Weekly Exercises 3

Room: 02.05.023

Tuesday, 10.11.2015, 14:15-15:45

Submission deadline: Tuesday, 10.11.2015, 11:15 , Room 02.09.023

Gaussian Mixture Models (8 Points)

Exercise 1 (Probability, 2 Points)**.** Prove the following identities:

a)
$$\int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t = \sqrt{2\pi}.$$

b)
$$\int_{-\infty}^{\infty} t \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t = 0.$$

Note: we define $\int_{-\infty}^{\infty} f(t)\ \mathrm{d}t := \lim_{x\to\infty} \int_{-x}^{x} f(t)\ \mathrm{d}t$.

Solution.

a)

$$\begin{aligned}
\left[\int_{-\infty}^{\infty} \exp\left(\frac{-t^2}{2\sigma^2}\right)\ \mathrm{d}t\right]^2 &= \int_{-\infty}^{\infty} \exp\left(\frac{-x^2}{2\sigma^2}\right)\ \mathrm{d}x \int_{-\infty}^{\infty} \exp\left(\frac{-y^2}{2\sigma^2}\right)\ \mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(\frac{-(x^2+y^2)}{2\sigma^2}\right)\ \mathrm{d}x\mathrm{d}y \\
&= \int_{0}^{2\pi}\int_{0}^{\infty} r\exp\left(\frac{-r^2}{2\sigma^2}\right)\ \mathrm{d}r\mathrm{d}\theta \\
&= 2\pi\int_{0}^{\infty} r\exp\left(\frac{-r^2}{2\sigma^2}\right)\ \mathrm{d}r \\
&= 2\pi\int_{-\infty}^{0} \exp\left(s\right)\ \mathrm{d}s \qquad \left(s = -\frac{r^2}{2\sigma^2} \Rightarrow \mathrm{d}r = -\sigma^2\mathrm{d}s.\right) \\
&= 2\pi\sigma^2
\end{aligned}$$

b)

$$\begin{aligned}
\int_{-\infty}^{\infty} t\exp\left(-\frac{t^2}{2\sigma^2}\right)\mathrm{d}t &= \frac{1}{2}\int_{-\infty}^{0}\exp\left(-\frac{t^2}{2\sigma^2}\right)\mathrm{d}t^2 + \frac{1}{2}\int_{0}^{\infty}\exp\left(-\frac{t^2}{2\sigma^2}\right)\mathrm{d}t^2 \\
&= -\sigma^2\left[\exp\left(-\frac{t^2}{2\sigma^2}\right)\right]_{-\infty}^{0} - \sigma^2\left[\exp\left(-\frac{t^2}{2\sigma^2}\right)\right]_{0}^{\infty} = -\sigma^2 - (-\sigma^2) = 0.
\end{aligned}$$

Exercise 2 (Probability, 2 Point)**.** Show that the expected value of the univariate Gaussian distribution

$$\mathcal{N}(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},$$

is given by $\mathbb{E}[x] = \mu$.

Solution.

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \,\mathrm{d}x = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (t+\mu) \exp\left(-\frac{t^2}{2\sigma^2}\right) \,\mathrm{d}t$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}} \left(\mu \int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2\sigma^2}\right) \,\mathrm{d}t + \int_{-\infty}^{\infty} t \exp\left(-\frac{t^2}{2\sigma^2}\right) \,\mathrm{d}t\right)$$
$$\overset{1a) + 1b)}{=} \frac{1}{\sqrt{2\pi\sigma^2}} (\mu\sqrt{2\pi}\sigma + 0) = \mu.$$

Exercise 3 (EM, 4 Points)**.** Use the general expectation maximization principle to derive an algorithm for maximizing the likelihood of the Gaussian mixture model. For that, we introduce *latent* binary variables z_{nk} which are 1 if the k-th component generated the n-th data point, and zero otherwise. With these latent variables, the likelihood takes the form:

$$p(X, Z \mid \theta) = \prod_{n=1}^{N} \prod_{k=1}^{K} \pi_k^{z_{nk}} \mathcal{N}(x_n \mid \mu_k, \Sigma_k)^{z_{nk}}, \qquad \theta = \{\pi, \mu, \Sigma\}.$$

a) **E step:** Using Bayes' theorem, show that the posterior is given by:

$$p(Z \mid X, \theta^{\text{old}}) \propto \prod_{n=1}^{N} \prod_{k=1}^{K} \left[\pi_k^{\text{old}} \mathcal{N}(x_n \mid \mu_k^{\text{old}}, \Sigma_k^{\text{old}})\right]^{z_{nk}}.$$

b) **M step:** Derive a closed form solution for the maximization problem:

$$\theta^{\text{new}} = \underset{\theta}{\operatorname{argmax}} \left\{\sum_{Z} p(Z \mid X, \theta^{\text{old}}) \ln p(X, Z \mid \theta)\right\}.$$

Solution.

a) Using Bayes' theorem, we get:

$$p(z_n \mid x_n, \theta^{\text{old}}) = \frac{p(x_n \mid \theta^{\text{old}}, z_n) p(z_n \mid \theta^{\text{old}})}{p(x_n \mid \theta^{\text{old}})}. \tag{1}$$

The individual terms are given as:

$$\begin{aligned} p(x_n \mid \theta^{\text{old}}, z_n) &= \prod_{k=1}^{K} \mathcal{N}(x_n \mid \Sigma_k^{\text{old}}, \mu_k^{\text{old}})^{z_{nk}}, \\ p(z_n \mid \theta^{\text{old}}) &= \prod_{k=1}^{K} \left(\pi_k^{\text{old}}\right)^{z_{nk}}. \end{aligned} \tag{2}$$

Since the data points x_n are identically and independently distributed, combining (2) into (1) yields the desired result.

b)

$$\begin{aligned}
\mathcal{Q}(\theta, \theta^{\text{old}}) &= \sum_Z p(Z \mid X, \theta^{\text{old}}) \ln p(X, Z \mid \theta) \\
&= \mathbb{E}_Z \left[\ln p(X, Z \mid \theta) \mid X, \theta^{\text{old}} \right] \\
&= \mathbb{E}_Z \left[\sum_{n=1}^{N} \sum_{k=1}^{K} (\ln \pi_k + \ln \mathcal{N}(x_n \mid \mu_k, \Sigma_k)) z_{nk} \middle| X, \theta^{\text{old}} \right] \\
&= \sum_{n=1}^{N} \sum_{k=1}^{K} \mathbb{E}_{z_n} \left[z_{nk} | X, \theta^{\text{old}} \right] (\ln \pi_k + \ln \mathcal{N}(x_n \mid \mu_k, \Sigma_k)) \\
&= \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma_{nk} (\ln \pi_k + \ln \mathcal{N}(x_n \mid \mu_k, \Sigma_k)).
\end{aligned} \tag{3}$$

To see the last step in the calculation above:

$$\begin{aligned}
\mathbb{E}_{z_n}[z_{nk}|X, \theta] := \sum_{z_n} z_{nk} p(z_n \mid x_n, \theta) &= \frac{\sum_{z_n} z_{nk} p(x_n|\theta, z_n) p(z_n|\theta)}{p(x_n \mid \theta)} \\
&= \frac{\sum_{z_n} z_{nk} \prod_{k'=1}^{K} [\mathcal{N}(x_n \mid \mu_{k'}, \Sigma_{k'})]^{z_{nk'}} [\pi_{k'}]^{z_{nk'}}}{\sum_{z_n} \prod_{j=1}^{K} [\pi_j \mathcal{N}(x_n \mid \mu_j, \Sigma_j)]^{z_{nj}}} \\
&= \frac{\sum_{z_n} z_{nk} \prod_{k'=1}^{K} [\pi_{k'} \mathcal{N}(x_n \mid \mu_{k'}, \Sigma_{k'})]^{z_{nk'}}}{\sum_{z_n} \prod_{j=1}^{K} [\pi_j \mathcal{N}(x_n \mid \mu_j, \Sigma_j)]^{z_{nj}}} \\
&= \frac{\pi_k \mathcal{N}(x_n \mid \mu_k, \Sigma_k)}{\sum_{j=1}^{K} \pi_j \mathcal{N}(x_n \mid \mu_j, \Sigma_j)} =: \gamma_{nk}.
\end{aligned} \tag{4}$$

Following the lecture we have:

$$\frac{\partial \mathcal{Q}}{\partial \pi_l} \stackrel{!}{=} 0 \Rightarrow \pi_l = \frac{\sum_{n=1}^{N} \gamma_{nl}}{N}. \tag{5}$$

$$\frac{\partial \mathcal{Q}}{\partial \Sigma_l} \stackrel{!}{=} 0 \Rightarrow \Sigma_l = \frac{\sum_{n=1}^{N} \gamma_{nl} (x_n - \mu_l)(x_n - \mu_l)^T}{\sum_{n=1}^{N} \gamma_{nl}}. \tag{6}$$

$$\frac{\partial \mathcal{Q}}{\partial \mu_l} \stackrel{!}{=} 0 \Rightarrow \mu_l = \frac{\sum_{n=1}^{N} \gamma_{nl} x_n}{\sum_{n=1}^{N} \gamma_{nl}}. \tag{7}$$

Maximum Flow and Minimum Cut (7 Points)

Exercise 4 (Min-Cut, 3 Points)**.** Show that for any $x \in \mathbb{B}^n$, it holds

$$\overline{x}_1 + \left[\sum_{i=1}^{n-1} x_i \overline{x}_{i+1} - \overline{x}_i x_{i+1} \right] + x_n = 1.$$

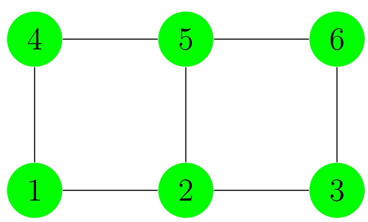
Solution. For $n = 1$ the proposition is correct as $\overline{x}_1 + x_1 = 1$ for any $x \in \mathbb{B}$. Note that $a\overline{b} - \overline{a}b + b = a$ for any $a, b \in \mathbb{B}$. Now assume the proposition is correct for $n - 1$, it holds

$$
\begin{aligned}
&\overline{x}_1 + [\sum_{i=1}^{n-1} x_i\overline{x}_{i+1} - \overline{x}_i x_{i+1}] + x_n \\
=&\overline{x}_1 + [\sum_{i=1}^{n-2} x_i\overline{x}_{i+1} - \overline{x}_i x_{i+1}] + x_{n-1}\overline{x}_n - \overline{x}_{n-1}x_n + x_n \\
=&\overline{x}_1 + [\sum_{i=1}^{n-2} x_i\overline{x}_{i+1} - \overline{x}_i x_{i+1}] + x_{n-1} = 1.
\end{aligned}
$$

$\square$

Exercise 5 (Graph Cut, 4 Points)**.**

a) Let $P = \{1, \ldots, 6\}$ be a set of pixels with a neighbourhood as depicted below:



Let $f_1, \ldots, f_6 \in \mathbb{R}$ be pixel weights and let $c_{ij} \in \mathbb{R}_{\geq 0}, i \in \{1, \ldots, 6\}, j \in \mathcal{N}(i)$, be weights for the length term. Consider the minimization problem:

$$\min_{x \in \mathbb{B}^6} \sum_{i=1}^{6} f_i x_i + \sum_{i=1}^{6} \sum_{j \in \mathcal{N}(i)} c_{ij} x_i \overline{x}_j. \tag{8}$$

Draw a network, such that the minimum s, t-cut of the network gives a solution to Problem (8).

b) Now consider the general problem. Let $P = \{1, \ldots, n\}, n \in \mathbb{N}$, be a set of pixels with a given neighbourhood structure $\mathcal{N} : \{1, \ldots, n\} \to \mathcal{P}(\{1, \ldots, n\})$. Construct a network that gives a solution to the minimization problem

$$\min_{x \in \mathbb{B}^n} \sum_{i=1}^{n} f_i x_i + \sum_{i=1}^{n} \sum_{j \in \mathcal{N}(i)} c_{ij} x_i \overline{x}_j, \tag{9}$$

where $f_1, \ldots, f_n \in \mathbb{R}$ and $c_{ij} \in \mathbb{R}_{\geq 0}, i \in \{1, \ldots, n\}, j \in \mathcal{N}(i)$, are given.

c) Let $n = 6$ and consider the same neighbourhood $\mathcal{N}$ as in Part a). Let $f_1 = 15, f_2 = -5, f_3 = 1, f_4 = 2, f_5 = -4, f_6 = -10$ and let $c_{ij} = 1$ for all $i, j \in \mathcal{N}(i)$.

Compute the maximum flow of the network from Part a) using the Ford-Fulkerson algorithm. Draw the flow at each iteration. How do you select the augmenting path at each iteration to terminate as early as possible? What is the minimum energy? What is the resulting segmentation?

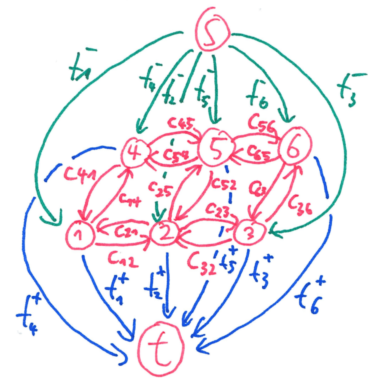
Solution. b) Define the netwok $(V, E, \hat{c}, s, t)$, where

$$\begin{aligned} V &= P \cup \{s, t\}, \\ E &= E_s \cup E_t \cup E_{\mathcal{N}}, \\ E_s &= \{(s, i) | i \in P\}, \\ E_t &= \{(i, t) | i \in P\}, \\ E_{\mathcal{N}} &= \{(i, j) | i \in P, j \in \mathcal{N}(i)\}, \\ \forall i \in P: \ \hat{c}(s, i) &= \max\{0, -f_i\}, \\ \forall i \in P: \ \hat{c}(i, t) &= \max\{0, f_i\} \text{ and} \\ \forall i \in P: \ \forall j \in \mathcal{N}(i): \ \hat{c}(i, j) &= c_{ij}. \end{aligned}$$

We can convert the minimum s, t-cut S, T of this network into a minimizer $x^* \in \mathbb{B}^n$ of Minimization Problem (9) by setting

$$x_i^* = \begin{cases} 0 & \text{if } i \in S \text{ and} \\ 1 & \text{otherwise.} \end{cases}$$

a)

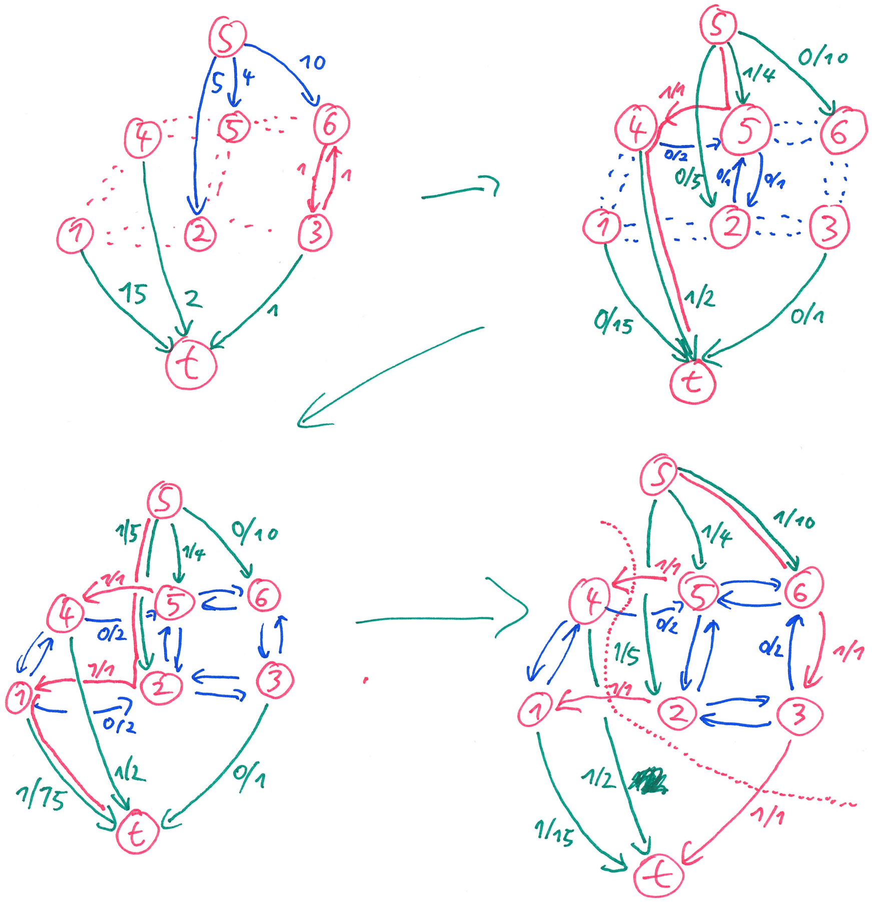


where $t_i = t_i^+ - t_i^-$

$t_i^+ = \max\{0, t_i\}$

$t_i^- = \max\{0, -t_i\}$

c)



Programming (15 Points)

Exercise 6 (5 Points)**.** Train two Gaussian mixture models which model the probability of a pixel belonging to the foreground respectively the background, using the `VLFeat`[1] library. Your program should take the following parameters as a command line argument:

- The filename of the input image.
- Coordinates of a bounding box (x_1, y_1, x_2, y_2), roughly indicating the foreground region.
- Number of components K in the Gaussian mixture model.

It should perform the following tasks:

1) Read the input image.

2) Train two Gaussian mixtures models $p_F(x)$ and $p_B(x)$ with K components each, which model the probability distribution of a pixel i belonging to the foreground and respectively the background. The points for the foreground GMM correspond to the RGB values of the pixels inside the rectangle (x_1, y_1, x_2, y_2), and the points for the background GMM correspond to the RGB values outside the rectangle.

3) Using the two GMMs, compute the *dataterm* at each pixel, which is the difference between the log-likelihood of a pixel belonging to the foreground or the background, i.e.

$$f_i = -\log \frac{p_F(x_i)}{p_B(x_i)} = \log(p_B(x_i)) - \log(p_F(x_i)).$$

Compute a global minimizer $\widehat{x} \in \mathbb{B}^n$ of the modular energy:

$$\widehat{x} = \underset{x \in \mathbb{B}^n}{\operatorname{argmin}} \sum_{i=1}^{n} f_i x_i,$$

and visualize the result. Try out different values of K (start with $K = 5$) and images. A possible output on `banana3.bmp` for $(x_1, y_1, x_2, y_2) = (160, 280, 510, 350)$ and $K = 5$ is shown below:

Hint: You can use the file `gmm.cpp` in `03_supp.zip` as a starting point for your program. `03_supp.zip` can be downloaded from the lecture homepage.

[1] `http://www.vlfeat.org/api/gmm.html`

Exercise 7 (5 Points)**.** The goal of this exercise is to use OpenGM to minimize the following submodular energy:

$$E(x) = \sum_{i=1}^{n} f_i x_i + \sum_{i=1}^{n} \sum_{j \in \mathcal{N}(i)} c_{ij} x_i \overline{x}_j, \tag{10}$$

where $x \in \mathbb{B}^n$, f_i denotes the dataterm from the previous exercise, $\mathcal{N}$ is the 8-neighbourhood as derscribed in the lecture and $c_{ij} = \lambda \frac{\pi}{8}$ is constant scaled by parameter $\lambda \in \mathbb{R}$.

Extend your program from Exercise 6. The extended exercise should do the following extra things:

- Build a graphical model with OpenGM that represents energy E from Equation (10). Note that you need to transform E into a slightly different form. In OpenGM you would specify the energy in the following form:

  $$E(x) = \sum_{i=1}^{n} f_i(x_i) + \sum_{i=1}^{n} \sum_{j \in \mathcal{N}(i): i<j} c_{ij}(x_i, x_j),$$

  where $x \in \mathbb{B}^n$, $f_i : \mathbb{B} \to \mathbb{R}$ are functions with one argument and $c_{ij} : \mathbb{B} \times \mathbb{B} \to \mathbb{R}$ are functions with two arguments.

- Find the minimizer of E with the help of the inference algorithms build into OpenGM.

- Display the segmentation result.

You can use the file `imagesegmentation.cpp` in `03_supp.zip` as a starting point for your program. `imagesegmentation.cpp` includes example code for everything you need to accomplish the exercise.

Experiment with different values for parameter λ. How does the segmentation change? Additionally you can try the image dependent length term $c_{ij} = \lambda \exp(-\frac{\|I(i)-I(j)\|^2}{2\sigma^2})$, where $I(i), I(j) \in \mathbb{R}^3$ are the RGB values at pixels i, j. Choose the standard deviation $\sigma \in \mathbb{R}$ carefully.

Exercise 8 (5 Points)**.** In Exercise 6 the user-specified rectangle served as a rough initial guess of the foreground region. The idea of this exercise is to use the segmentation result from Exercise 7 as a better guess for the foreground region when computing the Gaussian mixture model.

Write a program which performs the two previous exercises, estimating the Gaussian mixture model and computing the segmentation, in an alternating Expectation Maximization fashion. The E-step corresponds to the estimation of the GMM and the M-step is the segmentation algorithm.

Hint: Keep the means, covariances and responsibilities from the previous outer iteration as an initialization to warm-start the GMM optimization algorithm. Use the command `vl_set_initializion(gmm, VlGMMCustom);` to achieve that.