SLAM 101

Lecture 00 MLE & MAP

Ayoung Kim

Estimator

Estimator – MLE and MAP

likelihood

$$p(\theta|\mathrm{x}) = \frac{p(\mathrm{x}|\theta)\ p(\theta)}{p(\mathrm{x})}$$

Posterior probability

- Maximum likelihood estimator (MLE)
  - Estimate θ that maximize the likelihood $\theta \mapsto p(\mathrm{x}|\theta)$

$$\hat{\theta}_{ML}(\mathrm{x}) = \mathrm{argmax}\ p(\mathrm{x}|\theta)$$

- Maximum a posteriori (MAP) estimation
  - Estimate θ that maximize the posterior prob $\theta \mapsto p(\theta|\mathrm{x})$

$$\begin{aligned}\hat{\theta}_{MAP}(\mathrm{x}) &= \mathrm{argmax}\ p(\theta|\mathrm{x}) = \mathrm{argmax}\ \frac{p(\mathrm{x}|\theta)\ p(\theta)}{p(\mathrm{x})} = \mathrm{argmax}\ \frac{p(\mathrm{x}|\theta)\ p(\theta)}{\int p(\mathrm{x}|\theta)\ p(\theta) d\theta} \\ &= \mathrm{argmax}\ p(\mathrm{x}|\theta)\ p(\theta)\end{aligned}$$

[Example] MLE for Gaussian

- Parameters in Gaussian = mean and covariance
- We have a set of observation $\mathrm{x} = \{x_1, x_2, \cdots, x_N\}$
  - We think the distribution is Gaussian
  - Want to find mean using MLE $\hat{\theta}_{ML}(\mathrm{x}) = \mathrm{argmax}\ p(\mathrm{x}|\theta)$

- Given
  - A set of i.i.d. observation $\mathrm{x} = \{x_1, x_2, \cdots, x_N\}$

    independent and identically distributed

- Find the mean θ

$$p(x|\mu, \sigma) = \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

$$p(x|\theta) = \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp\left\{-\frac{1}{2\sigma^2}(x-\theta)^2\right\}$$

[Example] MLE in Gaussian

- Likehihood

$$\mathrm{x} = \{x_1, x_2, \cdots, x_N\}$$

$$\begin{aligned}\mathcal{L}(\theta) &= p(\mathrm{x}|\theta) = \prod_{i=1}^{N} p(x_i|\theta) \\ &= \left\{\frac{1}{\sqrt{(2\pi\sigma^2)}}\right\}^N \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\theta)^2\right\}\end{aligned}$$

- The negative log likelihood (log preserves the max)

$$\ln\mathcal{L} = \ln p(\mathrm{x}|\theta) = -\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\theta)^2 - \frac{N}{2}\ln\sigma^2 - \frac{N}{2}\ln(2\pi)$$

- Differentiate

$$\frac{\partial}{\partial\theta}(-ln\mathcal{L}) = \frac{\partial\ell}{\partial\theta} = 0 \quad \rightarrow \quad \frac{\partial}{\partial\theta}\left\{\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\theta)^2\right\} = \frac{1}{2\sigma^2}\sum_{i=1}^{N} -2(x_i-\theta) = \frac{-1}{\sigma^2}\sum_{i=1}^{N}(x_i-\theta)$$

$$\rightarrow \quad \theta_{ML} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

[Example] MAP for Gaussian

- Parameters in Gaussian = mean and covariance
- We have a set of observation $\mathrm{x} = \{x_1, x_2, \cdots, x_N\}$
  - We think the distribution is Gaussian
  - Want to find mean using MAP

$$\hat{\theta}_{MAP}(\mathrm{x}) = \mathrm{argmax}\ p(\mathrm{x}|\theta)\ p(\theta)$$

- Given
  - A set of i.i.d. observation $\mathrm{x} = \{x_1, x_2, \cdots, x_N\}$
  - Parameter prior $\theta = \mathcal{N}(\mu, 1)$

- Find the mean θ

[Example] MAP for Gaussian

$$\hat{\theta}_{MAP}(\mathrm{x}) = \operatorname{argmax}\, p(\mathrm{x}|\theta) = \operatorname{argmax}\, p(\mathrm{x}|\theta)\; p(\theta)$$
$$= \operatorname{argmin}\; (-\ln p(\mathrm{x}|\theta) - \ln p(\theta))$$

Take log

$$\frac{\partial}{\partial\theta} = 0 \quad \rightarrow \quad \frac{\partial}{\partial\theta}\{-\ln p(\mathrm{x}|\theta) - \ln p(\theta)\} = 0$$

$$\frac{\partial}{\partial\theta} - \ln p(\mathrm{x}|\theta) = \frac{-1}{\sigma^2}\sum_{i=1}^{N}(x_i - \theta)$$

… MLE

$$\frac{\partial}{\partial\theta} - \ln p(\theta) = \frac{\partial}{\partial\theta}\left\{\frac{1}{2}\ln 2\pi + \frac{1}{2}(\theta-\mu)^2\right\} \qquad \theta \sim \mathcal{N}(\mu, 1) = \frac{1}{\sqrt{2\pi}}\exp\left\{\tfrac{1}{2}(\theta-\mu)^2\right\}$$
$$= \frac{1}{2}\cdot 2(\theta-\mu) = -(\mu-\theta)$$

$$-\frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i-\theta) - (\mu-\theta) = 0 \quad \rightarrow \quad (\sum_{i=1}^{N}\frac{x_i}{\sigma^2} + \mu) - (\frac{N}{\sigma^2} + 1)\theta$$
$$\rightarrow \quad \theta_{MAP} = \frac{\sum_{i=1}^{N}\frac{x_i}{\sigma^2} + \mu}{\frac{N}{\sigma^2} + 1}$$

ANY QUESTIONS?