Deep Learning

AlejandroMllo

This document serves as a very brief summary of the topics covered in each chapter of the book Deep Learning [1].

Disclaimer: This document is completely extracted from [1], the author does not attribute any ownership over the material.

13 Linear Factor Models

- Models that use probabilistic inference, many of which use latent variables $\boldsymbol{h}$, with $p_{model}(\boldsymbol{x}) = \mathbb{E}_{\boldsymbol{h}} p_{model}(\boldsymbol{x}|\boldsymbol{h})$.
- They are defined by the use of a stochastic linear decoder function that generates $\boldsymbol{x}$ by adding noise to a linear transformation of $\boldsymbol{h}$.
- Its data-generation process is described as follows:
  - The explanatory factors $\boldsymbol{h}$ are sampled from a distribution $\mathbf{h} \sim p(\boldsymbol{h})$, where $p(\boldsymbol{h})$ is a factorial distribution, with $p(\boldsymbol{h}) = \prod_i p(h_i)$.
  - Next, the real-valued observable variables are sampled given the factors $\boldsymbol{x} = \boldsymbol{W}\boldsymbol{h} + \boldsymbol{b} +$ noise.
- Probabilistic PCA, factor analysis and other linear factor models are special cases of the above equations.
- Independent Component Analysis (ICA) is an approach to modeling linear factors that seeks to separate an observed signal into many underlying signals that are scaled and added together to form the observed data. These signals are intended to be fully independent.
  - This is commonly used to recover low-level signals that have been mixed together.
- Slow Feature Analysis (SFA) is a linear factor model that uses information from time signals to learn invariant features. The underlaying idea is that the important characteristics of scenes change very slowly compared to the individual measurements that make up a description of a scene.
  - It is possible to theoretically predict which features SFA will learn.
- Sparse coding is an unsupervised feature learning and feature extraction mechanism. Strictly speaking, "sparse coding" refers to the process of inferring the value of $\boldsymbol{h}$ in a model, while "sparse modeling" refers to the process of designing and learning the model, but the term "sparse coding" is often used to refer to both.
- Linear factor models including PCA and factor analysis can be interpreted as learning a manifold.

References

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