Exercise 12 of Machine Learning [IN 2064]

Name: Yiman Li
Matr-Nr: 03724352
cooperate with Kejia Chen(03729686)

Problem 1

To be seen in the end.

Problem 2

According to [1], let the linear encoder be

$$\boldsymbol{h} = f(\boldsymbol{x}) = \boldsymbol{W}^T(\boldsymbol{x} - \boldsymbol{\mu}) \tag{1}$$

The encoder computers a low-dimensional representation of $\boldsymbol{h}$. With the autoencoder view, we have a decoder computing the reconstruction:

$$\hat{\boldsymbol{x}} = g(\boldsymbol{h}) = \boldsymbol{b} + \boldsymbol{V}\boldsymbol{h} \tag{2}$$

The choice of linear encoder and decoder that minimize reconstruction error $\mathbb{E}[||\boldsymbol{x} - \hat{\boldsymbol{x}}||^2]$ correspond to $\boldsymbol{V} = \boldsymbol{W}$, $\boldsymbol{\mu} = \boldsymbol{b} = \mathbb{E}[\boldsymbol{x}]$ and the columns of $\boldsymbol{W}$ form an orthogonal basis which spans the same subspace as the principle eigenvectors of the covariance matrix

$$\boldsymbol{C} = \mathbb{E}[(\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^T] \tag{3}$$

what we've already known is that the eigenvalue λ_i of $\boldsymbol{C}$ corresponds to the variance of $\boldsymbol{x}$ in the direction of eigenvector $\boldsymbol{v}^{(i)}$. If $\boldsymbol{x} \in \mathbb{R}^D$ and $\boldsymbol{h} \in \mathbb{R}^d$ with $K < D$, then the optimal reconstruction error(choosing $\boldsymbol{\mu}, \boldsymbol{b}, \boldsymbol{W}$ and $\boldsymbol{W}$ as above) is

$$\min \mathbb{E}[||\boldsymbol{x} - \hat{\boldsymbol{x}}||^2] = \sum_{i=K+1}^{D} \lambda_i \tag{4}$$

- So only on the condition that the covariance has rank K, the eigenvalues λ_{K+1} to λ_D are 0 and reconstruction error is 0.
- Thus, it is usually impossible to get zero reconstruction error in the setting that $K < D$.

Problem 3

For expected value, using iterated expectations, we have:

$$\begin{aligned} \mathbb{E}[\boldsymbol{x}] &= \mathbb{E}_{p(z)}[\mathbb{E}_{p(\boldsymbol{x}|z)}[\boldsymbol{x}|z]] \\ &= \sum_{k=1}^{K} \boldsymbol{\pi}_k \mathbb{E}_{p(\boldsymbol{x}|z)}[\boldsymbol{x}|z] \\ &= \sum_{k=1}^{K} \boldsymbol{\pi}_k \boldsymbol{\mu}_k \end{aligned} \tag{5}$$

For covariance, since

$$\begin{aligned}
\mathbb{E}[\boldsymbol{x}\boldsymbol{x}^T] &= \mathbb{E}_{p(z)}[\mathbb{E}_{p(\boldsymbol{x}\boldsymbol{x}^T|z)}[\boldsymbol{x}|z]] \\
&= \sum_{k=1}^{K} \boldsymbol{\pi}_k \mathbb{E}_{p(\boldsymbol{x}|z)}[\boldsymbol{x}\boldsymbol{x}^T|z] \\
&= \sum_{k=1}^{K} \boldsymbol{\pi}_k (\sigma_k + \boldsymbol{\mu}_k \boldsymbol{\mu}_k^T)
\end{aligned} \tag{6}$$

So we have

$$\begin{aligned}
Cov[\boldsymbol{x}] &= \mathbb{E}[\boldsymbol{x}\boldsymbol{x}^T] - \mathbb{E}[\boldsymbol{x}]\mathbb{E}[\boldsymbol{x}]^T \\
&= \sum_{k=1}^{K} \boldsymbol{\pi}_k (\sigma_k + \boldsymbol{\mu}_k \boldsymbol{\mu}_k^T) - \sum_{i=1}^{K}\sum_{j=1}^{K} \boldsymbol{\pi}_i \boldsymbol{\pi}_j \boldsymbol{\mu}_i \boldsymbol{\mu}_j
\end{aligned} \tag{7}$$

Problem 4

a) The generative process is showed as below:

- Draw k from the categorical distribution on $1, \cdots, K_x$ with probability from $\boldsymbol{\pi}^x$
- Draw $\tilde{\boldsymbol{x}}$ from the normal distribution $\mathcal{N}(\boldsymbol{\mu}_k^x, \boldsymbol{\Sigma}_k^x)$
- Draw l from the categorical distribution on $1, \cdots, K_y$ with probability from $\boldsymbol{\pi}^y$
- Draw $\tilde{\boldsymbol{y}}$ from the normal distribution $\mathcal{N}(\boldsymbol{\mu}_l^y, \boldsymbol{\Sigma}_l^y)$
- Return $\tilde{\boldsymbol{z}} := \tilde{\boldsymbol{x}} + \tilde{\boldsymbol{y}}$

b) The sum of a two Gaussian distribution is still a Gasussian distribution

c) The distribution of $\boldsymbol{z}$ is showed below:

$$p(\boldsymbol{z}|\boldsymbol{\theta}^x, \boldsymbol{\theta}^y) = \sum_{k=l}^{K_x}\sum_{l=1}^{K_y} \pi_k^x \pi_l^y \mathcal{N}(\boldsymbol{z}|\boldsymbol{\mu}_k^x + \boldsymbol{\mu}_l^y, \boldsymbol{\Sigma}_k^x + \boldsymbol{\Sigma}_l^y) \tag{8}$$

Problem 5

To be seen in the end.

References

[1] Ian J. Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, Cambridge, MA, USA, 2016. `http://www.deeplearningbook.org`.