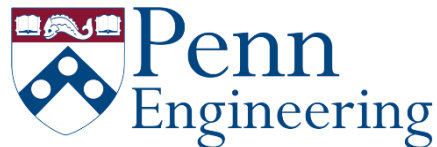


Robotics

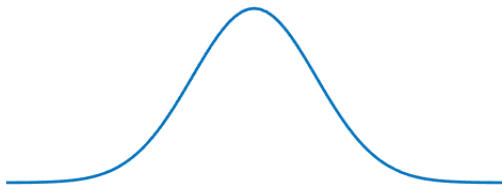
Estimation and Learning
with Dan Lee

Week 1. Gaussian Model Learning

1.4.2 GMM Parameter Estimation via EM



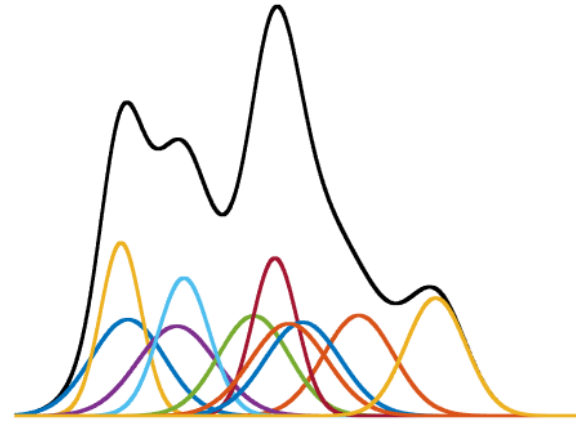
- Single Gaussian



μ

Σ

- Mixture of Gaussians



$$\mu = \{\mu_1, \mu_2, \dots, \mu_K\}$$



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

$$w = 1/K$$

K : Given

Learning GMM Parameters

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

Observed data Unknown parameters

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

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

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

Learning GMM Parameters

- Objective

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



Assuming independence of observations,

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

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!

$$\arg \max \textit{likelihood} \leftrightarrow \arg \max \ln(\textit{likelihood})$$

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

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

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!

$$p(\mathbf{x}) = \sum_{k=1}^K w_k g_k(\mathbf{x} | \boldsymbol{\mu}_k, \Sigma_k) \quad g_k: \text{Gaussian with } \boldsymbol{\mu}_k \text{ and } \Sigma_k$$



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

Learning GMM Parameters

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

where

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

→ No closed form solution exist.

Expectation-Maximization (EM)

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

EM for GMM

Initial μ and Σ

Latent variable z

EM for GMM

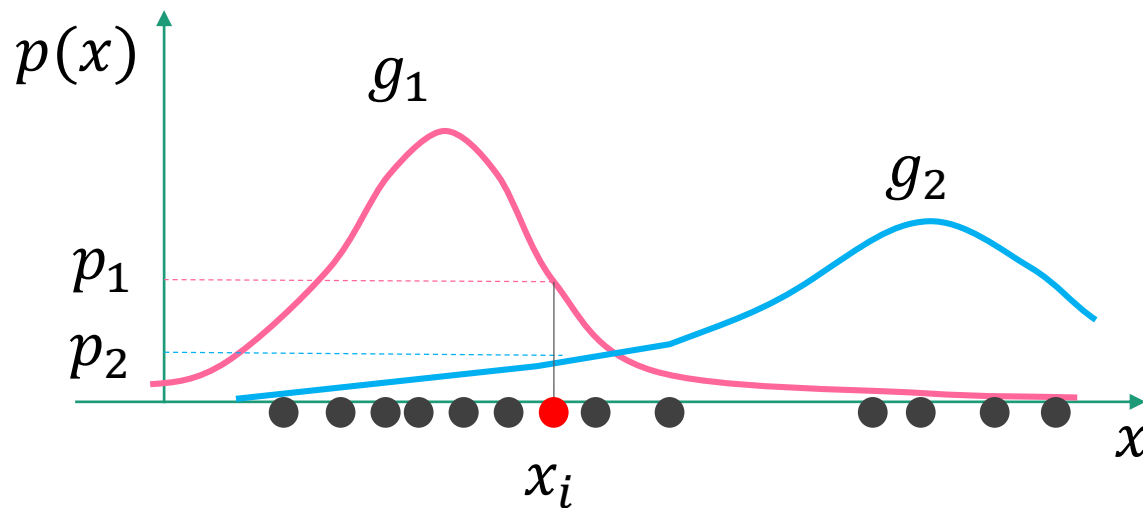
- Latent Variable

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

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

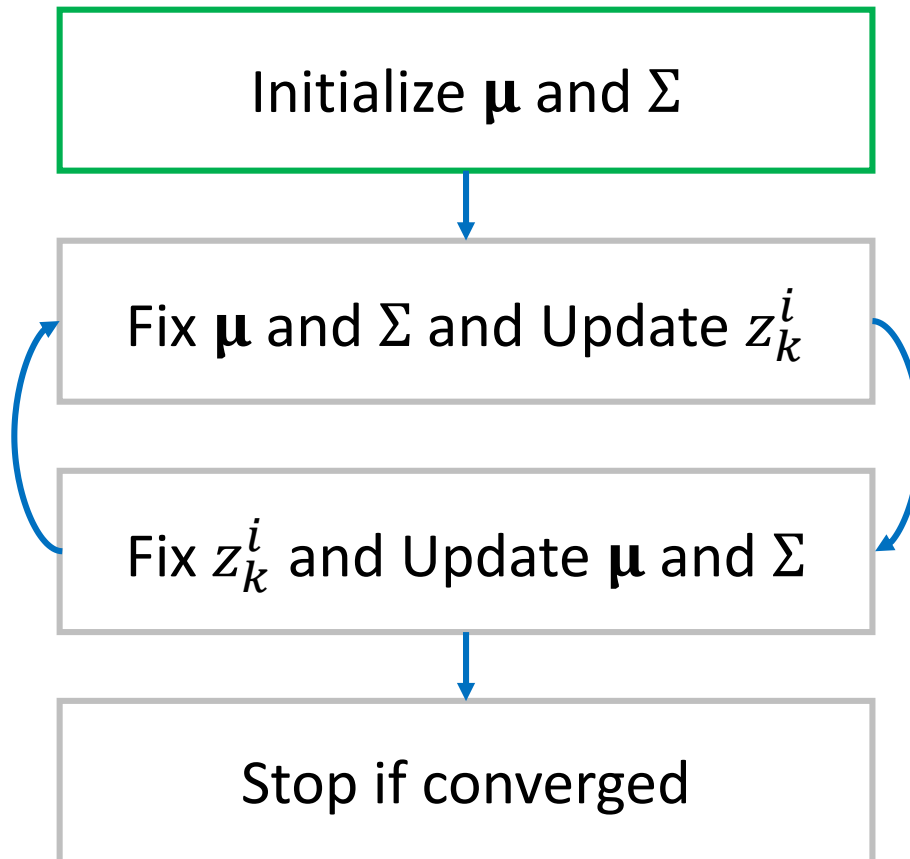
- Mean and Covariance Matrix

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

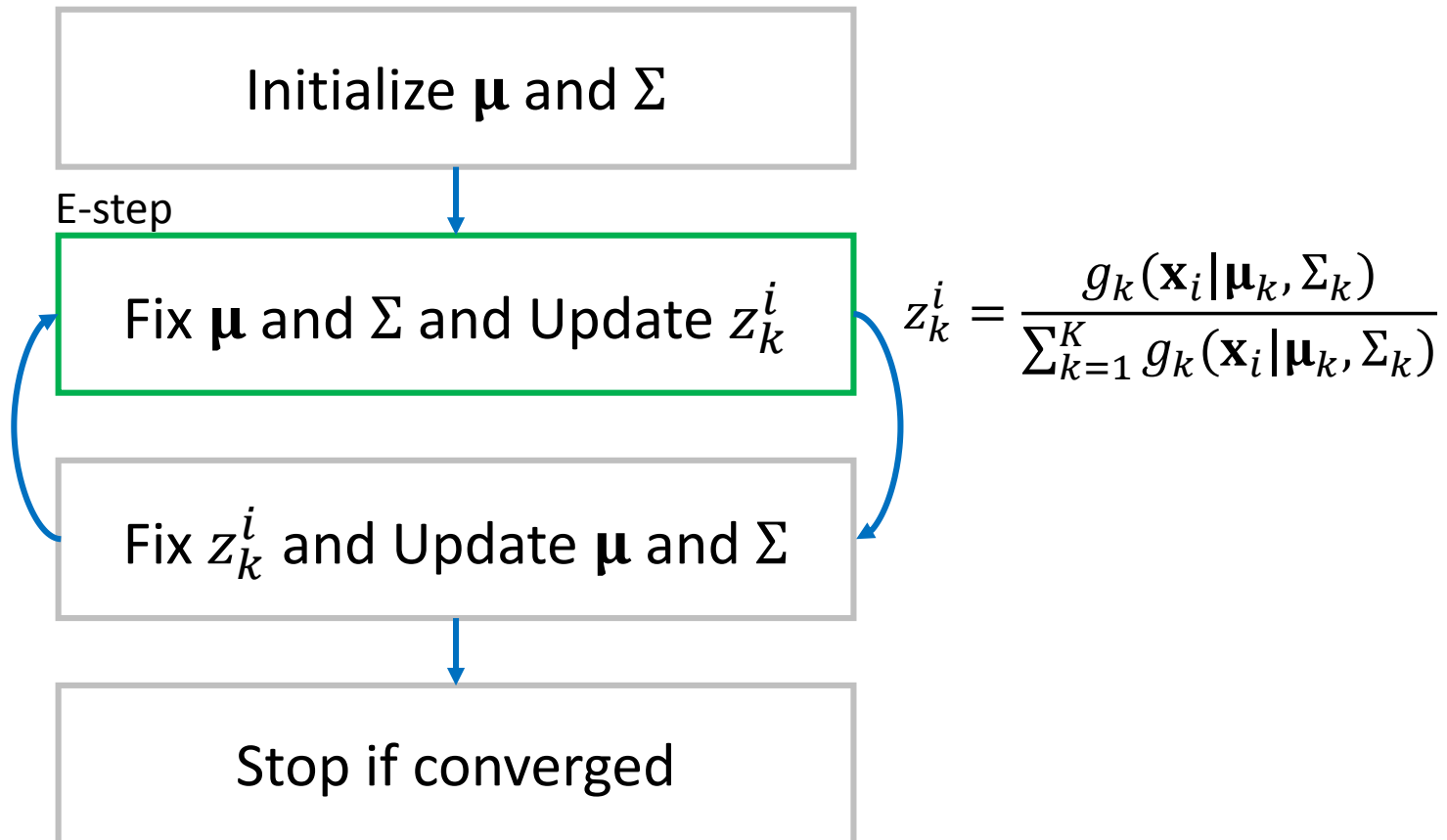
$$\hat{\boldsymbol{\Sigma}}_k = \frac{1}{z_k} \sum_{i=1}^N z_k^i (\mathbf{x}_i - \hat{\boldsymbol{\mu}}_k)(\mathbf{x}_i - \hat{\boldsymbol{\mu}}_k)^\top$$

$$z_k = \sum_{i=1}^N z_k^i$$

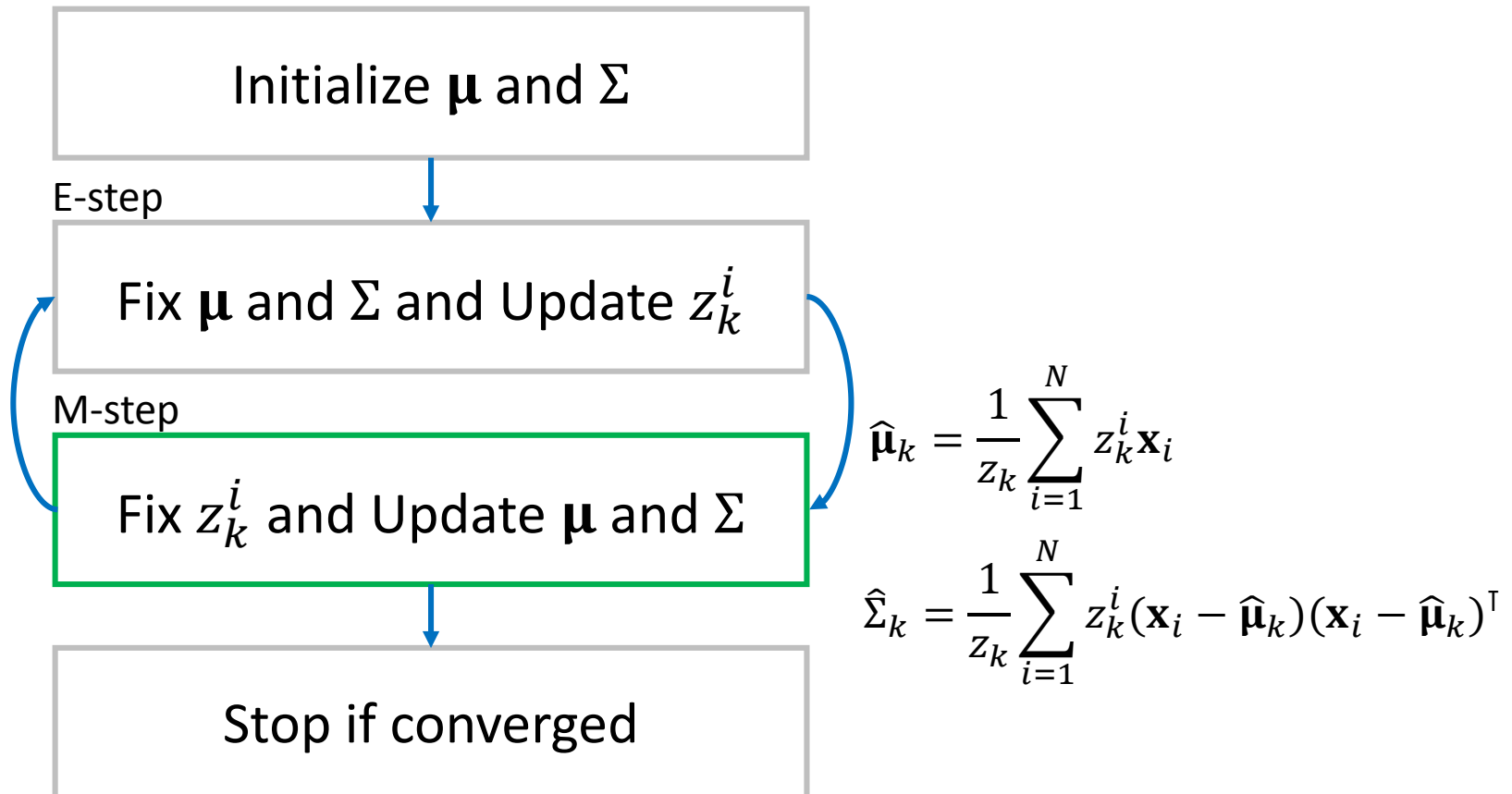
EM for GMM



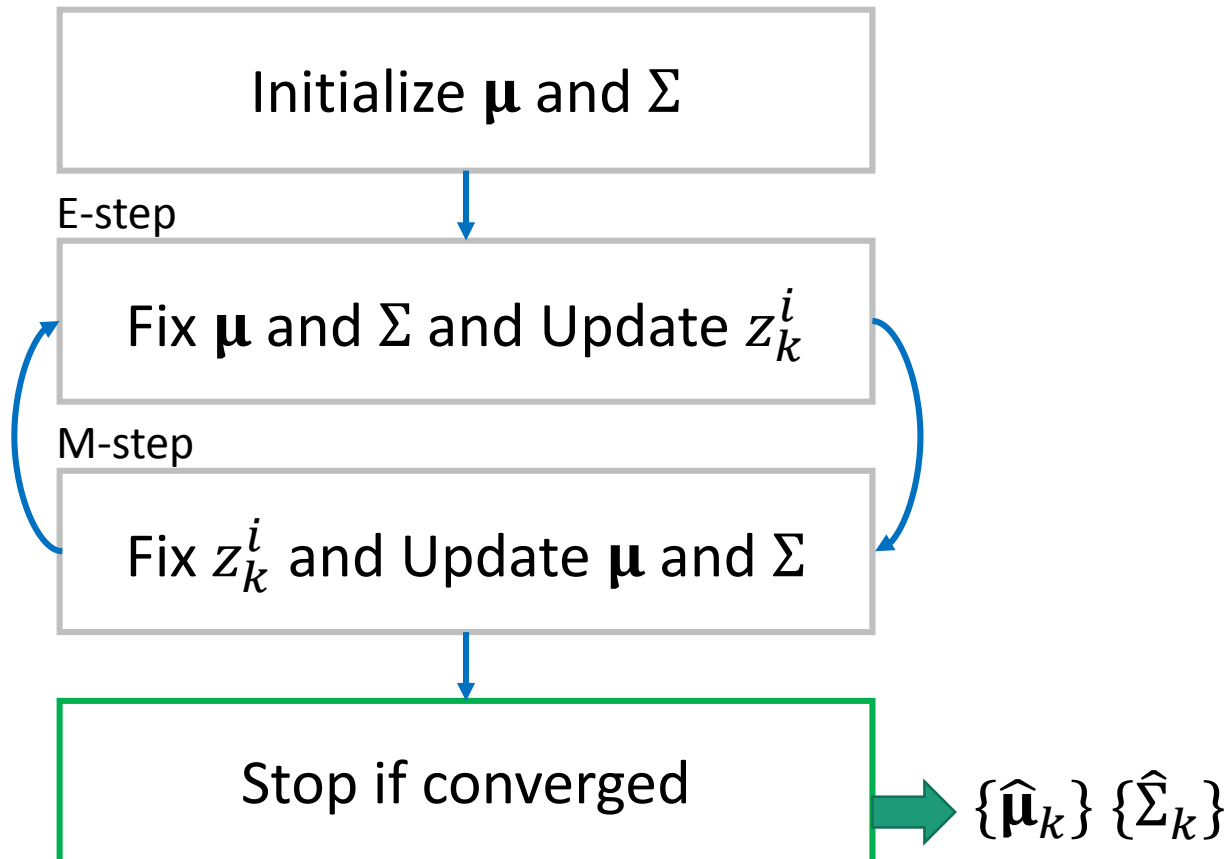
EM for GMM



EM for GMM



EM for GMM



EM for GMM

