# 02- Least Squares

→ Intro (linear case)

$$\underset{m\times n}{A} P = x$$

$m < n \rightarrow \infty$, no sol$^n$

$m = n \rightarrow 1$, no

$m > n \rightarrow 1$, no

→ no sol$^n$ when $x \notin$ column space of $A$

→ can happen when we have some noise

⇒ we would like to estimate $P$

$$AP \sim x$$

if closest estimate → $P^o$

$$\boxed{A^T(AP^o - x) = 0}$$

→ why?

$\|AP^o - x\|$ is least when

$AP^o - x$ is $\perp$ to $A$ (vector space)

<u>NOTE</u> → $A^TA$ is non-singular if cols of $A$ are linearly indep.

Non linear case → see this for derivation but the SLAM prob. paper notes for motivation & understanding

→ Taylor series

* $f(x) = f(a) + \frac{f'(a)}{1!}(x-a)^1 \ldots$

$$\boxed{f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n}$$

*

$$\boxed{f(x) = f(a) + \frac{\nabla f(a)^T (x-a)}{1!} + \frac{(x-a)^T H_f (x-a)}{2!} \ldots}$$

* point → minimizing $F(x)$

$$F(x) = \frac{1}{2} \| f(x) \|_2^2$$

↳ multivariate multivalued ($x = \partial$)

Easy way → calculating $\frac{\partial F}{\partial x} = 0$

(But, computationally ( expen... )

...pensive)

∴ very complex.

Our way → ① take an initial estimate $p_0$.

② Calculate $\Delta p_k$ such that
$k$ → $k^{th}$ iteration

$$F(p_k + \Delta p_k) < F(p_k)$$

③ stop if $\Delta p_k < \varepsilon$
else, $p_{k+1} \leftarrow p_k + \Delta p_k$

* $x = \vec{x}$

$$F(x + \Delta x) \simeq F(x) + \nabla J_F^T \Delta x$$

And, we want

$$F(x + \Delta x) < F(x)$$

So, we take

$$\boxed{\Delta x = -\lambda J_f}$$

→ gradient descent

→ problem → ① too slow
② may overshoot minima.

# For better accuracy, expand taylor (combatting overshooting)

F( ... )

$$F(x+\Delta x) = F(x) + J^T \Delta x + \frac{1}{2} \Delta x^T H \Delta x$$

$$\Delta F = J^T \Delta x + \frac{1}{2} \Delta x^T H \Delta x$$

We want $\Delta x^* = \underset{\Delta x}{\operatorname{argmin}} \Delta F$

$$\frac{\partial \Delta F}{\partial \Delta x} = J_F + H_F \Delta x = 0$$

$$H_F \Delta x = -J_F$$

Now, $f(x+\Delta x) = f(x) + J_f \Delta x$

We approximate $f(x)$ as well for faster computation ($H_F$, $J_F \to$ complex)

$$\Delta x^* = \underset{\Delta x}{\operatorname{argmin}} \underbrace{\frac{1}{2} \| f(x) + J^T \Delta x \|_2^2}$$

$$\to \frac{1}{2} (f(x) + J^T \Delta x)^T (f(x) + J^T x)$$

$$= \frac{1}{2} (f^T(x) + \Delta x^T J)(f(x) + J^T \Delta x)$$

$$= \frac{1}{2} ( \| f(x) \|^2 + f^T(x) J^T \Delta x$$

$$+ \Delta x^T J f(x)$$
$$+ \Delta x^T J J^T \Delta x)$$

all are scalars
(so, rearrange)

$$= \frac{1}{2} ( \|f(x)\|^2 + 2 f(x) J^T \Delta x$$
$$+ \Delta x^T J J^T \Delta x)$$

diff wrt $\Delta x$:

$$J f(x) + J J^T \Delta x = 0$$

$$J_f J_f^T \Delta x = - J_f f(x)$$

→ Gauss Newton

→ Problems → ① $J J^T \geq 0$
(not necessarily invertible) → ∵ some eigenvalues may be 0 & det = Π eigen values

② approximate

# LM

$$(J_f J_f^T + \lambda I) \Delta x = - J_f f(x)$$

(continued next part 07)