# Robotics

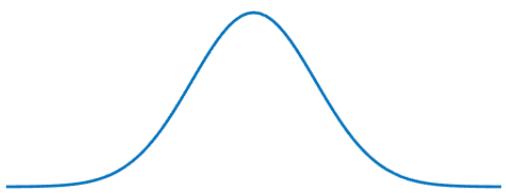
**Estimation and Learning**
**with Dan Lee**

# Week 1.
# Gaussian Model Learning

## 1.4.2 GMM Parameter Estimation via EM

Penn Engineering

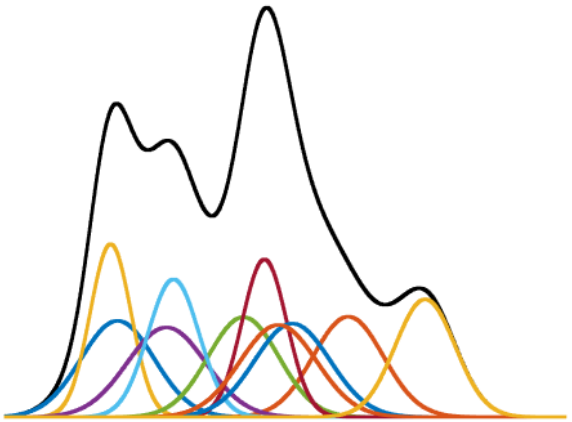
- Single Gaussian



$\boldsymbol{\mu}$

$\Sigma$

- Mixture of Gaussians



$$\boldsymbol{\mu} = \{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \dots, \boldsymbol{\mu}_K\}$$

$$\Sigma = \{\Sigma_1, \Sigma_2, \dots, \Sigma_K\}$$

$w = 1/K$

$K$: Given

# Learning GMM Parameters

Likelihood: $p(\{\mathbf{x}_i\}|\boldsymbol{\mu}, \Sigma)$

Observed data    Unknown parameters

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

$$\boldsymbol{\mu} = \{\boldsymbol{\mu}_k\}$$

$$\Sigma = \{\Sigma_k\} \quad k = 1,2, \dots, K$$

# Learning GMM Parameters

- Objective

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

Assuming independence of observations,

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

# Learning GMM 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!

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

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

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

# Learning GMM Parameters

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

(2) Gaussian Mixture Model!

$$p(\mathbf{x}) = \sum_{k=1}^{K} w_k g_k(\mathbf{x} | \boldsymbol{\mu}_k, \Sigma_k)$$

$g_k$: Gaussian with $\boldsymbol{\mu}_k$ and $\Sigma_k$

$$\boxed{\hat{\boldsymbol{\mu}}, \hat{\Sigma} = \arg\max_{\boldsymbol{\mu},\Sigma} \sum_{i=1}^{N} \ln \left\{ \frac{1}{K} \sum_{k=1}^{K} g_k(\mathbf{x}_i | \boldsymbol{\mu}_k, \Sigma_k) \right\}}$$

# Learning GMM Parameters

$$\hat{\boldsymbol{\mu}}, \hat{\Sigma} = \arg\max_{\boldsymbol{\mu},\Sigma} \sum_{i=1}^{N} \ln \left\{ \frac{1}{K} \sum_{k=1}^{K} g_k(\mathbf{x}_i | \boldsymbol{\mu}_k, \Sigma_k) \right\}$$

where

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

→ No closed form solution exist.

# Expectation-Maximization (EM)

1) Special case : EM for GMM Parameter Estimation
2) General EM Algorithm

# EM for GMM

Initial $\boldsymbol{\mu}$ and $\Sigma$

Latent variable $z$

# EM for GMM

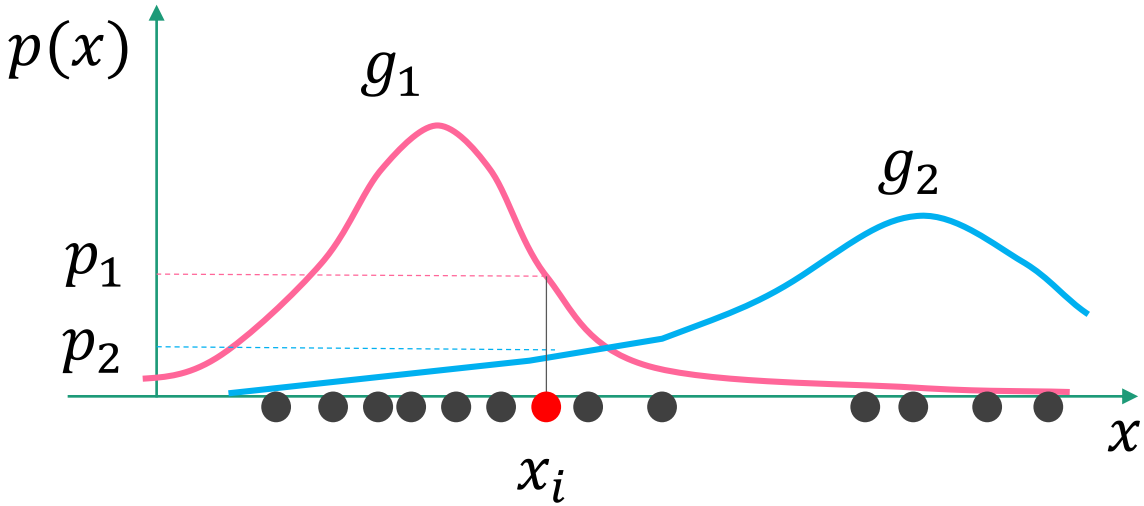
- Latent Variable

$$z_k^i = \frac{g_k(\mathbf{x}_i|\boldsymbol{\mu}_k, \Sigma_k)}{\sum_{k=1}^{K} g_k(\mathbf{x}_i|\boldsymbol{\mu}_k, \Sigma_k)}$$

# EM for GMM

- Latent Variable Example

$$z_k^i = \frac{g_k(\mathbf{x}_i|\boldsymbol{\mu}_k, \Sigma_k)}{g_1(\mathbf{x}_i|\boldsymbol{\mu}_1, \Sigma_1) + g_2(\mathbf{x}_i|\boldsymbol{\mu}_2, \Sigma_2)}$$




$$z_1^i = \frac{p_1}{p_1 + p_2}$$

$$z_2^i = \frac{p_2}{p_1 + p_2}$$

# EM for GMM

- Mean and Covariance Matrix

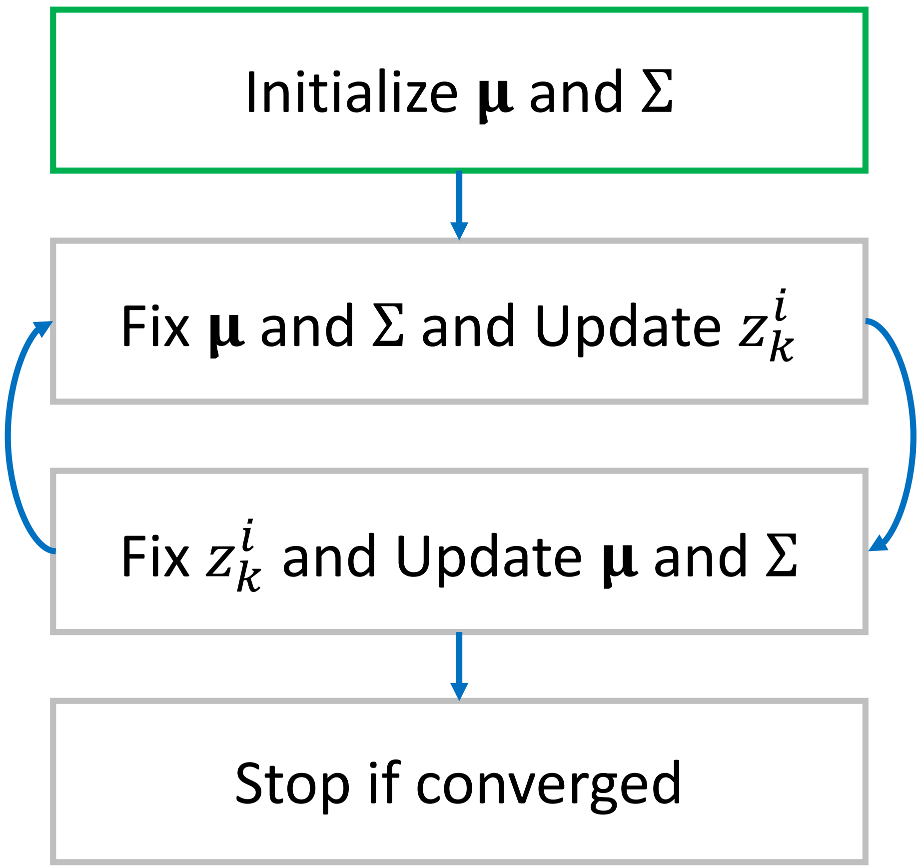
$$\widehat{\boldsymbol{\mu}}_k = \frac{1}{z_k}\sum_{i=1}^{N} z_k^i \mathbf{x}_i$$

$$\widehat{\Sigma}_k = \frac{1}{z_k}\sum_{i=1}^{N} z_k^i (\mathbf{x}_i - \widehat{\boldsymbol{\mu}}_k)(\mathbf{x}_i - \widehat{\boldsymbol{\mu}}_k)^\top$$

$$z_k = \sum\nolimits_{i=1}^{N} z_k^i$$

# EM for GMM

很像 MEC3456 里面一个方法：猜想和实际值之间的误差会被用来修正猜想。
具体是Jacobi, Gauss-Seidel,



# EM for GMM

Initialize $\boldsymbol{\mu}$ and $\Sigma$

E-step

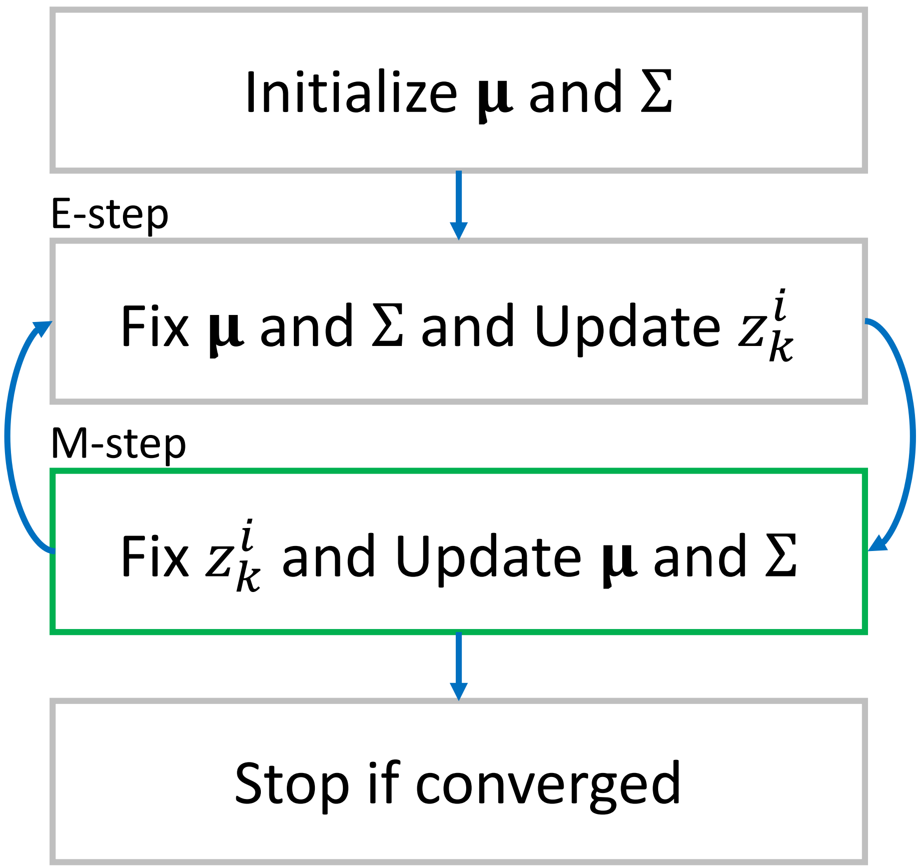
Fix $\boldsymbol{\mu}$ and $\Sigma$ and Update $z_k^i$

$$z_k^i = \frac{g_k(\mathbf{x}_i | \boldsymbol{\mu}_k, \Sigma_k)}{\sum_{k=1}^{K} g_k(\mathbf{x}_i | \boldsymbol{\mu}_k, \Sigma_k)}$$

Fix $z_k^i$ and Update $\boldsymbol{\mu}$ and $\Sigma$

Stop if converged

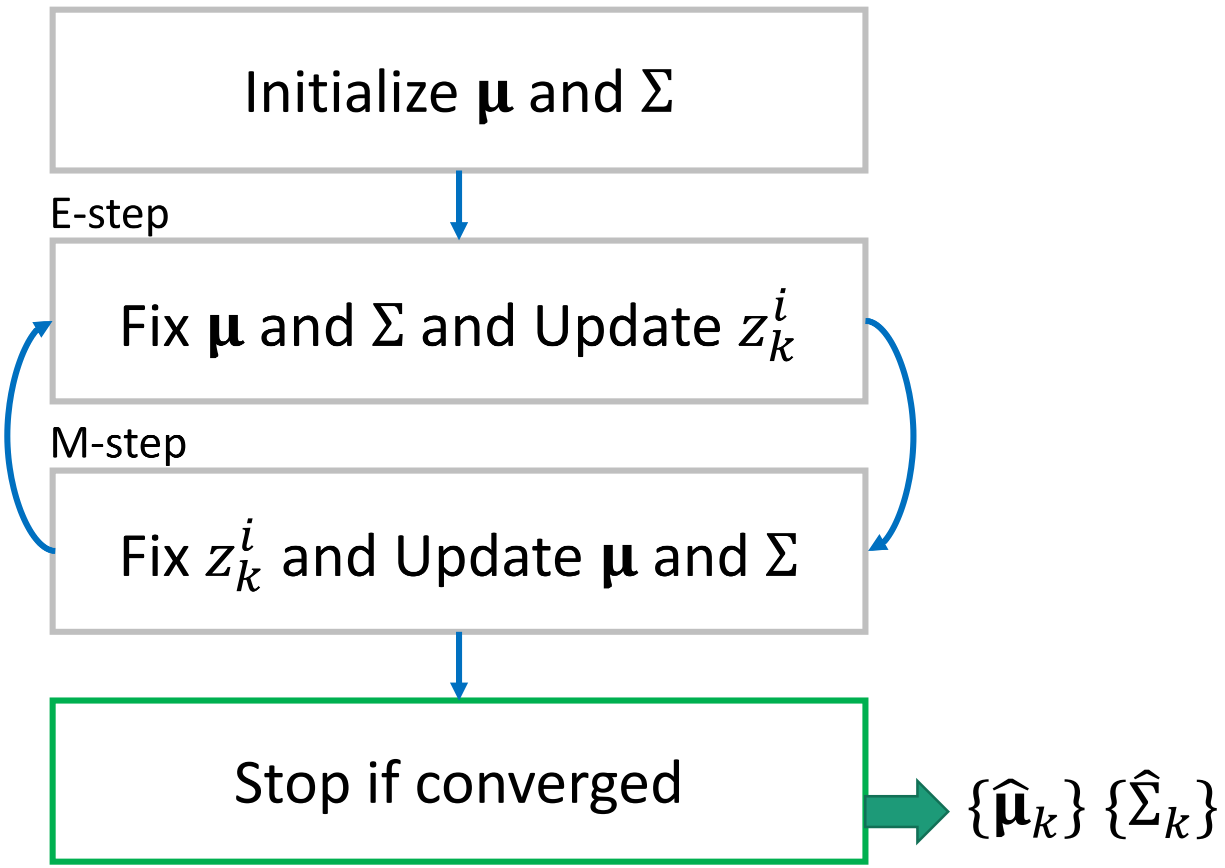
# EM for GMM




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

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

# EM for GMM


Initialize $\boldsymbol{\mu}$ and $\Sigma$
E-step
Fix $\boldsymbol{\mu}$ and $\Sigma$ and Update $z_k^i$
M-step
Fix $z_k^i$ and Update $\boldsymbol{\mu}$ and $\Sigma$
Stop if converged
$\{\hat{\boldsymbol{\mu}}_k\}$ $\{\hat{\Sigma}_k\}$

# EM for GMM

$\boldsymbol{\mu}_{\mathbf{0}}, \Sigma_0$

E-step: Update $z_k^i$

M-step: Update $\boldsymbol{\mu}$ and $\Sigma$

$\{\hat{\boldsymbol{\mu}}_k\}\ \{\hat{\Sigma}_k\}$