Robotics

Estimation and Learning with Dan Lee

Week 1. Gaussian Model Learning

1.3.2 MLE of Multivariate Gaussian

Penn Engineering

Maximum Likelihood Estimate of Multivariate Gaussian Parameters

- Objective

$$\hat{\boldsymbol{\mu}}, \hat{\Sigma} = \arg\max_{\boldsymbol{\mu},\boldsymbol{\Sigma}} p(\{\mathbf{x}_i\}|\boldsymbol{\mu}, \Sigma)$$

Maximum Likelihood Estimate of Multivariate Gaussian Parameters

- Objective

$$\hat{\boldsymbol{\mu}}, \hat{\Sigma} = \arg\max_{\boldsymbol{\mu}, \Sigma} p(\{\mathbf{x}_i\} | \boldsymbol{\mu}, \Sigma)$$

Assuming independence of observations,

$$\hat{\boldsymbol{\mu}}, \hat{\Sigma} = \arg\max_{\boldsymbol{\mu}, \Sigma} \prod_{i=1}^{N} p(\mathbf{x}_i | \boldsymbol{\mu}, \Sigma)$$

Maximum Likelihood Estimate of Multivariate Gaussian Parameters

$$\hat{\boldsymbol{\mu}}, \hat{\Sigma} = \arg\max_{\boldsymbol{\mu},\Sigma} \prod_{i=1}^{N} p(\mathbf{x}_i | \boldsymbol{\mu}, \Sigma)$$

(1) Take the log!

$$\arg\max\, likelihood \leftrightarrow \arg\max \ln(likelihood)$$

$$\log(x_1 \times x_2 \times \cdots \times x_k) = \log(x_1) + \log(x_1) + \cdots + \log(x_k)$$

$$\boxed{\arg\max_{\boldsymbol{\mu},\Sigma} \prod_{i=1}^{N} p(\mathbf{x}_i | \boldsymbol{\mu}, \Sigma) \quad \Rightarrow \quad \arg\max_{\boldsymbol{\mu},\Sigma} \sum_{i=1}^{N} \ln p(\mathbf{x}_i | \boldsymbol{\mu}, \Sigma)}$$

Maximum Likelihood Estimate of Multivariate Gaussian Parameters

$$\hat{\boldsymbol{\mu}}, \hat{\Sigma} = \arg\max_{\boldsymbol{\mu}, \Sigma} \sum_{i=1}^{N} \underline{\ln p(\mathbf{x}_i | \boldsymbol{\mu}, \Sigma)}$$

(2) Gaussian!

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^{T} \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right\}$$

$$\boxed{\ln p(\mathbf{x}_i | \boldsymbol{\mu}, \Sigma)} \longrightarrow \left\{-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu})^{T} \Sigma^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) - \frac{1}{2}\ln|\Sigma| + c\right\}$$

$c = -D/2 \ln(2\pi)$

Maximum Likelihood Estimate of Multivariate Gaussian Parameters

$$\hat{\boldsymbol{\mu}}, \hat{\Sigma} = \arg\max_{\boldsymbol{\mu},\Sigma} \sum_{i=1}^{N} \left\{ -\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu})^{\boldsymbol{T}} \Sigma^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) - \frac{1}{2} \ln|\Sigma| + c \right\}$$

$$\Updownarrow$$

$$\boxed{\hat{\boldsymbol{\mu}}, \hat{\Sigma} = \arg\min_{\boldsymbol{\mu},\Sigma} \sum_{i=1}^{N} \left\{ \frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu})^{\boldsymbol{T}} \Sigma^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) + \frac{1}{2} \ln|\Sigma| \right\}}$$

Maximum Likelihood Estimate of Multivariate Gaussian Parameters

$$\hat{\boldsymbol{\mu}}, \hat{\Sigma} = \arg\min_{\boldsymbol{\mu},\Sigma} \underbrace{\sum_{i=1}^{N} \left\{ \frac{1}{2} (\mathbf{x}_i - \boldsymbol{\mu})^{\boldsymbol{T}} \Sigma^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) + \frac{1}{2} \ln|\Sigma| \right\}}_{J(\boldsymbol{\mu},\Sigma)}$$

- At optimum,

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

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

Maximum Likelihood Estimate of Multivariate Gaussian Parameters

- In summary, we have

$$\hat{\boldsymbol{\mu}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i$$

$$\hat{\Sigma} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^\intercal$$

Multi-dimension Distribution: Example

- Ball color in multi-channels

$$\hat{\boldsymbol{\mu}} = \begin{bmatrix} 227 & 74 \end{bmatrix}$$

$$\hat{\Sigma} = \begin{bmatrix} 150 & -48 \\ -48 & 46 \end{bmatrix}$$