Machine Learning Worksheet 9

Latent Variable Models

1 K-Means and MoG

Problem 1. Consider a mixture of K isotropic Gaussians, each with the same covariance $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}$. In the limit $\sigma^2 \to 0$ show that the EM algorithm for MoG converges to the K-Means algorithm.

Problem 2. Consider a mixture of K Gaussians

$$p(\boldsymbol{x}) = \sum_k \pi_k \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$

Derive $E(\boldsymbol{x})$ and $Cov(\boldsymbol{x})$. It is helpful to remember the identity $Cov(\boldsymbol{x}) = E(\boldsymbol{x}\boldsymbol{x}^T) - E(\boldsymbol{x})E(\boldsymbol{x})^T$.

2 FA/pPCA and PCA

Problem 3. Consider the latent space distribution

$$p(\boldsymbol{z}) = \mathcal{N}(\boldsymbol{z}|\boldsymbol{0}, \boldsymbol{I})$$

and a conditional distribution for the observed variable $\boldsymbol{x} \in \mathbb{R}^d$,

$$p(\boldsymbol{x}|\boldsymbol{z}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{W}\boldsymbol{z} + \boldsymbol{\mu}, \boldsymbol{\Phi})$$

where $\boldsymbol{\Phi}$ is an arbitrary symmetric, positive-definite noise covariance variable. Furthermore, $\boldsymbol{A}$ is a non-singular $d \times d$ matrix and $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$. Show that for the maximum likelihood solution for the parameters of the model for $\boldsymbol{y}$ specific constraints on $\boldsymbol{\Phi}$ are preserved in the following two cases: (i) $\boldsymbol{A}$ is a diagonal matrix and $\boldsymbol{\Phi}$ is a diagonal matrix (this corresponds to the case of Factor Analysis). (ii) $\boldsymbol{A}$ is orthogonal and $\boldsymbol{\Phi} = \sigma^2 \boldsymbol{I}$ (this corresponds to pPCA).

Problem 4. Show that in the limit $\sigma^2 \to 0$ the posterior mean for the probabilistic PCA model becomes an orthogonal projection onto the same principal subspace as in PCA.