Math189R SP19
Homework 6
Monday, Apr 1, 2019

Consulted Solution for 1 and 2

Feel free to work with other students, but make sure you write up the homework and code on your own (no copying homework *or* code; no pair programming). Feel free to ask students or instructors for help debugging code or whatever else, though.

The starter files for problem 2 can be found under the Resource tab on course website. Please print out all the graphs generated by your own code and submit them together with the written part, and make sure you upload the code to your Github repository.

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

$$\boldsymbol{\mu}_k = \frac{\sum_i r_{ik}\mathbf{x}_i}{r_k}$$
$$\boldsymbol{\Sigma}_k = \frac{1}{r_k}\sum_i r_{ik}(\mathbf{x}_i - \boldsymbol{\mu}_k)(\mathbf{x}_i - \boldsymbol{\mu}_k)^\top = \frac{1}{r_k}\sum_i r_{ik}\mathbf{x}_i\mathbf{x}_i^\top - r_k\boldsymbol{\mu}_k\boldsymbol{\mu}_k^\top.$$

For a mixture of Gaussians we know that the log likelihood is

$$\begin{aligned}\ell(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) &= \sum_k\sum_i r_{ik}\log\mathbb{P}(\mathbf{x}_i|\theta_k) \\ &= -\frac{1}{2}\sum_i r_{ik}(\log|\boldsymbol{\Sigma}_k| + (\mathbf{x}_i - \boldsymbol{\mu}_k)^\top\boldsymbol{\Sigma}_k^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_k)).\end{aligned}$$

To find the optimization for $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ we solve for the roots of the derivative of the log likelihood with respect to each variable.

Let us first solve for $\boldsymbol{\mu}_k$.

$$\begin{aligned}\frac{\partial\ell}{\partial\boldsymbol{\mu}_k} &= -\frac{1}{2}\sum_i -2 * r_{ik}\boldsymbol{\Sigma}_k^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_k) \\ &= \boldsymbol{\Sigma}_k^{-1}\sum_i r_{ik}(\mathbf{x}_i - \boldsymbol{\mu}_k).\end{aligned}$$

The roots of this expression - solving for $\boldsymbol{\mu}_k$ - are then

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

as desired.

Let us solve for $\boldsymbol{\Sigma}_k$.

$$
\frac{\partial \ell}{\partial \boldsymbol{\Sigma}_k} = -\frac{1}{2} \sum_i r_{ik} (\boldsymbol{\Sigma}_k^{-1} - (\mathbf{x}_i - \boldsymbol{\mu}_k)^\top (\mathbf{x}_i - \boldsymbol{\mu}_k) \boldsymbol{\Sigma}_k^{-2})
$$

The roots of this expression - solving for $\boldsymbol{\Sigma}_k$ - are then

$$
\begin{aligned}
0 &= -\frac{1}{2} \sum_i r_{ik} (\boldsymbol{\Sigma}_k^{-1} - (\mathbf{x}_i - \boldsymbol{\mu}_k)^\top (\mathbf{x}_i - \boldsymbol{\mu}_k) \boldsymbol{\Sigma}_k^{-2}) \\
\sum_i r_{ik} \boldsymbol{\Sigma}_k^{-1} &= \sum_i r_{ik} (\mathbf{x}_i - \boldsymbol{\mu}_k)^\top (\mathbf{x}_i - \boldsymbol{\mu}_k) \boldsymbol{\Sigma}_k^{-2} \\
\boldsymbol{\Sigma}_k &= \boxed{\frac{\sum_i r_{ik} (\mathbf{x}_i - \boldsymbol{\mu}_k)^\top (\mathbf{x}_i - \boldsymbol{\mu}_k)}{\sum_i r_{ik}},}
\end{aligned}
$$

we have then show that the the M step for ML estimation of a mixture of Gaussians is given by the formulas derived above. ■

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 $k \in \{2, 10, 20\}$.

Finished code but was not able to run because commands are deprecated. ■