5A

Machine Learning II

ID: 5684926 Tristan Scheidemann

5A-1. a

Explicit calculation:

$$
\begin{aligned}
p(\mathbf{x}|\boldsymbol{\mu}_k, \Sigma_k) &= (2\pi)^{-\frac{D}{2}} |\Sigma^{-1}|^{\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^T \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu}_k)} \\
&= (2\pi)^{-\frac{D}{2}} \left|\frac{1}{\epsilon}\mathrm{I}\right|^{\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^T \frac{1}{\epsilon}\mathrm{I} (\mathbf{x}-\boldsymbol{\mu}_k)} \\
&= (2\pi\epsilon)^{-\frac{D}{2}} e^{-\frac{1}{2\epsilon}(\mathbf{x}-\boldsymbol{\mu}_k)^T (\mathbf{x}-\boldsymbol{\mu}_k)}.
\end{aligned}
$$

5A-1. b

Explicit calculation:

$$
\begin{aligned}
p(\mathbf{z}_k = 1|\mathbf{x}_n) &= \frac{p(\mathbf{x}_n|\mathbf{z}_k = 1)\, p(\mathbf{z}_k = 1)}{p(\mathbf{x}_n)} \\
&= \frac{\pi_k (2\pi\epsilon)^{-\frac{D}{2}} e^{-\frac{1}{2\epsilon}(\mathbf{x}_n-\boldsymbol{\mu}_k)^T (\mathbf{x}_n-\boldsymbol{\mu}_k)}}{(2\pi\epsilon)^{-\frac{D}{2}} \sum_{i=1}^{K} \pi_i e^{-\frac{1}{2\epsilon}(\mathbf{x}-\boldsymbol{\mu}_i)^T (\mathbf{x}-\boldsymbol{\mu}_i)}} \\
&= \frac{\pi_k e^{-\frac{1}{2\epsilon}(\mathbf{x}_n-\boldsymbol{\mu}_k)^T (\mathbf{x}_n-\boldsymbol{\mu}_k)}}{\sum_{i=1}^{K} \pi_i e^{-\frac{1}{2\epsilon}(\mathbf{x}-\boldsymbol{\mu}_i)^T (\mathbf{x}-\boldsymbol{\mu}_i)}}.
\end{aligned}
$$

Now we let $\epsilon \to 0$.

Auxiliary calculation

Change of exponential's basis:

$$
\begin{aligned}
\alpha^x &= e^{-\frac{1}{2\alpha}\|\mathbf{x}_n-\boldsymbol{\mu}_i\|^2} \\
\Leftrightarrow \quad x \ln \alpha &= -\frac{1}{2\alpha}\|\mathbf{x}_n - \boldsymbol{\mu}_i\|^2 \\
\Leftrightarrow \quad x &= -\frac{\frac{1}{2\alpha}\|\mathbf{x}_n - \boldsymbol{\mu}_i\|^2}{\ln \alpha}.
\end{aligned}
$$

Furthermore, for any $b \in \mathbb{R}^+/\{0\}$:

$$
\lim_{\substack{\alpha \to 0 \\ \alpha > 0}} \log_\alpha b = 0.
$$

This is easy to see, because for any $\alpha \in [0,1]$

$$0 < \log_\alpha b < 1 \implies \alpha^{-1} > b$$

which converges to zero if $\alpha \to 0$.

Additionally, we have the Maslov dequantization

$$\min_i b_i = \lim_{\substack{\alpha\to 0 \\ \alpha>0}} \log_\alpha \sum_i \alpha^{b_i}.$$

Sometimes we can recover a tropicalization even if the terms are not in the correct form:

$$\log_\alpha \pi_{i^*} + \min_i b_i = \lim_{\substack{\alpha\to 0 \\ \alpha>0}} \log_\alpha \pi_i \alpha^{b_i},$$

where i^* be the minimum term index of $\min_i b_{i^*}$ and π_i is positive.

First, we evaluate the denominator:

$$\begin{aligned}
\lim_{\substack{\alpha\to 0 \\ \alpha>0}} \sum_{i=1}^{K} \pi_i e^{-\frac{1}{2\alpha}(\mathbf{x}-\mu_i)^T(\mathbf{x}-\mu_i)} &= \lim_{\substack{\alpha\to 0 \\ \alpha>0}} \underbrace{\sum_{i=1}^{K} \pi_i \alpha^{-\frac{\frac{1}{2\alpha}\|\mathbf{x}_n-\boldsymbol{\mu}_i\|^2}{\ln\alpha}}}_{change\ of\ basis} \\
&= \lim_{\substack{\alpha\to 0 \\ \alpha>0}} \alpha^{\log_\alpha\left[\sum_{i=1}^{K} \pi_i \alpha^{-\frac{\frac{1}{2\alpha}\|\mathbf{x}_n-\boldsymbol{\mu}_i\|^2}{\ln\alpha}}\right]} \\
&= \lim_{\substack{\alpha\to 0 \\ \alpha>0}} \alpha^{\underbrace{\bigoplus_{i=1}^{K}\left[-\frac{\frac{1}{2\alpha}\|\mathbf{x}_n-\boldsymbol{\mu}_i\|^2}{\ln\alpha}\right]}_{Maslov\ dequantization}+\log_\alpha \pi_{i^*}} \\
&= \lim_{\substack{\alpha\to 0 \\ \alpha>0}} \alpha^{\min_i\left[-\frac{\frac{1}{2\alpha}\|\mathbf{x}_n-\boldsymbol{\mu}_i\|^2}{\ln}\cdot\right]+\log_\alpha \pi_{i^*}} \\
&= \lim_{\substack{\alpha\to 0 \\ \alpha>0}} \alpha^{-\frac{\frac{1}{2\alpha}\min_i\|\mathbf{x}_n-\boldsymbol{\mu}_i\|^2}{\ln}\cdot+\log_\alpha \pi_{i^*}} \\
&= \lim_{\substack{\alpha\to 0 \\ \alpha>0}} \pi_i^* \alpha^{-\frac{\frac{1}{2\alpha}\min_i\|\mathbf{x}_n-\boldsymbol{\mu}_i\|^2}{\ln\alpha}} \\
&= \lim_{\substack{\alpha\to 0 \\ \alpha>0}} \pi_i^* e^{-\frac{1}{2\alpha}\min_i\|\mathbf{x}_n-\boldsymbol{\mu}_i\|^2}.
\end{aligned}$$

In this case, $\bigoplus$ denotes addition in min-algebra. Evaluation of the limit now becomes simple:

$$\begin{aligned} \lim_{\substack{\alpha\to 0\\ \alpha>0}} p(\mathbf{z}_k = 1|\mathbf{x}_n) &= \frac{\lim_{\substack{\alpha\to 0\\ \alpha>0}} \pi_k e^{-\frac{1}{2\alpha}(\mathbf{x}_n-\mathbf{\mu}_k)^T(\mathbf{x}_n-\mathbf{\mu}_k)}}{\lim_{\substack{\alpha\to 0\\ \alpha>0}} \pi_i^* e^{-\frac{1}{2\alpha}\min_i\|\mathbf{x}_n-\boldsymbol{\mu}_i\|^2}}, \\ \Leftrightarrow \qquad r_{nk} &= \begin{cases} 1, if\ \|\mathbf{x}_n - \boldsymbol{\mu}_k\|^2 = \min_i \|\mathbf{x}_n - \boldsymbol{\mu}_i\|^2 \\ 0, if\ \|\mathbf{x}_n - \boldsymbol{\mu}_k\|^2 \neq \min_i \|\mathbf{x}_n - \boldsymbol{\mu}_i\|^2 \end{cases}. \end{aligned}$$

s