Z. Phys. B – Condensed Matter 68, 233–235 (1987)

Condensed Matter
Zeitschrift für Physik B

Probabilistic Characterization of the Ising Model

H. Thomas
snark ag, CH-8037 Zürich, Switzerland

Received April 29, 1987; revised version June 19, 1987

Dedicated to Professor Harry Thomas on the occasion of his 60th birthday

The object of this paper is to show that Ising models can be completely characterized by the probabilistic Markov property and to derive some conclusions from this relationship.

Introduction

The study of the Ising model continues to be a challenge for theoretical physicists and mathematicians. Although the concept of the Ising model is easily stated and understood, it turns out to be quite difficult to derive its properties or to generate it on a computer.

As a very simple model that exhibits phase transitions, it is of interest in solid state physics. As a probability-theoretical model for multidimensional arrays of random variables that are *not* independent but locally coupled, it may prove to be of interest in many other areas, e.g. image processing, sociology etc.).

The following is a mathematical discussion of the concept of the Ising model. A purely probabilistic characterization of the Ising model is presented which is independent of specific values of the random variables or of the details of their topological arrangement as well as of specific formulas for their joint probability. The proof is discussed and leads to interesting variants for the definition of the interaction potential (e.g. such that the partition function becomes the entropy of the system).

Notions and Notation

In the following we consider only regular finite families of random variables $(\xi_t)_{t\in T}$ with a finite set of values S_t at each point $t\in T$. Such a family is determined by the probability $P((x_t)_{t\in T}) = Pr(\xi_t = x_t \,|\, \forall t\in T)$ for every configuration $(x_t)_{t\in T}$ in the cartesian product of the S_t. We call $\mathscr{C}_A = \prod_{t\in A} S_t$ the set of configurations on the subset $A\subset T$. From now on ξ_A will be used for $(\xi_t)_{t\in A}$, $\mathbf{x}_A$ for $(x_t)_{t\in A}$, etc. If $P(\mathbf{x}_T)>0$ for all $\mathbf{x}_T\in\mathscr{C}_T$ then we call ξ_T regular. Of course, the probability is normed: $\sum_{x_T\in\mathscr{C}_T} P(\mathbf{x}_T)=1$. For sets $A, B\in T$ with $A\cap B=\emptyset$ we call ξ_A independent of ξ_B if $P_{A\cup B}(\mathbf{x}_A\circ\mathbf{x}_B) = P_A(\mathbf{x}_A)\cdot P_B(\mathbf{x}_B)$, where $\mathbf{x}_A\circ\mathbf{x}_B$ denotes the configuration on $A\cup B$ which is composed of $\mathbf{x}_A$ on A and $\mathbf{x}_B$ on B and $P_A(\mathbf{x}_A) = \sum_{y_{T\setminus A}\in\mathscr{C}_{T\setminus A}} P(\mathbf{x}_A\circ\mathbf{y}_{T\setminus A})$ is the marginal probability distribution of ξ_A. Similarly we call ξ_A (conditionally) independent of ξ_B given ξ_C if $P_{A\cup B\cup C}(\mathbf{x}_A\circ\mathbf{x}_B|\mathbf{x}_C) = P_{A\cup C}(\mathbf{x}_A|\mathbf{x}_C)\cdot P_{B\cup C}(\mathbf{x}_B|\mathbf{x}_C)$, where $P_{A\cup B}(\mathbf{x}_A|\mathbf{x}_B) = P_{A\cup B}(\mathbf{x}_A\circ\mathbf{x}_B)/P_B(\mathbf{x}_B)$. Thus ξ_A is conditionally independent of ξ_B given ξ_C if and only if

$$P_S(\mathbf{x}_A\circ\mathbf{x}_B\circ\mathbf{x}_C)\cdot P_S(\mathbf{y}_A\circ\mathbf{y}_B\circ\mathbf{x}_C) = P_S(\mathbf{y}_A\circ\mathbf{x}_B\circ\mathbf{x}_C)\cdot P_S(\mathbf{x}_A\circ\mathbf{y}_B\circ\mathbf{x}_C) \tag{1}$$

for all $\mathbf{x}_A, \mathbf{x}_B, \mathbf{x}_C, \mathbf{y}_A, \mathbf{y}_B$, where $S=A\cup B\cup C$.

We proceed to endow the index set T with a topology. We call $\mathscr{N}\subset T\times T$ a neighborhood relation, if $\mathscr{N}$ is symmetric and reflexive. Given a neighborhood relation we call a subset $S\subset T$ a simplex if it is not empty and any two $s, t\in S$ are neighbors. Let $\mathscr{S}\subset\mathscr{P}(T)$ be the set of all simplices. For any $S\subset T$ the set ∂S consisting of all $t\in T\setminus S$ which are neighbors of some $s\in S$ is called the boundary of S.

Markov Nets and Ising Models

We call ξ_T a ($\mathscr{N}$-)Markov net if for every $S\subset T$ ξ_S is conditionally independent of $\xi_{T\setminus(S\cup\partial S)}$ given $\xi_{\partial S}$.

In view of (1) for every fixed $\mathbf{z}_T \in \mathscr{C}_T$

$$P(\mathbf{x}_s \circ \mathbf{x}_t \circ \mathbf{z}_{T\setminus\{s,t\}}) \cdot P(\mathbf{z}_T) = P(\mathbf{x}_s \circ \mathbf{z}_{T\setminus\{s\}}) \cdot P(\mathbf{x}_t \circ \mathbf{z}_{T\setminus\{t\}}) \qquad \forall \mathbf{x}_s \in S_s,\ \forall \mathbf{x}_t \in S_t \tag{2}$$

holds for Markov nets, if s and t are not neighbors.

We call ξ_T a ($\mathcal{N}$-)Ising model if

$$P(\mathbf{x}_T) = \frac{1}{Z} \cdot \exp\left(\sum_{S \in \mathscr{S}} I_S(\mathbf{x}_S)\right),$$

where the sum is over all simplices and $\mathbf{x}_S$ is assumed to be the restriction of $\mathbf{x}_T$ to S. Z is a normalization factor that ensures that $P(\mathbf{x}_T)$ is normed. It is called the partition function. Obviously ξ_T, when defined in this manner, is regular.

It is clear how these definitions lead to a markov chain if the neighborhood relation $\mathcal{N}$ is chosen as a one-dimensional open-ended chain. Similarly T can be made into an n-dimensional torus as is usual in the Ising model of physics.

Möbius Inversion

Lemma. *For two real functions H and I of the subsets $S \in \mathscr{P}(T)$ of the power set of T*

$$H(S) = \sum_{S' \subset S} I(S') \qquad \forall S \subset T$$

and

$$I(S) = \sum_{S' \subset S} (-1)^{|S \setminus S'|} H(S') \qquad \forall S \subset T$$

are equivalent. This is also called the inclusion exclusion principle.

The proof of the Lemma is based on the fact that the number of sets S' of cardinality k which are subsets of S and supersets of S'' is given by the binomial coefficient $\binom{|S \setminus S''|}{k - |S''|}$ and therefore for fixed S and S''

$$\sum_{S'' \subset S' \subset S} (-1)^{|S' \setminus S''|} = \sum_{k=|S''|}^{|S|} \binom{|S' \setminus S''|}{k - |S''|} (-1)^{k - |S''|}$$
$$= \sum_{k=0}^{|S \setminus S''|} \binom{|S \setminus S''|}{k} (-1)^k = \delta(S'', S).$$

Thus

$$\sum_{S' \subset S} \sum_{S'' \subset S'} (-1)^{|S' \setminus S''|} H(S'')$$
$$= \sum_{S'' \subset S} H(S'') \cdot \sum_{S'' \subset S' \subset S} (-1)^{|S' \setminus S''|} = H(S)$$

and conversely

$$\sum_{S' \subset S} \sum_{S'' \subset S'} (-1)^{|S \setminus S'|} I(S'')$$
$$= \sum_{S'' \subset S} (-1)^{|S \setminus S''|} I(S'') \cdot \sum_{S'' \subset S' \subset S} (-1)^{-|S' \setminus S''|} = I(S).$$

Obviously the relations $H(S) = \sum_{S' \subset S} I(S')$ form a non-singular system of equations, even when we consider sublattices of the lattice of all subsets. The coefficients for the inverse expression for the $I(S)$ will depend, however, on the particular sublattice.

Equivalence of Markov Nets and Ising Models

Theorem: *The class of $\mathcal{N}$-Markov nets is identical with the class of $\mathcal{N}$-Ising models.*

In the proof we slightly generalize the reasoning of Sherman. We choose an arbitrary but fixed configuration $\mathbf{z}_T \in \mathscr{C}_T$. For each $S \subset T$ we define

$$H_S(\mathbf{x}_S) = \log(P(\mathbf{x}_S \circ \mathbf{z}_{T\setminus S})). \tag{3}$$

Each H_S is thus a function of the configurations on S. Using Möbius inversion we define interaction potentials

$$I_S(\mathbf{x}_S) = \sum_{S' \subset S} (-1)^{|S \setminus S'|} H_{S'}(\mathbf{x}_{S'}).$$

The probability distribution can then be written as

$$P(\mathbf{x}_T) = \exp\left(\sum_{S \in \mathscr{P}(T)} I_S(\mathbf{x}_S)\right). \tag{4}$$

If we set $I_\emptyset = -\log(Z)$ and all $I_S \equiv 0$ except for $S \in \mathscr{S} \cup \{\emptyset\}$, then P describes an Ising model.

It is obvious that every Ising model has the Markov property. To show the converse we examine the interaction potential of a set $S \notin \mathscr{S} \cup \{\emptyset\}$. There exist $s, t \in S$ which are not neighbors. Because of (2) we have

$$I_S(\mathbf{x}_S) = \sum_{S' \subset S} (-1)^{|S \setminus S'|} H_{S'}(\mathbf{x}_{S'})$$
$$= \sum_{S' \subset S \setminus \{s,t\}} (-1)^{|S \setminus S'|} \{H_{S' \cup \{s,t\}}(\mathbf{x}_{S' \cup \{s,t\}}) - H_{S' \cup \{s\}}(\mathbf{x}_{S' \cup \{s\}})$$
$$- H_{S' \cup \{t\}}(\mathbf{x}_{S' \cup \{t\}}) + H_{S'}(\mathbf{x}_{S'})\} = 0.$$

Therefore every Markov net is an Ising model.

Ramifications

In the proof of the above theorem we have defined the $H_S(\mathbf{x}_S) = H_S(\mathbf{x}_S, \mathbf{z}_T)$ using an arbitrary but fixed configuration $\mathbf{z}_T$ on $T\setminus S$. $H_\emptyset = I_\emptyset$, the negative logarithm of the partition function, becomes just $\log(P(\mathbf{z}_T))$. Obviously any other $\mathbf{z}_T$ could have been used as well as a linear combination

$$H_S(\mathbf{x}_S) = \sum_{\mathbf{z}_T \in \mathscr{C}_T} \lambda_{\mathbf{z}_T} H_S(\mathbf{x}_S, \mathbf{z}_T),$$

where $\sum_{\mathbf{z}_T \in \mathscr{C}_T} \lambda_{\mathbf{z}_T} = 1$.

The following choices of $\lambda_{\mathbf{z}_T}$ appear to be of interest:

1) $\lambda_{\mathbf{z}_T} = P(\mathbf{z}_T)$: with this choice the partition function becomes

$$-\log(Z) = I_\emptyset = \sum_{\mathbf{z}_T \in \mathscr{C}_T} P(\mathbf{z}_T) \log(P(\mathbf{z}_T)).$$

2) $\lambda_{\mathbf{z}_T} = 1/|\mathscr{C}_T|$. This yields the traditional Ising interaction potentials if the state set $S_t = \{-1, 1\}$ at each point $t \in T$. We then have

$$-\log(Z) = \frac{1}{|\mathscr{C}_T|} \cdot \sum_{\mathbf{z}_T \in \mathscr{C}_T} \log(P(\mathbf{z}_T)),$$

i.e. the partition function is the inverse of the geometrical mean of the probabilities of all configurations.

In the case of the open-ended one-dimensional Markov chain, we can choose the $I_S(\mathbf{x}_S)$ as functions of the marginal probabilities $P_S(\mathbf{x}_S)$. This greatly facilitates calculation and simulation of such random "nets".

In a different context R.F. Hauser has proposed to study the resulting "interaction potentials", with $H_S(\mathbf{x}_S) = \log(P_S(\mathbf{x}_S))$. Following the argument of the proof of the main theorem, it can be seen, that the interaction potentials I_S have to be defined in a different manner in order to achieve $I_S \equiv 0$ if S is not a simplex. For Markov chains this can be achieved if one plugs the lattice of connected subsets into the Möbius inversion theorem. This yields the inversion formula

$$I_{\{i,\ldots,i+k\}} = H_{\{i,\ldots,i+k\}} - H_{\{i,\ldots,i+k-1\}} - H_{\{i+1,\ldots,i+k\}} + H_{\{i+1,\ldots,i+k-1\}}.$$

This is easily seen to be zero for $k > 1$.

For cyclic Markov chains or two-dimensional Markov nets an expression for the total probability as a function of the simplex probabilities is not to be expected, as the simplex probabilities are not likely to be independent.

References

1. Averitsev, M.B.: Probl. Peredachi Inf. **6**, 100 (1970)
2. Dobrushin, R.L.: Theory of Probab. Appl. **13**, 197 (1968)
3. Grimmett, G.R.: Bull. London Math. Soc. **5**, 81 (1973)
4. Hauser, R.F.: Bestimmung des lokalen Informationsgehaltes endlicher Familien von diskreten Zufallsvariablen angewandt auf das Problem der Konturerkennung in digitalen Bildern. Diss. ETH Nr. 7468, Zürich 1984
5. Preston, C.J.: Adv. Appl. Prob. **5**, 242 (1973)
6. Sherman, S.: Isr. J. Math. **14**, 93 (1973)
7. Spitzer, F.: Am. Math. Monthly **78**, 142 (1971)

Hartwig Thomas
Kienastenwiesweg 31
CH-8053 Zürich
Switzerland