Theory and Methods for the Ferromagnetic Ising Model

Jay Shen
Department of Physics
University of Chicago
Chicago, IL 60637
jshe@uchicago.edu

Mark Lee
Department of Statistics
University of Chicago
Chicago, IL 60637
markyl@uchicago.edu

Abstract

The Ising model is a historically important problem in statistical mechanics. Although it was originally proposed as a crude approximation of ferromagnetic phenomena, it furnished statistical mechanics, and later probabilistic science as whole, with a new paradigm of graph-based modeling that would prove influential. Especially in the distilled, theoretical study of probabilistic graphical models (PGMs), which today is its own field, many methods developed for Ising models and spin-glasses have been repurposed and reinterpreted, as have the rich physical vocabulary of energies, entropy, and partition functions.

In this paper, we pay homage to the Ising model by applying modern methods to the original problem of ferromagnetism. We discuss the theory of the Ising model both from a PGMs and statistical mechanics point of view, examining the correspondence. We then evaluate a number of approaches to inference—Markov Chain Monte Carlo, belief propagation, and variational inference. Finally, we show the successes and limitations of the Ising model in the context of physical phenomena.

1 Theory of the Ising Model

A particle with magnetic moment $\vec{\mu}$ in an external magnetic field $\vec{B}$ has potential energy:

$$U = -\vec{\mu} \cdot \vec{B}$$

A particle also creates an instrinsic magnetic field defined by magnetic moment $\vec{\mu}$. The pairwise interactions between particles, described by exchange constants J_{ij}, thus also have potential energy:

$$U = -J_{ij}\vec{\mu_1} \cdot \vec{\mu_2}$$

We will ignore the short-range Pauli exclusion interactions.

Now, consider a collection of particles $\vec{\mu_i}$ in the prescence of an external magnetic field $\vec{B}$. The Hamiltonian, which specifies the total energy of the system, is defined by the sum of all pairwise and all magnetic field energies:

$$E = -\frac{1}{2}\sum_i \sum_j J_{ij}\vec{\mu_i} \cdot \vec{\mu_j} - \sum_i \vec{\mu_i} \cdot \vec{B} \tag{1}$$

In most cases, working with this Hamiltonian is intractable. Luckily, we can simplify it by making several strategic assumptions.

First, we assume all atoms are fermionic spin-$\frac{1}{2}$ and identical. Then the magnetic moments simplify to $\vec{\mu_i} = \mu\sigma_i$, where $\sigma_i \in \{-1, 1\}$ is the spin of the atom.

Preprint. Under review.

Second, we make a mean-field approximation and assume that the atoms are arranged in a regular lattice, and that the strength of pairwise interactions are negligible for non-adjacent pairs.

Coalescing constants, the Hamiltonian reduces to:

$$E = -J \sum_i \sum_{j \in adj(i)} \sigma_i \sigma_j - \sum_i \sigma_i B_i \tag{2}$$

Following the Boltzmann distribution, the probability of some state $\vec{\sigma}$ at inverse temperature β is:

$$P(\vec{\sigma}) = \frac{1}{Z} e^{-\beta E(\vec{\sigma})} \tag{3}$$

2 Inference on Ising Models

In most cases, we are interested in maximum a posteriori (MAP) inference on Ising models, namely, the most likely instantiation of variables. In the context of the physical interpretation, this corresponds to finding the equilibrium distribution of spins that minimize the energy.

We will now discuss some approaches to MAP inference on Ising models.

2.1 Sample Based

Sample based Monte Carlo methods aim to discover a MAP estimate by repeated sampling from the distribution. Markov Chain Monte Carlo (MCMC) improves on this approach by, at each step, conditioning the distribution on all previous samples and defining a Markov chain of states and transition probabilities.

In the context of the Ising model, sampling from the conditioned distribution corresponds to flipping spins of some spin state $\vec{\sigma}$ to get some new state $\vec{\sigma}'$. Intuitively, the transition to state $\vec{\sigma}'$ is favorable if $P(\vec{\sigma}') > P(\vec{\sigma})$. Thus, we define the transition "probability" to be:

$$P_{\vec{\sigma} \to \vec{\sigma}'} = \frac{P(\vec{\sigma}')}{P(\vec{\sigma})} = e^{-\beta(E(\vec{\sigma}') - E(\vec{\sigma}))}$$

In the case where only one spin σ_i flips, the energy difference in the exponential is:

$$E(\vec{\sigma}') - E(\vec{\sigma}) = 2\sigma_i (J \sum_{j \in adj(i)} \sigma_j + B_i)$$

So, we get:

$$P_{\vec{\sigma} \to \vec{\sigma}'} = \exp\Big[-2\beta\sigma_i (J \sum_{j \in adj(i)} \sigma_j + B_i) \Big]$$

Our policy is thus: if the transition lowers the energy, accept it. Otherwise, accept it with probability $P_{\vec{\sigma} \to \vec{\sigma}'}$.

References