Problem Set 7

Nov. 18, 2021

1. Suppose our dataset D comes from k Gaussian distributions,

$$p_{\mathrm{M}}(x) = \sum_{j=1}^{k} \alpha_j N(x|\boldsymbol{\mu_j}, \boldsymbol{\Sigma_j})$$

where α_j, $\boldsymbol{\mu_j}$, $\boldsymbol{\Sigma_j}$ are the ratio, mean and covariance matrix of the *j*-th component of the mixed Gaussian distribution, respectively. Try to derive the process of using EM algorithm to solve the parameters of the mixed Gaussian model:

(1) According to the Bayesian formula, find the posterior expectation and write the function form that needs to be maximized;

(2) Maximum the function in (1), write down the updating formula for α_j, $\boldsymbol{\mu_j}$, $\boldsymbol{\Sigma_j}$.

Due date: Nov. 24 (Wednesday) 23:00 Beijing time