# Week 10

# Agenda

- GMM Review
- 2. Case Study

For next week: Read 'Eigenface' paper

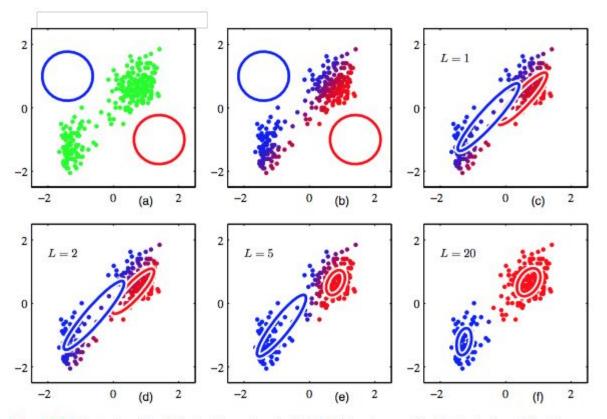


Figure 9.8 Illustration of the EM algorithm using the Old Faithful set as used for the illustration of the K-means algorithm in Figure 9.1. See the text for details.

#### **PARAMETERS**

- To build a K-Means model, how many parameters do we need to fit to the data?
- 2. What about for GMM?
- 3. What are the implications for choosing each of the four types of covariance matrices (at the left)?

#### **Pros of GMMs?**

1.

#### Cons of GMMs?

1

## EM for GMMs

• E-step: Evaluate the Responsibilities

$$\tau(z_{nk}) = \frac{\pi_k N(x_n | \mu_k, \Sigma_k)}{\sum_{j=1}^K \pi_j N(x_n | \mu_j, \Sigma_j)}$$

M-Step: Re-estimate Parameters

$$\mu_k^{new} = \frac{\sum_{n=1}^N \tau(z_{nk}) x_n}{N_k}$$

$$\Sigma_k^{new} = \frac{1}{N_k} \sum_{n=1}^{N} \tau(z_{nk}) (x_k - \mu_k^{new}) (x_k - \mu_k^{new})^T$$

$$\pi_k^{new} = \frac{N_k}{N}$$

#### Initialization

 Can use kmeans (faster). Mixing is proportional to number of examples in each cluster.

### E-Step

Compute the 'responsibilities'

#### M-Step

Update the parameters

### More 2-D Gaussians More 2-D Gaussians $\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ $\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ $\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ × 2 spherical diag Train accuracy: 88.3 Train accuracy: 88.3 Test accuracy: 92.3 Test accuracy: 94.9 x 1 x 1 full tied Train accuracy: 95.5 Train accuracy: 89.2 Test accuracy: 87.2 Test accuracy: 100.0 versicolor · · · virginica

#### **Covariance Types**

Spherical (7)(1) [ 7.5 0 0 0 7.5 0 0 0 7.5]

Diagnol (21)(3) [4 0 0 0 8.4 0 0 0 6]

Full (42)(6) [4 1 5 1 8.4 6 5 6 6] Case Study...

# Final Thoughts?