

Math 189R problem set 6

Adam Guo 2020-03-09

1. **(Murphy 11.2 - EM for Mixtures of Gaussians)** Show that the M step for ML estimation of a mixture of Gaussians is given by

$$\begin{aligned}\boldsymbol{\mu}_k &= \frac{\sum_i r_{ik} \mathbf{x}_i}{r_k} \\ \boldsymbol{\Sigma}_k &= \frac{1}{r_k} \sum_i r_{ik} (\mathbf{x}_i - \boldsymbol{\mu}_k)(\mathbf{x}_i - \boldsymbol{\mu}_k)^T = \frac{1}{r_k} \sum_i r_{ik} \mathbf{x}_i \mathbf{x}_i^T - r_k \boldsymbol{\mu}_k \boldsymbol{\mu}_k^T.\end{aligned}$$

Solution:

In Murphy, the complete data log likelihood is given by

$$\ell(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = -\frac{1}{2} \sum_i r_{ik} [\log |\boldsymbol{\Sigma}_k| + (\mathbf{x}_i - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_k)]$$

At each step, we want to find the expressions for $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ that maximise the log likelihood.

To maximise $\boldsymbol{\mu}_k$, we take the derivative of ℓ with respect to $\boldsymbol{\mu}_k$ and set it to 0.

$$\begin{aligned}\frac{d\ell}{d\boldsymbol{\mu}_k} &= \frac{d}{d\boldsymbol{\mu}_k} \left(-\frac{1}{2} \sum_i r_{ik} (-\mathbf{x}_i^T \boldsymbol{\Sigma}_k^{-1} \boldsymbol{\mu}_k - \boldsymbol{\mu}_k^T \boldsymbol{\Sigma}_k^{-1} \mathbf{x}_i + \boldsymbol{\mu}_k^T \boldsymbol{\Sigma}_k^{-1} \boldsymbol{\mu}_k) \right) \\ &= \frac{1}{2} \sum_i r_{ik} \mathbf{x}_i^T \boldsymbol{\Sigma}_k^{-1} + r_{ik} \mathbf{x}_i^T \boldsymbol{\Sigma}_k^{-T} - r_{ik} \boldsymbol{\mu}_k^T (\boldsymbol{\Sigma}_k^{-1} + \boldsymbol{\Sigma}_k^{-T}) \\ &= \left(\sum_i r_{ik} (\mathbf{x}_i^T - \boldsymbol{\mu}_k^T) \right) \boldsymbol{\Sigma}_k^{-1}, \quad \text{since } \boldsymbol{\Sigma}_k^{-1} = \boldsymbol{\Sigma}_k^{-T} \\ 0 &= \left(\sum_i r_{ik} (\mathbf{x}_i^T - \boldsymbol{\mu}_k^T) \right) \boldsymbol{\Sigma}_k^{-1} \\ &= \sum_i r_{ik} (\mathbf{x}_i^T - \boldsymbol{\mu}_k^T) \\ \sum_i r_{ik} \boldsymbol{\mu}_k^T &= \sum_i r_{ik} \mathbf{x}_i^T \\ r_k \boldsymbol{\mu}_k &= \sum_i r_{ik} \mathbf{x}_i \\ \boldsymbol{\mu}_k &= \frac{\sum_i r_{ik} \mathbf{x}_i}{r_k}, \quad \text{yielding the correct answer.}\end{aligned}$$

Performing the same steps to maximise Σ_k ,

$$\begin{aligned}
\frac{d\ell}{d\Sigma_k} &= \frac{d}{d\Sigma_k} \left(-\frac{1}{2} \sum_i r_{ik} \log |\Sigma_k| + r_{ik} (\mathbf{x}_i - \boldsymbol{\mu}_k)^T \Sigma_k^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_k) \right) \\
&= -\frac{1}{2} \sum_i (r_{ik} \Sigma_k^{-1} - r_{ik} \Sigma_k^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_k) (\mathbf{x}_i - \boldsymbol{\mu}_k)^T \Sigma_k^{-1}) \\
0 &= -\frac{1}{2} \sum_i (r_{ik} \Sigma_k^{-1} - r_{ik} \Sigma_k^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_k) (\mathbf{x}_i - \boldsymbol{\mu}_k)^T \Sigma_k^{-1}) \\
\sum_i r_{ik} \Sigma_k^{-1} &= \sum_i r_{ik} \Sigma_k^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_k) (\mathbf{x}_i - \boldsymbol{\mu}_k)^T \Sigma_k^{-1} \\
r_k \mathbf{I} &= \sum_i r_{ik} (\mathbf{x}_i - \boldsymbol{\mu}_k) (\mathbf{x}_i - \boldsymbol{\mu}_k)^T \Sigma_k^{-1} \\
\Sigma_k &= \frac{1}{r_k} \sum_i r_{ik} (\mathbf{x}_i - \boldsymbol{\mu}_k) (\mathbf{x}_i - \boldsymbol{\mu}_k)^T \\
&= \frac{1}{r_k} \sum_i r_{ik} (\mathbf{x}_i \mathbf{x}_i^T) - r_k \boldsymbol{\mu}_k \boldsymbol{\mu}_k^T, \quad \text{yielding the correct answer.}
\end{aligned}$$

2. **(SVD Image Compression)** In this problem, we will use the image of a scary clown online to perform image compression. In the starter code, we have already load the image into a matrix/array for you. However, you might need internet connection to access the image and therefore successfully run the starter code. The code requires Python library Pillow in order to run.

Plot the progression of the 100 largest singular values for the original image and a randomly shuffled version of the same image (all on the same plot). In a single figure plot a grid of four images: the original image, and a rank k truncated SVD approximation of the original image for $k \in \{2, 10, 20\}$.

Solution: