Hoang Nguyen - 15M54097

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. Therefore, the one of the challenges for future deep learing and cognitive science researchers is to find the "holy grail" activation function. On another matter, current models cannot acquire knowledge by themselves, but they must rely on man-made inputs and data structure. In the world, few Machine Learning groups are working on the problem of choosing training samples that focus on diversity by using submodularity set functions. Another demand for machine learning system is the capacity for performing multiple task and automatic logical reasoning. Finally, training a deep learning system requires high performance computing as well as long training time. The training time of deep learning system doesn't scale with the data produced everyday. As a result, many machine-based system still employ traditional learning technique such as decision tree, random forest, or SVM. One possible solution for this problem in near future is to employ quantum computing in the deep network training process. In 2016, D-Waves has introduced D2X with 1000 q-bits. D2X has produced result for graph coloring problem with more than 600 vertices in an instance. Furthermore, Google and D-Waves have been working together for quantum machine learning project. Their intial results on car recognition for self-driving car outperforms current graphic card based optimization. I believe deep learning techniques will soon leave the research labs and become the most essential part of mankind.