Image Denoising

Homework n°4 - Exercises

Exercise 7.4

Using Bayes' theorem, we have:

$$\begin{aligned}
\gamma_{kn}^{(t)} &= \mathbb{P}(z_n = k \mid x_n; \theta^{(t)}) \\
&= \frac{\mathbb{P}(z_n = k, x_n; \theta^{(t)})}{\mathbb{P}(x_n; \theta^{(t)})} \\
&= \frac{\mathbb{P}(z_n = k; \theta^{(t)})\mathbb{P}(x_n \mid z_n = k; \theta^{(t)})}{\mathbb{P}(x_n; \theta^{(t)})}.
\end{aligned}$$

Then, we use:

$$\mathbb{P}(z_n = k; \theta^{(t)}) = \pi_k^{(t)},$$

$$\mathbb{P}(x_n \mid z_n = k; \theta^{(t)}) = \mathcal{N}(x_n; \mu_k^{(t)}, \Sigma_k^{(t)}),$$

$$\mathbb{P}(x_n; \theta^{(t)}) = \sum_{k=1}^{K} \mathbb{P}(x_n, z_n = k; \theta^{(t)}) = \sum_{k=1}^{K} \mathbb{P}(z_n = k; \theta^{(t)})\mathbb{P}(x_n \mid z_n = k; \theta^{(t)}) = \sum_{k=1}^{K} \pi_k^{(t)} \mathcal{N}(x_n; \mu_k^{(t)}, \Sigma_k^{(t)})$$

to get

$$\boxed{\gamma_{kn}^{(t)} = \frac{\pi_k^{(t)} \mathcal{N}(x_n; \mu_k^{(t)}, \Sigma_k^{(t)})}{\sum\limits_{k=1}^{K} \pi_k^{(t)} \mathcal{N}(x_n; \mu_k^{(t)}, \Sigma_k^{(t)})}}.$$

Exercise 7.7

The MAP optimization problem to directly estimate θ by maximizing $L(x;\theta)$ is

$$\max_{\theta} \; L(\theta; x) = \prod_{n=1}^{N} \sum_{k=1}^{K} \pi_k \mathcal{N}(x_n; \mu_k, \Sigma_k)$$

or equivalently

$$\max_{\theta} \; \log L(\theta; x) = \sum_{n=1}^{N} \log \left(\sum_{k=1}^{K} \pi_k \mathcal{N}(x_n; \mu_k, \Sigma_k) \right). \tag{1}$$

The optimization problem solved in the M step is

$$\max_{\theta} \; Q(\theta \mid \theta^{(t)}) = \sum_{n=1}^{N} \sum_{k=1}^{K} \gamma_{kn}^{(t)} \left[\log \pi_k - \frac{1}{2} \log |\Sigma_k| - \frac{1}{2}(x_n - \mu_k)^T \Sigma_k^{-1} (x_n - \mu_k) - \frac{d}{2} \log(2\pi) \right]. \tag{2}$$

A difference is that in (1), μ_k and Σ_k appear inside an exponential (because the logarithm is outside the sum on k), so differentiating with respect to these parameters is not convenient. On the other hand, in (2), the different terms essentially appear in quadratic expressions, easy to differentiate (and with differentials easy to set to 0).

Exercise 9.2

$$\mathrm{EPLL}_{\mathbb{P}}(U) = \sum_{P \in \mathcal{P}} \log \mathbb{P}(PU).$$

with $\mathcal{P}$ a set of patch projectors.

We have

$$\mathrm{EPLL}_{\mathbb{P}}(U) = \log \prod_{p \in \mathcal{P}} \mathbb{P}(PU)$$

If we assume that the patches $\mathbf{PU}, \mathbf{P} \in \mathcal{P}$ are independent, then this can be rewritten

$$\mathrm{EPLL}_{\mathbb{P}}(U) = \log \mathbb{P}\left(\bigcap_{P \in \mathcal{P}} \{PU\} \right),$$

in which case $\mathrm{EPLL}_{\mathbb{P}}(\mathbf{U})$ can indeed be interpreted as a log-likelihood.

However, if we do not assume that the patches $PU, P \in \mathcal{P}$ are independent, then in general we cannot see $\mathrm{EPLL}_{\mathbb{P}}(U)$ as a log-likelihood.

And in general, the patches are not independent, since they overlap. This can be approximately true for small patches in a large image, but it remains an approximation.

Exercise 9.3

$$E(Q_P) = \beta \frac{\|PU - Q_P\|^2}{2\sigma^2} - \log \sum_{k=1}^{K} \pi_k \mathcal{N}(Q_P \mid \mu_k, \Sigma_k).$$

We have the following equivalent problems:

$$\begin{aligned}
\min E(Q_P) &\iff \min \ \beta \frac{\|PU - Q_P\|^2}{2\sigma^2} - \log \sum_{k=1}^{K} \pi_k \mathcal{N}(Q_P \mid \mu_k, \Sigma_k) \\
&\iff \max \ \exp\left(\log \sum_{k=1}^{K} \pi_k \mathcal{N}\left(Q_P \mid \mu_k, \Sigma_k\right) - \beta \frac{\|PU - Q_P\|^2}{2\sigma^2} \right) \\
&\iff \max \ \left(\sum_{k=1}^{K} \pi_k \mathcal{N}\left(Q_P \mid \mu_k, \Sigma_k\right) \right) \exp\left(-\beta \frac{\|PU - Q_P\|^2}{2\sigma^2} \right) \\
&\iff \max \ \mathbb{P}(Q_P)\mathbb{P}(PU \mid Q_P),
\end{aligned}$$

if we set $\mathbb{P}(Q_P) = \sum_{k=1}^{K} \pi_k \mathcal{N}\left(Q_P \mid \mu_k, \Sigma_k\right)$ and $\mathbb{P}(PU \mid Q_P) = \frac{1}{\sqrt{(2\pi)^d \sigma^2}} \exp\left(-\beta \frac{\|PU - Q_P\|^2}{2\sigma^2}\right)$ with d the right dimension (that is, given Q_P, PU is centered in Q_P, but with some noise of variance σ^2).

$$\boxed{\text{That is, minimizing } E(Q_P) \text{ is equivalent to a maximum a posteriori estimate of } Q \text{ assuming that it obeys a GMM with weights } \pi_k, \text{ means } \mu_k, \text{ and covariance matrices } \Sigma_k.}$$

Exercise 9.6

$$E_\beta\left(U, \{Q_P\}_{p\in\mathcal{P}} \mid \tilde{U}\right) = \sum_{P\in\mathcal{P}}\left[\frac{\|PU - P\tilde{U}\|^2}{2\sigma^2} + \beta\frac{\|PU - Q_P\|^2}{2\sigma^2} - \log\mathbb{P}(Q_P)\right] \tag{9.6}$$

For fixed Q_P, differentiating (9.6) with respect to U yields

$$\sum_{P\in\mathcal{P}}\left[\frac{2P^T(PU - P\tilde{U})}{2\sigma^2} + \beta\frac{2P^T(PU - Q_P)}{2\sigma^2}\right]$$

which can be rewritten

$$\frac{1}{\sigma^2}\sum_{P\in\mathcal{P}}\left[P^T(PU - P\tilde{U}) + \beta P^T(PU - Q_P)\right]$$

or

$$\frac{1}{\sigma^2}\left((1+\beta)\sum_{P\in\mathcal{P}}P^TP\right)U - \frac{1}{\sigma^2}\sum_{P\in\mathcal{P}}\left(P^TP\tilde{U} + \beta P^TQ_P\right)$$

We set this to 0:

$$\begin{aligned}
U &= \left((1+\beta)\sum_{P\in\mathcal{P}}P^TP\right)^{-1}\sum_{P\in\mathcal{P}}\left(P^TP\tilde{U} + \beta P^TQ_P\right)\\
&= (1+\beta)^{-1}\left(\sum_{P\in\mathcal{P}}P^TP\right)^{-1}\left(\left(\sum_{P\in\mathcal{P}}P^TP\right)\tilde{U} + \beta\sum_{P\in\mathcal{P}}P^TQ_P\right)\\
&= (1+\beta)^{-1}\left(\tilde{U} + \beta\left(\sum_{P\in\mathcal{P}}P^TP\right)^{-1}\sum_{P\in\mathcal{P}}P^TQ_P\right),
\end{aligned}$$

so

$$\boxed{U = \frac{\tilde{U} + \beta\left(\sum_{P\in\mathcal{P}}P^TP\right)^{-1}\sum_{P\in\mathcal{P}}P^TQ_P}{1+\beta}}.$$