Chapter I
Complex–Valued Boltzmann Manifold

Masaki Kobayashi
University of Yamanashi, Japan

ABSTRACT

Information geometry is one of the most effective tools to investigate stochastic learning models. In it, stochastic learning models are regarded as manifolds in the view of differential geometry. Amari applied it to Boltzmann Machines, which is one of the stochastic learning models. The purpose of this chapter is to apply information geometry to complex-valued Boltzmann Machines. First, we construct the complex-valued Boltzmann Machines. Next, the author describes information geometry. The author will know some important notions of information geometry, exponential families, mixture families, Kullback-Leibler divergence, connections, geodesics, Fisher metrics, potential functions and so on. Finally, they apply information geometry to complex-valued Boltzmann Machines. They will investigate the structure of complex-valued Boltzmann manifold and know the notions of the connections and Fisher metric. Moreover we will get an effective learning algorithm, what is called em algorithm, for complex-valued Boltzmann machines with hidden neurons.

INTRODUCTION

These days we can get massive information and it is hard to deal with it without computers. Machine learning is effective for computers to manage massive information. Machine learning uses various learning machine models, for instance, decision trees, Bayesian Networks, Support Vector Machine, Hidden Markov Model, normal mixed distributions, neural networks and so on. Some of them are stochastically constructed.

The neural network is one of the learning machine models. It consists of many units, which are called neurons. We often use binary neurons. Then each neuron takes only two states. The set of neurons, however, takes many states. Various types of neural networks have been proposed. Feed-forward types and symmetric types of neural networks are main models. Feed-forward types of neural networks are often applied to recognize given patterns and are so useful. Symmetric types of neural network are often applied as Associative Memories. The Hopfield Network is one of the most famous models. Boltzmann Machines are stochastic types of Hopfield Networks.

Neurons take only two states in most cases. McEliece, Posner, Rodemich and Venkatech (1987) is a highly recognized critique of the Hopfield memory low capacity. Baldi and Hormik (1989) rigorously showed existence of numerous local minima in objective function for learning a nonlinear perceptron. The presentation capacity of a neuron is too poor. We hope that the neuron models have multi-states. Some researchers have proposed such models. Multi-level neuron is one of them (Zurada, Cloete & Poel, 1996). **Complex-valued neurons** are also multi-states neuron models. Several models of complex-valued neurons have been proposed, for example, phasor neurons (Noest, 1988a), discrete-state phasor neurons (Noest, 1988b), amplitude-phase type of complex-valued neurons (Hirose, 1992; Kuroe, 2003), real part – imaginary part type of complex-valued neurons (Benvenuto & Piazza, 1992; Nitta & Furuta, 1991; Nitta, 1997) and so on (Nemoto & Kubono, 1996; Nemoto, 2003). In this chapter, we deal with the phasor neurons and the discrete phasor neurons. In this chapter, we call phasor neurons continuous phasor neurons to distinguish them clearly from discrete phasor neurons.

Several types of neural networks, feed-forward neural networks, Hopfield Networks and Boltzmann Machines, have been extended to the complex-valued neural networks. Noest proposed complex-valued Hopfield Networks (Noest, 1988a; Noest, 1988b). They are called continuous phasor or discrete phasor neural networks. Hirose proposed back propagation learning algorithms for the amplitude-phase type of complex-valued neural networks (Hirose, 1992). Benvenuto and Piazza (1992) and Nitta and Furuya (1991) independently proposed back propagation learning algorithms for the real part – imaginary part type of complex-valued neural networks. Boltzmann Machines were also extended to the **complex-valued Boltzmann Machines** (Zemel, Williams & Mozer, 1993; Zemel, Williams & Mozer, 1995). The complex-valued Boltzmann Machines proposed by Zemel et al. (1993) are continuous models. Kobayashi and Yamazaki proposed the discrete version of complex-valued Boltzmann Machines (Kobayashi & Yamazaki, 2003).

We have to know learning machines' behaviors to manage them. But the learning machines often behave strangely. For example local minima and plateaus appear. It is called to be caused by the singularities. The learning machines without singularities are called regular models. Regular machines behaviors are so simple, so we can expect it well. For example, AIC (Akaike Information Criterion) is effective for regular models. But we cannot deal with the singular ones by the similar way. The local minima and plateaus in the learning process appear in the feed-forward neural networks with hidden neurons. But they do not appear in a simple perceptron (Fukumizu & Amari, 2000). We need the effective tools to investigate the singularities. **Information geometry** is one of them.

Boltzmann Machines without hidden neurons are regular models. Boltzmann Machines with hidden neurons are singular ones. Complex-valued Boltzmann Machines are also singular. Information geometry is one of the good tools to investigate singularities. In this chapter, we investigate the complex-valued Boltzmann Machines in the view of information geometry. Information geometry provides an efficient learning method, *em* **algorithm** to complex-valued Boltzmann Machines.

BACKGROUND

Information geometry, which is a special type of differential geometry, is mainly constructed by Amari. It is an effective tool to investigate stochastic models. In fact, Amari applied it to the various stochastic models, normal distributions, discrete distributions, stochastic multilayer perceptron, normal mixture with hidden variables, normal multiplication model and Boltzmann Machines (Amari, Kurata & Nagaoka, 1992a; Amari, 1995). In information geometry, a family of probability distributions is considered as a manifold, what is called a stochastic manifold.

Information geometry provides the notions of distance between two given probability distributions to the family of probability distributions. And it also provides the notions of parallel movement on the family of probability distributions.

Related Content

Application and Some Fundamental Study of GNN In Forecasting

Arun Kumar Garov, A. K. Awasthi, Ram Kumarand Monica Sankat (2023). *Concepts and Techniques of Graph Neural Networks (pp. 132-154).*

www.irma-international.org/chapter/application-and-some-fundamental-study-of-gnn-in-forecasting/323826

Artificial Intelligence and Reliability Metrics in Medical Image Analysis

Yamini G.and Gopinath Ganapathy (2020). *Deep Neural Networks for Multimodal Imaging and Biomedical Applications (pp. 172-185).*

www.irma-international.org/chapter/artificial-intelligence-and-reliability-metrics-in-medical-image-analysis/259493

High Speed Optical Higher Order Neural Networks for Discovering Data Trends and Patterns in Very Large Databases

David R. Selviah (2009). *Artificial Higher Order Neural Networks for Economics and Business (pp. 442-465).*

www.irma-international.org/chapter/high-speed-optical-higher-order/5294

Introduction to Graph Neural Network: Types and Applications

Ganga Devi S. V. S. (2023). *Concepts and Techniques of Graph Neural Networks (pp. 33-42).*

www.irma-international.org/chapter/introduction-to-graph-neural-network/323820

Role of Artificial Neural Network for Prediction of Gait Parameters and Patterns

Kamalpreet Sandhuand Vikram Kumar Kamboj (2020). *AI Techniques for Reliability Prediction for Electronic Components (pp. 124-135).*

www.irma-international.org/chapter/role-of-artificial-neural-network-for-prediction-of-gait-parameters-and-patterns/240494