Probability Models for Complex Systems

by
Zhiyi Chi

Sc.M. in Mathematics, Brown University, Providence, Rhode Island, 1994

Thesis

Submitted in partial fulfillment of the requirements for
the Degree of Doctor of Philosophy
in Division of Applied Mathematics at Brown University

Brown University
Providence, Rhode Island
May 1998

Abstract of "Probability Models for Complex Systems," by Zhiyi Chi, Ph.D., Brown University, May 1998

This thesis is a collection of essays on probability models for complex systems.

Chapter 1 is an introduction to the thesis. The main point made here is the importance of probabilistic modeling to complex problems of machine perception.

Chapter 2 studies minimum complexity regression. The results include: (1) weak consistency of the regression, (2) divergence of estimates in L^2-norm with an arbitrary complexity assignment, and (3) condition on complexity measure to ensure strong consistency.

Chapter 3 proposes compositionality as a general principle for probabilistic modeling. The main issues covered here are: (1) existence of general compositional probability measures, (2) subsystems of compositional systems, and (3) Gibbs representation of compositional probabilities.

Chapter 4 and 5 establish some useful properties of probabilistic context-free grammars (PCFGs). The following problems are discussed: (1) consistency of estimated PCFGs, (2) finiteness of entropy, momentum, etc, of estimated PCFGs, (3) branching rates and re-normalization of inconsistent PCFGs, and (4) identifiability of parameters of PCFGs.

Chapter 6 proposes a probabilistic feature based model for languages. Issues dealt with in the chapter include: (1) formulation of such grammars using maximum entropy principle, (2) modified maximum-likelihood type scheme for parameter estimation, (3) a novel pseudo-likelihood type estimation which is more efficient for sentence analysis.

Chapter 7 develops a novel model on the origin of scale invariance of natural images. After presenting the evidence of scale invariance, the chapter goes on to: (1) argue for a $1/r^3$ law of size of object, (2) establish a 2D Poisson model on the origin of scale invariance, and (3) show numerical simulation results for this model.

Chapter 8 is a theoretical extension of Chapter 7. A general approach to construct scale and translation invariant distributions using wavelet expansion is formulated and applied to construct scale and translation invariant distributions on the spaces of generalized functions and functions defined on the whole integer lattice.

Preface

Probabilistic modeling, often called statistical modeling, is becoming increasingly important to the study of many areas of science. The reason for this is twofold. On the one hand, many problems in modern science and technology are so complicated that they cannot be solved accurately by using simple and deterministic rules. However, by introducing stochastic mechanism into the solution, it is possible to find good approximate answers to these problems. For instance, stochastic annealing processes have been used to attack a wide range of hard optimization problems. The performance of the stochastic approach depends largely on how well it incorporates the stochastic mechanism with the elements of the problems. On the other hand, many natural and social phenomena are characterized by a variety of randomness. As an example, in medicine, people observe that the number of the cases of a disease often varies from region to region and from time to time. Usually statistical methods are the main tools to study such phenomena and the effectiveness of these methods relies on how well they model the phenomena and their randomness. It is fair to say that probabilistic modeling is of fundamental importance to the implementation of statistical methods.

This thesis is a collection of essays which have a common theme: the study of complex systems by probabilistic modeling. Under this theme, the essays cover a range of problems which can be roughly divided into five categories: (1) statistical estimation, (2) methodology of probabilistic modeling, (3) probabilistic language model, (4) probabilistic vision model, and finally, (5) probability theory.

Chapter 1 is an introduction to the thesis. From the principle of Grenander's pattern theory, we give further arguments for the importance of probabilistic modeling to vision, speech recognition, or all of machine perception. We also point out the contribution of the results in this thesis to probabilistic modeling.

Chapter 2 is concerned with non-parametric estimation, which is a classical problem in statistics. We study regression based on the minimum complexity principle and establish several consistency results on this estimation method. The results demonstrate that in order to get strong consistency, complexity measures of functions should be tied with the actual behaviors of functions.

Chapter 3 proposes a general theory and methodology, the compositionality principle, for probabilistic modeling of patterns. We introduce the notion of compositionality and formulate composition systems mathematically. The main theoretical result in this chapter is the existence of general probabilistic composition systems. We also introduce the notion of subsystems and represent the compositional probability distributions in the form of Gibbs distribution.

Chapters 4, 5, and 6 study probability models for languages. The first two chapters are devoted to probabilistic context-free grammars, which are among the simplest grammars for languages. In Chapter 4, we demonstrate that estimated production probabilities of a probabilistic context-free grammar always impose a proper distribution on the set of finite parse trees. In Chapter 5, we generalize the results in last chapter and develop an array of other useful statistical results on probabilistic context-free grammars.

As is well known in linguistics, context-free grammars are too weak to well approximate natural languages. In Chapter 6, we introduce a much more general probabilistic language model called probabilistic feature based grammar, which incorporates the theory of unification grammars and the theory of Gibbs distributions. We introduce a pseudo-likelihood type scheme for parameter estimation, which is efficient for language analysis. We also study the more classical maximum-likelihood type estimation scheme and prove the consistency of both schemes.

Chapter 7 applies probabilistic modeling to another complex system — the space of natural images. As is widely believed, statistics of natural images are of fundamental importance to vision as well as image processing. One of the most distinguishing and intriguing statistical properties of natural images is scale invariance of many marginal distributions of images. We establish a model on the origin of scale invariance of natural images. Briefly speaking, the model is a combination of the Poisson point process and projective geometry. We also conduct numerical simulations for the model, and the results show satisfactory scale invariance.

Chapter 8 is an extensive theoretical study of scale invariance. Motivated by the model established in Chapter 7, we develop a general mathematical approach to construct scale and translation invariant distributions on the space of functions defined on the whole integer lattice as well as on the space of generalized functions.

Contents

List of Figures