

The flip-the-state transition operator for restricted Boltzmann machines

Kai Brügge · Asja Fischer · Christian Igel

Received: 14 February 2013 / Accepted: 3 June 2013 / Published online: 3 July 2013
© The Author(s) 2013

Abstract Most learning and sampling algorithms for restricted Boltzmann machines (RBMs) rely on Markov chain Monte Carlo (MCMC) methods using Gibbs sampling. The most prominent examples are Contrastive Divergence learning (CD) and its variants as well as Parallel Tempering (PT). The performance of these methods strongly depends on the mixing properties of the Gibbs chain. We propose a Metropolis-type MCMC algorithm relying on a transition operator maximizing the probability of state changes. It is shown that the operator induces an irreducible, aperiodic, and hence properly converging Markov chain, also for the typically used periodic update schemes. The transition operator can replace Gibbs sampling in RBM learning algorithms without producing computational overhead. It is shown empirically that this leads to faster mixing and in turn to more accurate learning.

Keywords Restricted Boltzmann machine · Markov chain Monte Carlo · Gibbs sampling · Mixing rate · Contrastive divergence learning · Parallel tempering

1 Introduction

Restricted Boltzmann machines (RBMs, Smolensky 1986; Hinton 2002) are undirected graphical models describing stochastic neural networks. They have raised much attention

Editors: Hendrik Blockeel, Kristian Kersting, Siegfried Nijssen, and Filip Železný.

K. Brügge

Department of Computer Science, University of Helsinki, P.O. Box 68, 00014 Helsinki, Finland
e-mail: Kai.Bruegge@cs.helsinki.fi

K. Brügge

Helsinki Institute for Information Technology HIIT, P.O. Box 68, 00014 Helsinki, Finland

A. Fischer

Institut für Neuroinformatik, Ruhr-Universität Bochum, 44780 Bochum, Germany
e-mail: asja.fischer@ini.rub.de

C. Igel (✉)

Department of Computer Science, University of Copenhagen, 2100 Copenhagen, Denmark
e-mail: igel@diku.dk