Uniwersytet Wrocławski
Wydział Matematyki i Informatyki
Instytut Matematyczny
specjalność: Analiza danych

Bartosz Chmiela

Locally-informed proposals in Metropolis-Hastings algorithm with applications

Praca magisterska
napisana pod kierunkiem
dr hab. Pawła Lorka

Wrocław 2022

Abstract

The Markov Chain Monte Carlo methods (abbrv. MCMC) are a family of algorithms used for sampling from a given probability distribution. They prove very effective when the state space is large. This fact can be used to solve many hard deterministic problems – one of them being *traveling salesmen problem*. It will be used in this thesis to test a new approach of *locally-informed propolsals* as a modification of well known *Metropolis-Hastings* algorithm. In this thesis we will present the implementation of modified algorithm, experiments based on it, results and a comparison of to previous MCMC methods.

Metody próbkowania Monte Carlo łańcuchami Markowa są rodziną algorytmow używanych do próbkowania z danego rozkładu prawdopodobieństwa. Okazują się efektywne zwłaszcza gdy przestrzeń stańw jest wielka. Ten fakt może być wykorzystany przy rozwiązywaniu wielu deterministycznych problemów – jednym z nich jest *problem komiwojażera*. Zostanie on użyty w tej pracy do przetestowania nowego podejścia *lokalnie poinformowawnego?*, jako modyfikacji dobrze znanego algorytmu *Metropolisa-Hastinigsa*. W tej pracy zaprezentujemy implementacje zmodyfikowanego algorytmu, eksperymentów bezujących na nim, wyników oraz porównania z poprzednimi metodami próbkowania Monte Carlo.

Contents

List of Tables

List of Figures

1 Introduction

The Markov Chain Monte Carlo methods (abbrv. MCMC) are a family of algorithms used for sampling from a given probability distribution. At first they do not seem useful for solving practical deterministic problems, but with some tweaks they can become a powerful tool. It happens especially when space of possible solutions is enormous and computing becomes infeasible for machines. These offer a shortcut for obtaining "close enough" answers.

At their core, MCMC methods generate a Markov Chain (abbrv. MC) with a defined distribution and sample using it. The convergence of the chain is assured by ergodic theorems. The most known of them is *Metropolis-Hastings* algorithm, which constructs a MC using another set of distributions, maybe simpler ones.

In this thesis we work on *locally-informed proposals*, which involve determining *local* distribution – which comes down to finding transition probabilities of the state. They are a bit more complex and computationally heavy, but offer better results with less iterations.

To test this method we will need a deterministic problem which quickly becomes infeasible for machines to compute – one of them is a well-known traveling salesman problem. The testing is carried out using its benchmark training set *tsplib95* and implementation is provided in *Python3*.

2 Markov chains

Markov chains are the very basic building blocks of the theory used within this thesis. They are a natural extension of independent stochastic processes, that assume a weak dependence between the presence and the past.

In this thesis we will focus only on stochastic processes with discrete time steps and finite state space, which fulfill(?) the Markov property. These are the ones that, we are able to simulate in computers.

2.1 Basic terminology and assumptions

We assume that the reader has a basic probabilistic background, so that we can freely use terminology from probability theory, like random or independent variables, stochastic processes, measure or σ-algebra.

Throughout the whole thesis we will be working on probabilistic space $\{\Omega, \mathcal{F}, P\}$:

Definition 2.1. *A probability space is a triplet: $\{\Omega, \mathcal{F}, P\}$, where Ω is some abstract sample space*

is the space of all possible Markov chains, $\mathcal{F}$ a σ-algebra on Ω and P is a probabilistic measure on this space.

Definition 2.2. *The state space of a Markov chain is a finite set S*

Definition 2.3. *and because time is increasing in discrete steps, we will be working set of indices I instead. In reality we have only finite amount of memory, so set I will be also finite.*

Most of the time a Markov chain will be associated with a transition matrix $\mathbf{P}$:

Definition 2.4. *A transition matrix* $\mathbf{P}$ *is a stochastic matrix, which means that the sum of rows is equal to* 1.

2.2 Definition and basic properties

Formally a Markov chain is a stochastic process

2.3 Stationarity

2.4 Ergodicity

3 Markov chain Monte Carlo methods

4 Traveling salesman problem

5 Decoding encrypted text

6 Code description?

7 Conclusions

References

[1] O. Häggström et al. *Finite Markov chains and algorithmic applications*, volume 52. Cambridge University Press, 2002.

[2] C. Karpiński. On use of monte carlo markov chains to decode encrypted text and to solve travelling salesman problem. 2020.

[3] C. J. Maddison, D. Duvenaud, K. J. Swersky, M. Hashemi, and W. Grathwohl. Oops i took a gradient: Scalable sampling for discrete distributions. 2021.

A Source code