Quiz 3 - Lecture 14 (Prof. Shinoda)

1. Prove that $p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$
2. Discuss the future prospect of deep learning and its related techniques.

Collaborators: None.

Exercise 3-1. Prove that $p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$

Solution: By definition, $\mathbf{z}$ is one-hot encoding representation, we have:

$$p(\mathbf{z}) = \prod_{k=1}^{K} \pi_k^{z_k}$$
$$p(\mathbf{x}|\mathbf{z}) = \prod_{k=1}^{K} \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)^{z_k}$$

By the product rule, we have the join probability of $\mathbf{x}$ and $\mathbf{z}$ as follow:

$$p(\mathbf{x}, \mathbf{z}) = p(\mathbf{x}|\mathbf{z})p(\mathbf{z})$$

Using the sum product to compute the marginal $p(\mathbf{x})$:

$$\begin{aligned} p(\mathbf{x}) &= \sum_{\mathbf{z}} p(\mathbf{x}|\mathbf{z})p(\mathbf{z}) \\ &= \sum_{\mathbf{z}} \prod_{k=1}^{K} \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)^{z_k} \pi_k^{z_k} \quad = \sum_{j=1}^{K} \prod_{k=1}^{K} (\pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)^{\delta_{jk}}, \end{aligned}$$

where δ_{jk} is the Kronecker delta. Simply rewrite the product keeping not-1 values, we have the desired result:

$$p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$

Exercise 3-2. Discuss the future prospect of deep learning and related techniques.

Solution: Recently, machine learning and especially deep learning technique have been employed into everyday life. Deep learning technique originates from the effort to model human's neural network. However, our best current models only can mimic a small fraction of biological brain work functions. The challenges include: finding an effective activation function for neurons, automate training data acquisition, and multi-task machine. Currently, our best models still use very simple activation function artificial neuron to keep the back-propagation computation cost tractable.