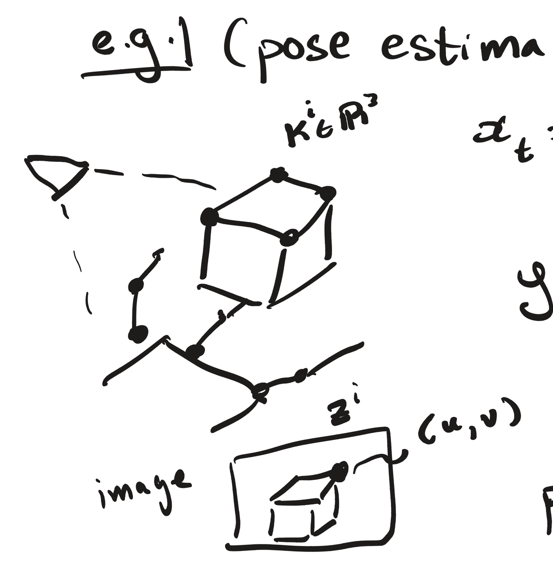
3/26: Estimation via Least-Squares

Announcements:
- HW3 due today (3/26)
- HW4 $\longrightarrow$ after SB
- milestone feedback ~ tomorrow

Last time:
- state estimation problem

Today:
- solution via least-squares

e.g.) (pose estimation)

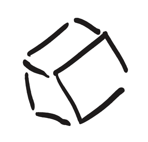


$x_t = [\, p_t,\ R_t,\ v_t,\ \omega_t \,]$
$\quad \in\mathbb{R}^3 \quad \in SO(3) \quad \in\mathbb{R}^3 \quad \in\mathbb{R}^3$

$y_t = \begin{bmatrix} z^1 \\ \vdots \\ z^8 \end{bmatrix} \in \mathbb{R}^{16}$

$p(x_{t+1} \mid x_t, u_t)$ dynamics
"predict using GT state"

$p(y_{t+1} \mid x_{t+1})$ measurement



"put cube into GT state"

$$\boxed{\hat{x}_{0:T} = \arg\max_{x_{0:T}} p(x_{0:T} \mid u_{0:T-1}, y_{1:T})}$$

MAP estimation
↪ unconstrained NLP

Recall:

i) Conditional independence

$X \perp Y \mid Z$ "X conditionally ind. of Y, given Z"

$p(x,y \mid z) = p(x\mid z)\, p(y \mid z)$

$$p(x \mid y, z) = \frac{p(x, y \mid z)}{p(y\mid z)} \quad \text{[cond. prob]}$$

$$= \frac{p(x\mid z)\cancel{p(y\mid z)}}{\cancel{p(y \mid z)}} = p(x \mid z)$$

"Y gives no information about X, given Z"

ii) Bayes' rule

$$p(x\mid y) = \frac{p(y\mid x)\, p(x)}{p(y)} \propto p(y \mid x)\, p(x)$$

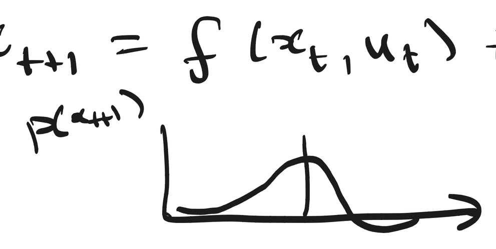
$p(y)$ not a function of x

$$\arg\max_x p(x\mid y) = \arg\max_x p(y\mid x)\, p(x)$$

$$\arg\max_{x_{0:T}} p(x_{0:T} \mid u_{0:T-1}, y_{1:T})$$
$$= \arg\min_{x_{0:T}} -\log p(x_{0:T} \mid u_{0:T-1}, y_{1:T})$$
$$= \arg\min_{x_{0:T}} -\log p(x_0) + \sum_{t=0}^{T-1} -\log p(x_{t+1} \mid x_t, u_t) - \log p(y_{t+1} \mid x_{t+1})$$

Important special case:

$x_{t+1} = f(x_t, u_t) + w_t \qquad w_t \sim \mathcal{N}(0, Q)$



$y_{t+1} = g(x_{t+1}) + v_{t+1} \qquad v_{t+1} \sim \mathcal{N}(0, R)$

under this model:

$p(x_{t+1} \mid x_t, u_t) = \mathcal{N}(f(x_t, u_t), Q)$

$p(y_{t+1} \mid x_{t+1}) = \mathcal{N}(g(x_{t+1}), R)$

$$\mathcal{N}(x; \mu, \Sigma) = \underbrace{\frac{1}{\sqrt{(2\pi)^n |\Sigma|}}}_{\eta} \cdot \exp\left(\frac{-1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

$$\log p(x; \mu, \Sigma) = \log \eta - \frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)$$

$(\Sigma^{-1/2} \cdot \Sigma^{-1/2}) = \Sigma^{-1}$ cholesky decomp/ matrix square root

$$\log p(x; \mu, \Sigma) = \log \eta - \frac{1}{2}\|\Sigma^{-1/2}(x-\mu)\|^2$$

If we choose a prior $p(x_0) = \mathcal{N}(\mu_0, \Sigma_0)$

$$-\log p(x_{0:T} \mid u_{0:T-1}, y_{1:T}) = \|\Sigma_0^{-1/2}(x_0 - \mu_0)\|^2 + \sum_{t=0}^{T-1} \|Q^{-1/2}(x_{t+1} - f(x_t, u_t))\|^2 + \|R^{-1/2}(y_{t+1} - g(x_{t+1}))\|^2$$

$$\arg\max p(x_{0:T} \mid h_T) = \sum_{i=1}^{N} \|r_i(x)\|^2 \quad \text{(nonlinear LS!)}$$

Recall: Gauss-Newton

$x^{(0)} \qquad J(x^{(k)})^T J(x^{(k)})\, \delta x^{(k)} = -J(x^{(k)})\, r(x^{(k)})$

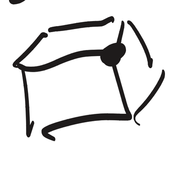
$x^{(k+1)} \leftarrow x^{(k)} + \delta x^{(k)}$

Fun fact: one iteration of GN = EKS

Important point: does this converge?
is it unique?

↪ observability "is the estimation problem well-posed?"

e.g. (pose estimation)



single keypoint

If J doesn't have full column rank $\exists$ multiple δx that solve the normal equations.