LEAST SQUARES AND UNCERTAINTY

Why filtering → least squares ?

↓ until now focused on optimization!

$$x^* = \underset{x}{\operatorname{argmax}}\; p(x|z)$$

the most likely

x conditioned by z (measurement)

Bayes Rule

consider uniform dist

$$p(x|z) = \frac{p(z|x)\,p(x)}{p(z)}$$

only a normalizer

$$\propto p(z|x)$$

$$= \prod_i p(z^{[i]}|x)$$

independence assumption

Gaussian assumption

$$p(z_i|x) = \mathcal{N}(z^{[i]};\, h^{[i]}(x),\, \Sigma^{[i]})$$

$$\propto \exp(-e^{[i]} \dots)$$

max ⇒ minimization

exp(...) $\oplus$ $e^T \Omega e$ → $\|e^{[i]}(x)\|^2_\Omega$

$\prod \Rightarrow \sum$

Ω L2-norm

GOAL: obtain some statistics of X ⤻ to estimate uncertainty of our solution

X^* - solution

each measurement z^* = prediction ≡ evaluate the prediction function h at optimum x^*

if I want to take a state and project through prediction function:

CONDITIONAL

$$p(z^{[i]}|\Delta x + x^*) \sim \mathcal{N}(J^{[i]}\Delta x + z^{[i]*},\; \Omega^{[i]-1})$$

↑ error propagation ?...

↓ describes the conditional distribution of the measurement given perturbation around the optimal state

⇓

JOINT DISTRIBUTION { 1. joint distribution over the entire measurement (stack of all possible measurements)

ICP: stack of all predicted 2D points

$$p(z \mid \Delta x + x^*) \sim \mathcal{N}(\mu_z, \Omega_z^{-1})$$

$$\Updownarrow$$

$$\mu_z = \begin{pmatrix} J^{[1]}\Delta x + z^{[1]*} \\ \vdots \\ J^{[K]}\Delta x + z^{[K]*} \end{pmatrix} = \underbrace{\begin{pmatrix} J^{[1]}\Delta x \\ \vdots \\ J^{[K]}\Delta x \end{pmatrix}}_{\tilde{J}} \Delta x + \underbrace{\begin{pmatrix} z^{[1]*} \\ \vdots \\ z^{[K]*} \end{pmatrix}}_{\tilde{z}^*}$$

$$\Omega_z = \begin{pmatrix} \Omega^{[1]} & & 0 \\ & \Omega^{[2]} & \\ 0 & & \ddots & \Omega^{[K]} \end{pmatrix}$$

assuming measurements independent

2. Joint x, z (are in information form) → but are equivalent to the other

$$p(x) = \mathcal{N}(x; x^*, \Sigma_x)$$

$$p(z|x)$$

known

$$\Downarrow$$

$$p(x, z) \longrightarrow \mu_{x,z} = \begin{pmatrix} x^* \\ z^* \end{pmatrix}$$

$$\Omega_{x,z} = \begin{pmatrix} J^T \Omega_z J + \Omega_x & -J^T \Omega_z \\ -\Omega_z J & \Omega_z \end{pmatrix}$$

$$\Downarrow$$

CHAIN RULE

do not know nothing about my state

$$\Sigma_x = \infty \Rightarrow \Omega_x = 0 \text{ (no information)}$$

$$p(\Delta x) = \mathcal{N}(0, \Sigma_x)$$

given that a perturbation around optimal!

(NOTE
WE
ARE
AT THE
OPTIMUM)

$$p(\Delta x, z) \sim \mathcal{N}\left([x, z^*]^T, \ \Sigma_{x,z}\right)$$

$$\Omega_{x,z} = \begin{pmatrix} J^T \Omega_z J & -J^T \Omega_z \\ -\Omega_z J & \Omega_z \end{pmatrix}$$

should sound familiar given that is a part of last equation

No Ω_x given that we set $\Omega_x = \emptyset$

$$\Omega_{xx} = J^T \Omega_z J$$

$$= H$$ → same matrix from last equation.....

$\Omega_{x|z} = \Omega_{xx}$ → so, we do not need the other cross-correlation

$$\Downarrow$$

$$\mu_{x|z} = x^* + \Sigma_{xz} \Sigma_{zz}^{-1} (z - z^*)$$

$z \cong z^*$ → we are at the optimum

$$\Downarrow$$

$$\mu_{x|z} \cong x^*$$

$$\Sigma_{x|z} = \Omega_{xx}^{-1}$$

$$\cong H^{-1}$$

so, we can tell how certain the solution is by least-square

Notes:

- if I have measurement dimension 1 & state = 10, each time integrate a measurement into machinery, will be

$$J^{T}_{10\times1}\ R_{1\times1}\ J^{m}_{1\times10} = [\ \ ]_{10\times10} \longrightarrow \text{rank} = 1 \,//$$

↓ in the best of the cases, will qualify only 1 dimension of my [illegible]

- at minimum, would need 10 measurement to be able to resolve the state

[min(# measurement) ⟶ dim(measurement) × # measurement ≥ dim(state)]

but is not a sufficient condition

is it why then I need 3 points or something else?

2 points ⊕ information

1 coordinate of the third point for example