmatrix} \frac{\partial^2 e}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 e}{\partial x_n \partial x_1} \\ \vdots & \ddots & \\ \frac{\partial^2 e}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 e}{\partial x_n \partial x_n} \end{bmatrix} = J_x\left[\left(J_{\underline{x}}[e(\underline{x}_0)]\right)^T\right]$$

$$H_x[e(\underline{x}_0)] = J_{\underline{x}}\left\{\left(2r(x_0)^T J_x[\underline{r}(\underline{x}_0)]\right)^T\right\}$$

$$= J_{\underline{x}}\left\{2 J_{\underline{x}}[\underline{r}(\underline{x}_0)]^T \underline{r}(x_0)\right\}$$

$$(AB)^T = B^T A^T$$

$$= 2\, J_x\left[\underline{r}(\underline{x}_0)\right]^T J_x\left[\underline{r}(\underline{x}_0)\right] + \underbrace{2\, J_{\underline{x}}\left\{J_{\underline{x}}\left[\underline{r}(\underline{x}_0)\right]\right\}^T}_{\text{second derivation}}$$

Gauss Newton

$$H_x\left[c(\underline{x}_0)\right] \approx 2\, J_x\left[\underline{r}(\underline{x}_0)\right]^T J_x\left[r(x_0)\right]$$

$$\underline{x}_1 = \underline{x}_0 - H_{\underline{x}}\left[c(\underline{x}_0)\right]^{-1} J_{\underline{x}}\left[c(\underline{x}_0)\right]$$

$$= \underline{x}_0 - \left(2 J_x\left[\underline{r}(x_0)\right]^T J_x\left[\underline{r}(x_0)\right]\right)^{-1}\left(2 J_x\left[\underline{r}(x_0)\right]^T \underline{r}(x_0)\right)$$

$$= \underline{x}_0 - \underbrace{\left(J^T J\right)^{-1} J^T}_{J^{+} \leftarrow \text{pseudo inverse}} \underline{r}$$