Neural Networks 34 (2012) 1–9

Contents lists available at SciVerse ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

On the equivalence of Hopfield networks and Boltzmann Machines

Adriano Barra [a,*], Alberto Bernacchia [b], Enrica Santucci [c], Pierluigi Contucci [d]

[a] Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185, Roma, Italy
[b] Department of Neurobiology, Yale University, 333 Cedar street, CT06510, New Haven, USA
[c] Dipartimento di Matematica, Università degli Studi dell'Aquila, Via Vetoio 1, 67010, L'Aquila, Italy
[d] Dipartimento di Matematica, Alma Mater Studiorum Università di Bologna, Piazza di Porta San Donato 5, 40126, Bologna, Italy

ARTICLE INFO

Article history:
Received 28 May 2011
Received in revised form 5 January 2012
Accepted 12 June 2012

Keywords:
Statistical mechanics
Hopfield networks
Boltzmann Machines

ABSTRACT

A specific type of neural networks, the Restricted Boltzmann Machines (RBM), are implemented for classification and feature detection in machine learning. They are characterized by separate layers of visible and hidden units, which are able to learn efficiently a generative model of the observed data. We study a "hybrid" version of RBMs, in which hidden units are analog and visible units are binary, and we show that thermodynamics of visible units are equivalent to those of a Hopfield network, in which the N visible units are the neurons and the P hidden units are the learned patterns. We apply the method of stochastic stability to derive the thermodynamics of the model, by considering a formal extension of this technique to the case of multiple sets of stored patterns, which may act as a benchmark for the study of correlated sets.

Our results imply that simulating the dynamics of a Hopfield network, requiring the update of N neurons and the storage of $N(N-1)/2$ synapses, can be accomplished by a hybrid Boltzmann Machine, requiring the update of $N+P$ neurons but the storage of only NP synapses. In addition, the well known glass transition of the Hopfield network has a counterpart in the Boltzmann Machine: it corresponds to an optimum criterion for selecting the relative sizes of the hidden and visible layers, resolving the trade-off between flexibility and generality of the model. The low storage phase of the Hopfield model corresponds to few hidden units and hence a overly constrained RBM, while the spin-glass phase (too many hidden units) corresponds to unconstrained RBM prone to overfitting of the observed data.

1. Introduction

A common goal in Machine Learning is to design a device able to reproduce a given system, namely to estimate the probability distribution of its possible states (Honavar & Uhr, 1994). When a satisfactory model of the system is not available, and its underlying principles are not known, this goal can be achieved by the observation of a large number of samples (Coolen, Kuehn, & Sollich, 2005). A well studied example is the visual world, the problem of estimating the probability of all possible visual stimuli (Pitkow, 2010). A fundamental ability for the survival of living organisms is to predict which stimuli will be encountered and which are more or less likely to occur. For this purpose, the brain is believed to develop an internal model of the visual world, to estimate the probability and respond to the occurrence of various events (Bernacchia & Amit, 2007; Bernacchia, Seo, Lee, & Wang, 2011).

Ising-type neural networks have been widely used as generative models of simple systems (Barra, 2008; Hertz, Krogh, & Palmer, 1991). Those models update the synaptic weights between neurons according to a specific learning rule, depending on the neural activity driven by a given set of observations; after learning, the network is able to generate a sequence of states whose probabilities match those of the observations. Popular examples of Ising models, characterized by a quadratic energy function and a Boltzmann distribution of states, are the Hopfield model (Amit, 1992; Hopfield, 1982) and Boltzmann Machines (BM) (Hinton, 2007). Boltzmann Machines (BM) have been designed to capture the complex statistics of arbitrary systems by dividing neurons in two subsets, visible and hidden units: marginalizing the Boltzmann distribution over the hidden units allows the BM to reproduce, through the visible units, arbitrarily complex distributions of states, by learning the appropriate synaptic weights (Hinton, 2007). State-of-the-art feature detectors and classifiers implement a specific type of BM, the Restricted Boltzmann Machine (RBM), because of its efficient learning algorithms (Bengio, 2009). The RBM is characterized by a bipartite topology in which hidden and visible units are coupled, but there is no interaction within either set of visible or hidden units (Hinton & Salakhutdinov, 2006).

All neurons of RBMs are binary, both the visible and the hidden units. The analog equivalent of RBMs, the Restricted Diffusion

* Corresponding author.
E-mail address: adriano.barra@roma1.infn.it (A. Barra).

0893-6080/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2012.06.003

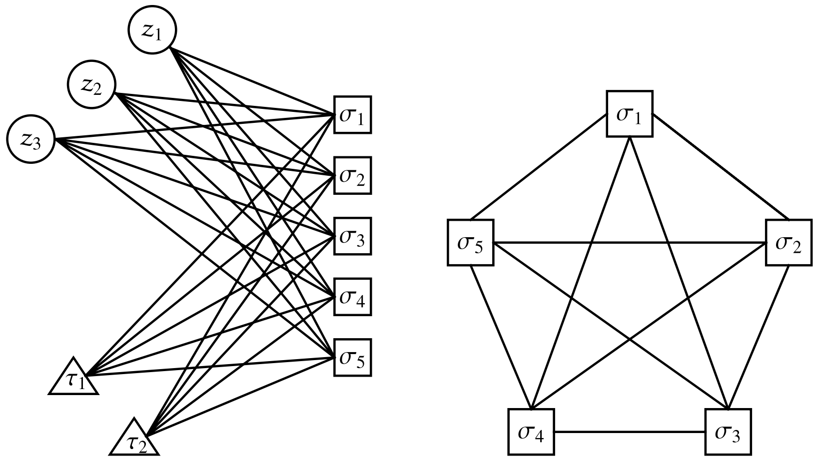

Fig. 1. Left panel: schematic representation of a Hybrid Boltzmann Machine (HBM) where the hidden units are analog (z, τ variables) and the visible units are binary (σ variables). The two sets of hidden units, z and τ, represent two feature sets that are both connected to the layer of visible units σ. The layers of hidden and visible units are reciprocally connected, but there are no intra-layer connections, thus forming a bipartite topology. Right panel: schematic representation of the equivalent Hopfield neural network built upon the visible units only, with an internal fully connected structure.

Networks, have all analog units and have been described in Bengio (2009) and Marks and Movellan (2001). Here we study the case of a "hybrid" Boltzmann Machine (HBM), in which the hidden units are analog and the visible units are binary (Fig. 1 left). We show that the HBM, when marginalized over the hidden units, is equivalent to a Hopfield network (Fig. 1 right), where the N visible units are the neurons and the P hidden units are the learned patterns. Although the Hopfield network can generate probability distributions in a limited space, it has been widely studied for its associative and retrieval properties. The exact mapping proven here introduces a new way to simulate Hopfield networks, and allows a novel interpretation of the spin glass transition, which translates into an optimal criterion for selecting the relative size of the hidden and visible layers in the HBM.

We use the method of stochastic stability to study the thermodynamics of the system in the case of analog synapses. This method has been previously described in Aizenman and Contucci (1998) and Barra, Genovese, and Guerra (2010), and offers an alternative approach to the replica trick for studying Ising-type neural networks, including the Hopfield model and the HBM. We analyze the model with two non-interacting sets of hidden units in the HBM, which correspond to two sets of uncorrelated patterns in the Hopfield network, and study the thermodynamics with the assumption of replica symmetry. We extend the theory to cope with two sets of interconnected hidden layers, corresponding to sets of correlated patterns, and we show that their interaction acts as a noise source for retrieval.

2. Statistical equivalence of HBM and Hopfield networks

We define a "hybrid" Boltzmann Machine (HBM, see Fig. 1 left) as a network in which the activity of units in the visible layer is discrete, $\sigma_i = \pm 1, i \in (1, \dots, N)$ (digital layer), and the activity in the hidden layer is continuous (analog layer). The layers of hidden and visible units are reciprocally connected, but there are no intra-layer connections, thus forming a bipartite topology. We assume that the layer of hidden units is further divided into two sets, both described by continuous variables, $z_\mu, \tau_\nu \in \Re,\ \mu \in (1, \dots, P),\ \nu \in (1, \dots, K)$. We will consider the case of interacting hidden units (connections between z and τ) in the next section. In order to maintain a parsimonious notation, in this section we consider a single hidden layer, e.g. only the layer defined by the variables z.

The synaptic connections between units in the two layers are fixed and symmetric, and are defined by the synaptic matrix ξ_i^μ. The input to unit σ_i in the visible (digital) layer is the sum of the activities in the hidden (analog) layer weighted by the synaptic matrix, i.e. $\sum_\mu \xi_i^\mu z_\mu$. The input to unit z_μ in the hidden (analog) layer is the sum of the activities in the visible (digital) layer, weighted by the synaptic matrix, i.e. $\sum_i \xi_i^\mu \sigma_i$. In the following, we denote by z the set of all hidden $\{z_\mu\}$ variables, and by σ the set of all visible $\{\sigma_i\}$ variables.

The dynamics of the activity is different in the two layers; in the analog layer it changes continuously in time, while in the digital layer it changes in discrete steps. The activity in the hidden (analog) layer follows the stochastic differential equation

$$T\frac{dz_\mu}{dt} = -z_\mu(t) + \sum_i \xi_i^\mu \sigma_i + \sqrt{\frac{2T}{\beta}}\,\zeta_\mu(t), \tag{1}$$

where ζ is a white Gaussian noise with zero mean and covariance $\langle \zeta_\mu(t)\zeta_\nu(t')\rangle = \delta_{\mu\nu}\ \ \delta(t-t')$. The parameter T quantifies the timescale of the dynamics, and the parameter β determines the strength of the fluctuations. The first term in the right hand side is a leakage term, the second term is the input signal and the third term is a noise source. Since noise is uncorrelated between different hidden units, they evolve independently. Eq. (1) describes an Ornstein–Uhlembeck diffusion process (Tuckwell, 1988) and, for fixed values of σ, the equilibrium distribution of z_μ is a Gaussian distribution centered around the input signal, which is equal to

$$\Pr(z_\mu|\sigma) = \sqrt{\frac{\beta}{2\pi}}\exp\left[-\frac{\beta}{2}\left(z_\mu - \sum_i \xi_i^\mu \sigma_i\right)^2\right]. \tag{2}$$

In order for this equilibrium distribution to hold, the activity of digital units σ must be constant, while in fact it depends on time. However, we assume that the timescale of diffusion T is much faster than the rate at which the digital units are updated. Therefore, a different equilibrium distribution for z, characterized by different values of σ, holds between each subsequent update of σ. Since hidden units are independent, their joint distribution is the product of individual distributions, i.e. $\Pr(z|\sigma) = \prod_{\mu=1}^P \Pr(z_\mu|\sigma)$.

The activity in the visible (digital) layer follows a standard Glauber dynamics for Ising-type systems (Amit, 1992). At a specified sequence of time intervals (much larger than T), the activity of units in the digital layer is updated randomly according to a probability that depends on their input. While updating the digital units σ, the analog variables z are fixed, namely the update of digital units is instantaneous. The activity of a unit σ_i is