Robotics

Estimation and Learning
with Dan Lee

Supplementary Notes on MLE for Univariate Gaussian

Penn Engineering

Maximum Likelihood Estimate of Gaussian Model Parameters

- Objective

  Estimate the mean and the variance given observed data

$$\hat{\mu}, \hat{\sigma} = \arg\max_{\mu,\sigma} p(\{x_i\}|\mu, \sigma)$$

- Assuming independence of observations,

$$\hat{\mu}, \hat{\sigma} = \arg\max_{\mu,\sigma} \prod_{i=1}^{N} p(x_i|\mu, \sigma)$$

Maximum Likelihood Estimate of Gaussian Model Parameters

- To obtain $\hat{\mu}, \hat{\sigma} = \arg\max_{\mu,\sigma} \prod_{i=1}^{N} p(x_i|\mu, \sigma)$

(1) $$\arg\max_{\mu,\sigma} \prod_{i=1}^{N} p(x_i|\mu, \sigma) = \arg\max_{\mu,\sigma} \sum_{i=1}^{N} \ln p(x_i|\mu, \sigma)$$

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

Maximum Likelihood Estimate of Gaussian Model Parameters

$$\hat{\mu}, \hat{\sigma} = \arg\max_{\mu,\sigma} \prod_{i=1}^{N} p(x_i | \mu, \sigma)$$

$$= \arg\max_{\mu,\sigma} \sum_{i=1}^{N} \ln p(x_i | \mu, \sigma)$$ from (1)

$$= \arg\max_{\mu,\sigma} \sum_{i=1}^{N} \ln \left\{ \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{(x_i - \mu)^2}{2\sigma^2} \right\} \right\}$$ from (2)

Maximum Likelihood Estimate of Gaussian Model Parameters

(continued)

$$\hat{\mu}, \hat{\sigma} = \arg\max_{\mu,\sigma} \sum_{i=1}^{N} \ln\left\{\frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x_i - \mu)^2}{2\sigma^2}\right\}\right\}$$

$$= \arg\max_{\mu,\sigma} \sum_{i=1}^{N} \left\{-\frac{(x_i - \mu)^2}{2\sigma^2} - \ln\sqrt{2\pi}\sigma\right\}$$

$$= \arg\min_{\mu,\sigma} \sum_{i=1}^{N} \left\{\frac{(x_i - \mu)^2}{2\sigma^2} + \ln\sigma + const\right\}$$

Maximum Likelihood Estimate of Gaussian Model Parameters

Let $$J(\mu, \sigma) = \sum_{i=1}^{N} \left\{ \frac{(x_i - \mu)^2}{2\sigma^2} + \ln\sigma \right\}$$

Then $$\hat{\mu}, \hat{\sigma} = \arg\min_{\mu,\sigma} J(\mu, \sigma)$$

① $\dfrac{\partial J}{\partial \mu} = 0 \longrightarrow \hat{\mu}$ ② $\dfrac{\partial J(\hat{\mu}, \sigma)}{\partial \sigma} = 0 \longrightarrow \hat{\sigma}$

Maximum Likelihood Estimate of Gaussian Model Parameters

① $\boxed{\dfrac{\partial J}{\partial \mu} = 0 \longrightarrow \hat{\mu}}$

$$\frac{\partial J}{\partial \mu} = \frac{\partial}{\partial \mu}\sum_{i=1}^{N}\left\{\frac{(x_i-\mu)^2}{2\sigma^2} + \ln\sigma\right\}$$

$$= \sum_{i=1}^{N}\left\{\frac{\partial}{\partial \mu}\frac{(x_i-\mu)^2}{2\sigma^2}\right\}$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i-\mu) = 0$$

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i \longleftarrow$$

Maximum Likelihood Estimate of Gaussian Model Parameters

② $\boxed{\dfrac{\partial J(\hat{\mu},\sigma)}{\partial \sigma} = 0 \longrightarrow \hat{\sigma}}$

$$\frac{\partial J}{\partial \sigma} = \frac{\partial}{\partial \sigma}\sum_{i=1}^{N}\left\{\frac{(x_i-\hat{\mu})^2}{2\sigma^2} + \ln\sigma\right\}$$

$$= \left(\frac{\partial}{\partial \sigma}\frac{1}{2\sigma^2}\right)\left(\sum_{i=1}^{N}(x_i-\hat{\mu})^2\right) - \frac{N}{\sigma}$$

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\hat{\mu})^2 \longleftarrow = \frac{1}{\sigma}\left(N - \frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i-\hat{\mu})^2\right) = 0$$

Maximum Likelihood Estimate of Gaussian Model Parameters

- In summary, we have

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N} (x_i - \hat{\mu})^2$$