CSCI 5521: Homework 3

Christopher White

April 7, 2019

Problem 1

Probability Density Function

$$p\left(x|C_i\right) = \frac{1}{2b_i}\exp\left(-\frac{|x-\mu_i|}{b_i}\right), b_i > 0 \tag{1}$$

Mixture Density

$$P\left(x\right) = \sum_{i=1}^{K}\pi_i\frac{1}{2b_i}\exp\left(-\frac{|x-\mu_i|}{b_i}\right) \tag{2}$$

Likelihood Function

$$\begin{aligned}\mathcal{L}\left(\Phi|\mathcal{X}\right) &= \sum_t \log\sum_{i=1}^{k} p\left(x^t|C_i\right)P\left(C_i\right)\\ &= \sum_t \log\sum_{i=1}^{k}\pi_i\frac{1}{2b_i}\exp\left(-\frac{|x^t-\mu_i|}{b_i}\right)\end{aligned} \tag{3}$$

Complete Likelihood Function

$$\begin{aligned}\mathcal{L}_c\left(\Phi|\mathcal{X},\mathcal{Z}\right) &= \sum_t\sum_i z_i^t\left[\log\pi_i + \log p_i(x^t|\Phi^l)\right]\\ &= \sum_t\sum_i z_i^t\left[\log\pi_i + \log\frac{1}{2b_i}\exp\left(-\frac{|x^t-\mu_i|}{b_i}\right)\right]\\ &= \sum_t\sum_i z_i^t\left[\log\pi_i - \log 2b_i - \left(\frac{|x^t-\mu_i|}{b_i}\right)\right]\end{aligned} \tag{4}$$

E-Step:

$$\mathcal{Q}(\Phi|\Phi^l) = \sum_t\sum_i E[z_i^t|\mathcal{X},\Phi^l][\log\pi_i + \log p_i(x^t|\Phi^l)] \tag{5}$$

$$E[z_i^t|\mathcal{X},\Phi^l] = \frac{p(x^t|C_i,\Phi^l)P(C_i)}{\sum_j p(x^t|C_j,\Phi^l)P(C_j)} \equiv \gamma(z_i^t) \tag{6}$$

$$\gamma(z_i^t) \equiv h_i^t = \frac{\pi_i\frac{1}{2b_i}\exp\left(-\frac{|x^t-\mu_i|}{b_i}\right)}{\sum_{j=1}^{k}\pi_j\frac{1}{2b_j}\exp\left(-\frac{|x^t-\mu_j|}{b_j}\right)} \tag{7}$$

M-Step:

$$\begin{aligned}
\mathcal{Q}(\Phi|\Phi^l) &= \sum_t \sum_i h_i^t \left[\log \pi_i + \log p_i(x^t|\Phi^l)\right] \\
&= \sum_t \sum_i h_i^t \left[\log \pi_i + \log \left(\frac{1}{2b_i} \exp\left(-\frac{|x^t - \mu_i|}{b_i}\right)\right)\right] \\
&= \sum_t \sum_i h_i^t \left[\log \pi_i - \log 2b_i - \frac{|x^t - \mu_i|}{b_i}\right]
\end{aligned} \tag{8}$$

$$\sum_i \pi_i = 1$$

Estimate of π_i:

$$\frac{\partial}{\partial \pi_i} \sum_t \log \sum_i \pi_i \left(\frac{1}{2b_i} \exp\left(-\frac{|x - \mu_i|}{b_i}\right)\right) - \alpha(\sum_i \pi_i - 1) = 0 \tag{9}$$

$$\sum_t \frac{\frac{1}{2b_i} \exp\left(-\frac{|x-\mu_i|}{b_i}\right)}{\sum_j \pi_j \left(\frac{1}{2b_j} \exp\left(-\frac{|x^t-\mu_j|}{b_j}\right)\right)} = \alpha \tag{10}$$

$$\sum_t \frac{\pi_i \frac{1}{2b_i} \exp\left(-\frac{|x-\mu_i|}{b_i}\right)}{\sum_j \pi_j \left(\frac{1}{2b_j} \exp\left(-\frac{|x^t-\mu_j|}{b_j}\right)\right)} = \alpha \pi_i = \sum_t \gamma(z_i^t) \tag{11}$$

$$\pi_i = \frac{N_i}{N} \tag{12}$$

Estimate of b_i:

$$\frac{\partial}{\partial b_i}\sum_t \log \sum_i \pi_i \left(\frac{1}{2b_i}\exp\left(-\frac{|x-\mu_i|}{b_i}\right)\right) = 0 \tag{13}$$

$$\sum_t \frac{\pi_i \frac{2b_i \frac{|x^t-\mu_i|}{b_i^2}\exp\left(-\frac{|x^t-\mu_i|}{b_i}\right) - 2\exp\left(-\frac{|x^t-\mu_i|}{b_i}\right)}{4b_i^2}}{\sum_j \pi_j \left(\frac{1}{2b_j}\exp\left(-\frac{|x^t-\mu_j|}{b_j}\right)\right)} = 0 \tag{14}$$

$$\sum_t \frac{\pi_i \frac{\left(\frac{|x^t-\mu_i|}{b_i} - 1\right)\exp\left(-\frac{|x^t-\mu_i|}{b_i}\right)}{2b_i^2}}{\sum_j \pi_j \left(\frac{1}{2b_j}\exp\left(-\frac{|x^t-\mu_j|}{b_j}\right)\right)} = 0 \tag{15}$$

$$\sum_t \frac{\pi_i \frac{\exp\left(-\frac{|x^t-\mu_i|}{b_i}\right)}{2b_i}}{\sum_j \pi_j \left(\frac{1}{2b_j}\exp\left(-\frac{|x^t-\mu_j|}{b_j}\right)\right)} \frac{\left(\frac{|x^t-\mu_i|}{b_i} - 1\right)}{b_i} = 0 \tag{16}$$

$$\sum_t \frac{\pi_i \frac{\exp\left(-\frac{|x^t-\mu_i|}{b_i}\right)}{2b_i}}{\sum_j \pi_j \left(\frac{1}{2b_j}\exp\left(-\frac{|x^t-\mu_j|}{b_j}\right)\right)} \left(|x^t-\mu_i| - 1\right) = 0 \tag{17}$$

$$\sum_t \gamma(z_i^t)\left(|x^t-\mu_i| - b_i\right) = 0 \tag{18}$$

$$\sum_t \gamma(z_i^t)|x^t-\mu_i| - \gamma(z_i^t) b_i = 0 \tag{19}$$

$$\sum_t \gamma(z_i^t) b_i = \sum_t \gamma(z_i^t)|x^t-\mu_i| \tag{20}$$

$$b_i = \frac{\sum_t \gamma(z_i^t)|x^t-\mu_i|}{\sum_t \gamma(z_i^t)} \tag{21}$$

Estimate of μ_i:

$$\frac{\partial}{\partial \mu_i}\sum_t \log \sum_i \pi_i \left(\frac{1}{2b_i}\exp\left(-\frac{|x^t-\mu_i|}{b_i}\right)\right) = 0 \tag{22}$$

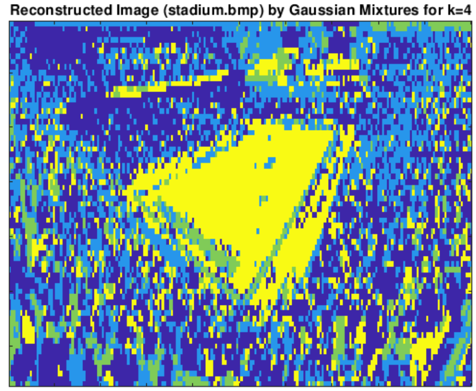
$$\sum_t \gamma(z_i^t)\frac{\partial}{\partial \mu_i}\left(-\frac{|x^t-\mu_i|}{b_i}\right) = 0 \tag{23}$$

If $\gamma(z_i^t) = 1$ for $i = argmax_j \gamma(z_j^t)$ and $\sum_t |\theta - x^t|$ is minimized for $\theta = x^{N/2}$, then if x^t are sorted the $\mu_i = x^{N_i/2}$.

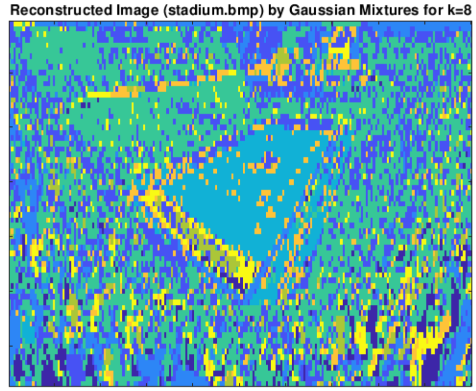
Problem 2

(a) The following images are for the EM algorithm for $k = \{4, 8, 12\}$.

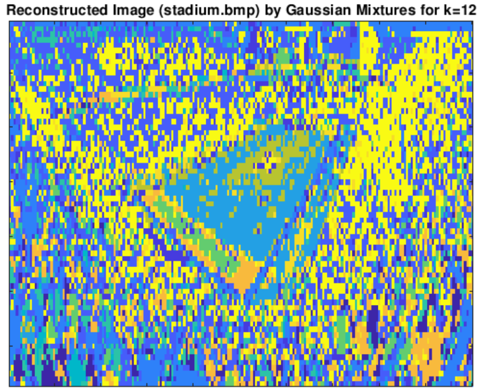
Reconstructed Image (stadium.bmp) by Gaussian Mixtures for k=4



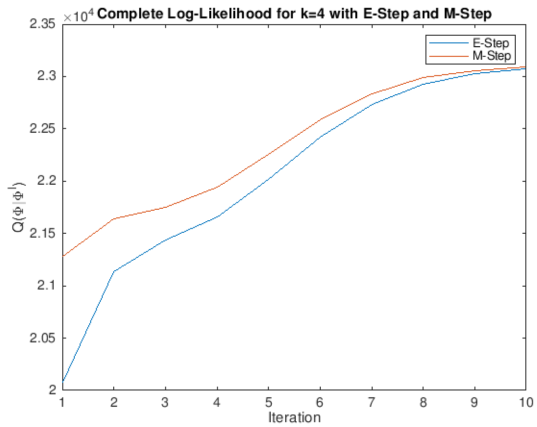
Reconstructed Image (stadium.bmp) by Gaussian Mixtures for k=8

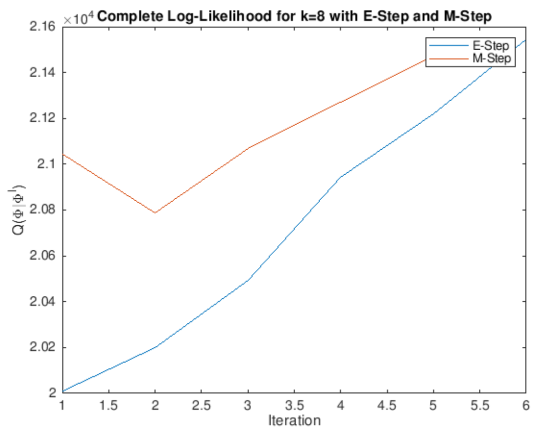


Reconstructed Image (stadium.bmp) by Gaussian Mixtures for k=12



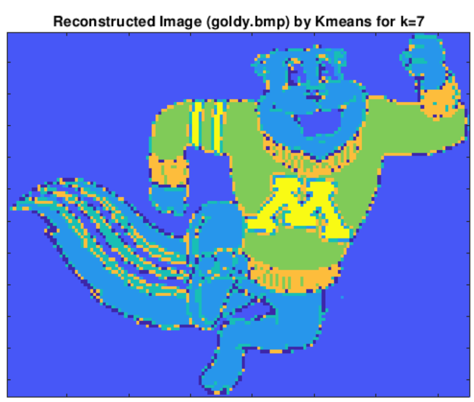
(b) The graphs below show the log likelihood function, $Q(\Phi|\Phi^l)$, for the E-step and M-step of the algorithm for $k = \{4, 8, 12\}$. The function did not plot for the $k = 12$ setup because the $\mathcal{Q}$ grew so quickly that only the first point was plotted.





(c) The kmeans image is show below. The EM algorithm failed because the covariance matrix was singular, which means that it cannot be inverted. Because it is invertible, it cannot be used in the PDF calculation. Therefore, the clustering was unable to proceed.

The kmeans did not fail. The kmeans algorithm does not rely on a covariance matrix that needs inverting. Therefore, it is safe from singular matrices.



(d) The following is the derivation of the regularized Σ_is.

$$\mathcal{L}_c\left(\mu, \Sigma|\mathcal{X}\right) = \sum_t \sum_{i=1}^{k} h_i^t \left(\log \pi_i + \log \mathcal{N}\left(x^t|\mu_i, \Sigma_i\right) - \frac{\lambda}{2}\sum_{j=1}^{d}(\Sigma_i^{-1})_{jj}\right) \tag{24}$$

$$\mathcal{L}_c\left(\mu, \Sigma|\mathcal{X}\right) = \sum_t \sum_{i=1}^{k} h_i^t \left(\log \pi_i - \frac{d}{2}\log 2\pi - \frac{1}{2}\log|\Sigma_i| - \frac{1}{2}(x^t-\mu_i)^T\Sigma_i^{-1}(x^t-\mu_i) - \frac{\lambda}{2}\sum_{j=1}^{d}(\Sigma_i^{-1})_{jj}\right) \tag{25}$$

$$\frac{\partial\mathcal{L}_c\left(\mu, \Sigma|\mathcal{X}\right)}{\partial\Sigma_i^{-1}} = \sum_t h_i^t \frac{\partial}{\partial\Sigma_i^{-1}} \left(\log \pi_i - \frac{d}{2}\log 2\pi - \frac{1}{2}\log|\Sigma_i| - \frac{1}{2}(x^t-\mu_i)^T\Sigma_i^{-1}(x^t-\mu_i) - \frac{\lambda}{2}\sum_{j=1}^{d}(\Sigma_i^{-1})_{jj}\right) \tag{26}$$

$$= \frac{N_i}{2}\Sigma_i - \frac{\lambda N_i I}{2} - \frac{1}{2}\sum_t z_i^t(x^t-\mu_i)(x^t-\mu_i)^T \tag{27}$$

$$\Sigma_i = \lambda I + \frac{1}{N_i}\sum_t h_i^t(x^t-\mu_i)(x^t-\mu_i)^T \tag{28}$$

(e) The final image generated by the modified code is below. This image did not have the singular matrix failure. The colors assigned to the pixels appear less blurry or jagged compared to the prior kmeans image.

Reconstructed Image (goldy.bmp) by Gaussian Mixtures for k=7