8 Blocks, Markov chains, Ising model

8a Introductory remarks

The following three questions are related more closely than it may seem.

8a1 Question. 100 children stay in a ring, 40 boys and 60 girls. Among the 100 pairs of neighbors, 20 pairs are heterosexual (a girl and a boy); others are not. What about the number of all such configurations?

8a2 Question. A Markov chain with two states (0 and 1) is given via its 2×2-matrix of transition probabilities. What about the probability that the state 1 occurs 60 times among the first 100?

8a3 Question. (*Ising model*) A one-dimensional array of n spin-1/2 particles is described by the configuration space $\{-1, 1\}^n$. Each configuration $(s_1, \dots, s_n) \in \{-1, 1\}^n$ has its energy

$$H_n(s_1, \dots, s_n) = -\frac{1}{2}(s_1 s_2 + \dots + s_{n-1} s_n) - h(s_1 + \dots + s_n) \, ;$$

here $h \in \mathbb{R}$ is a parameter. (It is the strength of an external magnetic field, while the strength of the nearest neighbor coupling is set to 1.) What about the dependence of the energy and the mean spin $(s_1 + \dots + s_n)/n$ on h and the temperature?

Tossing a fair coin n times we get a random element $(\beta_1, \dots, \beta_n)$ of $\{0, 1\}^n$, and may consider the $n-1$ pairs $(\beta_1, \beta_2), (\beta_2, \beta_3), \dots, (\beta_{n-1}, \beta_n)$. We introduce *pair frequencies*

$$\frac{K'}{n-1} = \Big(\frac{K'_{00}}{n-1}, \frac{K'_{01}}{n-1}, \frac{K'_{10}}{n-1}, \frac{K'_{11}}{n-1} \Big) \in P\big(\{0,1\}^2\big) \, ,$$
$$K'_{ab} = \#\{ i = 1, \dots, n-1 : \beta_i = a, \beta_{i+1} = b \} \, ,$$

and their (joint) distribution

$$\mu'_n \in P\big(P(\{0,1\}^2)\big)\,,$$
$$\int f \, d\mu'_n = \frac{1}{2^n} \sum_{\beta \in \{0,1\}^n} f\Big(\frac{K'_{00}}{n-1}, \frac{K'_{01}}{n-1}, \frac{K'_{10}}{n-1}, \frac{K'_{11}}{n-1}\Big)\,.$$

Alternatively, we may consider n pairs $(\beta_1, \beta_2), (\beta_2, \beta_3), \dots, (\beta_{n-1}, \beta_n), (\beta_n, \beta_1)$, the corresponding pair frequencies $\frac{K''}{n} = \big(\frac{K''_{00}}{n}, \frac{K''_{01}}{n}, \frac{K''_{10}}{n}, \frac{K''_{11}}{n}\big)$ and their (joint) distribution μ''_n.

8a4 Exercise. LD-convergence of $(\mu'_n)_n$ is equivalent to LD-convergence of $(\mu''_n)_n$, and their rate functions (if exist) are equal.

Prove it.

Hint: recall 5d.

You may say that what we call μ'_n should be called μ'_{n-1} instead; but it does not matter in the following sense.

8a5 Exercise. Let μ_n be probability measures on a compact metrizable space K. Then LD-convergence of $(\mu_n)_n$ is equivalent to LD-convergence of $(\mu_{n+1})_n$, and their rate functions (if exist) are equal.

Prove it.

Hint: similar to 2a17.

8a6 Exercise. Explain, why LD-convergence of $(\mu'_n)_n$ cannot be derived from Theorem 5a9 (Mogulskii's theorem) combined with Theorem 2b1 (the contraction principle).

8a7 Exercise. If the rate function I for $(\mu'_n)_n$, $(\mu''_n)_n$ exists then

$$\min\{I(x_{00}, x_{01}, x_{10}, x_{11}) : x_{01} + x_{10} = z\} = I_{0.5}(z)$$

for all $z \in [0,1]$. (See (3a5) for $I_{0.5}$.)

Prove it.

Hint: consider the measure preserving map $\{0,1\}^n \to \{0,1\}^{n-1}$, $(\beta_1, \dots, \beta_n) \mapsto (\beta_1 \oplus \beta_2, \beta_2 \oplus \beta_3, \dots, \beta_{n-1} \oplus \beta_n)$; here '$\oplus$' stands for the sum mod 2 (called also XOR = 'exclusive or').

We turn to Markov chains. Let a 2×2-matrix

$$\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$$

be given, $p_{ab} \in [0,1]$, $p_{00}+p_{01} = 1$, $p_{10}+p_{11} = 1$. In addition, let $p_0, p_1 \in [0,1]$ be given such that $p_0 + p_1 = 1$. We define the probability of a history $(s_0, \dots, s_n) \in \{0,1\}^{n+1}$ by

$$P_n(s_0, \dots, s_n) = p_{s_0} p_{s_0,s_1} p_{s_1,s_2} \dots p_{s_{n-1},s_n} ;$$

clearly, we get a probability measure P_n on $\{0,1\}^{n+1}$. The pair frequencies K/n get their distribution ν_n,

$$\int f \, d\nu_n = \sum_{s \in \{0,1\}^{n+1}} f\Big(\frac{K_{00}}{n}, \frac{K_{01}}{n}, \frac{K_{10}}{n}, \frac{K_{11}}{n} \Big) P_n(s) .$$

8a8 Exercise. LD-convergence of $(\nu_n)_n$ does not depend on p_0, p_1 as long as $p_0, p_1 \ne 0$. Also the rate function (if exists) does not depend.

Prove it.

Hint: use 8a9 below.

8a9 Exercise. Let μ_n, ν_n be probability measures on a compact metrizable space K. Assume that there exists $C \in (0, \infty)$ such that $\mu_n \le C\nu_n$ and $\nu_n \le C\mu_n$ for all n. Then LD-convergence of $(\mu_n)_n$ is equivalent to LD-convergence of $(\nu_n)_n$, and their rate functions (if exist) are equal.

Prove it.

Hint: $C^{1/n} \to 1$.

8a10 Exercise. Assuming that $p_{00}, p_{01}, p_{10}, p_{11}$ do not vanish, remove the restriction $p_0, p_1 \ne 0$ in 8a8.

Hint: similarly to 8a4, the pair (s_0, s_1) does not matter.

8a11 Exercise. LD-convergence of $(\nu_n)_n$ does not depend on $p_{00}, p_{01}, p_{10}, p_{11}$ as long as they do not vanish.

Prove it.

Hint: similarly to 3a, 3b use Theorem 2c1 (titled LDP).

The rate function (if exists) does not depend on the initial probabilities p_a, but does depend on the transition probabilities p_{ab}; namely, the rate function must contain (additively) the terms

$$-x_{00} \ln p_{00} - x_{01} \ln p_{01} - x_{10} \ln p_{10} - x_{11} \ln p_{11} .$$

It means that we may restrict ourselves to the simplest matrix

$$\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} ,$$

thus reducing 8a2 to 8a1.

We turn to the array of spin-1/2 particles. The energy $H_n(s_1,\dots,s_n)$ depends on the spin configuration $(s_1,\dots,s_n)\in\{-1,1\}^n$ only via pair frequencies,

$$H_n(s_1,\dots,s_n)=(n-1)\Big(\frac{K'_{+-}}{n-1}+\frac{K'_{-+}}{n-1}-\frac{K'_{++}}{n-1}-\frac{K'_{--}}{n-1}\Big)\,.$$

Similarly to 3d, we have the uniform distribution U_n and the Gibbs measure G_n on $\{-1,1\}^n$; $\mathrm{d}G_n/\mathrm{d}U_n=\mathrm{const}_n\cdot\mathrm{e}^{-\beta H_n}$. The distribution of $\frac{K'}{n-1}$ w.r.t. U_n is μ'_n; the distribution of $\frac{K'}{n-1}$ w.r.t. G_n is ν_n,[1]

$$\nu_n=\mathrm{const}_n\cdot\exp\bigg(-\beta(n-1)\Big(\frac{K'_{+-}}{n-1}+\frac{K'_{-+}}{n-1}-\frac{K'_{++}}{n-1}-\frac{K'_{--}}{n-1}\Big)\bigg)\cdot\mu'_n\,.$$

If $(\mu'_n)_n$ satisfies LDP with a rate function I, then $(\nu_n)_n$ satisfies LDP with the rate function J,

$$J\Big(\frac{K'_{++}}{n-1},\frac{K'_{+-}}{n-1},\frac{K'_{-+}}{n-1},\frac{K'_{--}}{n-1}\Big)=$$
$$=I\Big(\frac{K'_{++}}{n-1},\frac{K'_{+-}}{n-1},\frac{K'_{-+}}{n-1},\frac{K'_{--}}{n-1}\Big)+\beta\Big(\frac{K'_{+-}}{n-1}+\frac{K'_{-+}}{n-1}-\frac{K'_{++}}{n-1}-\frac{K'_{--}}{n-1}\Big)+\mathrm{const}\,,$$

and we may proceed as in 3d, taking into account that

$$\frac{s_1+\dots+s_n}{n}=\frac{K''_{++}}{n}+\frac{K''_{+-}}{n}-\frac{K''_{-+}}{n}-\frac{K''_{--}}{n}\approx\frac{K'_{++}}{n-1}+\frac{K'_{+-}}{n-1}-\frac{K'_{-+}}{n-1}-\frac{K'_{--}}{n-1}\,.$$

8b Pair frequencies: combinatorial approach

We consider the cyclic pair frequencies[2] $\frac{K}{n}$ for $\beta\in\{0,1\}^n$,

$$K_{ab}(\beta)=\#\{i=1,\dots,n:\beta_i=a,\beta_{i+1}=b\}\quad\text{for }a,b\in\{0,1\}\,,$$

where β_{n+1} is interpreted as β_1. Clearly, $K_{01}(\beta)=K_{10}(\beta)$ and $K_{00}(\beta)+K_{01}(\beta)+K_{10}(\beta)+K_{11}(\beta)=n$; thus, $K_{01}(\beta)=K_{10}(\beta)=\frac12\big(n-K_{00}(\beta)-K_{11}(\beta)\big)$.

Let us denote by $N(k_{00},k_{11})$ the number of all $\beta\in\{0,1\}^n$ such that $K_{00}(\beta)=k_{00}$ and $K_{11}(\beta)=k_{11}$.

[1] This ν_n is not related to the Markov chain…

[2] These $\frac{K}{n}$ are $\frac{K''}{n}$ of 8a.

8b1 Lemma. Let $k_{00}, k_{11} \in \{0,1,2,\dots\}$ satisfy $\frac12(n-k_{00}-k_{11}) \in \{1,2,\dots\}$, then

$$1 \le \frac{N(k_{00},k_{11})}{\binom{\frac12(n+k_{00}-k_{11})-1}{k_{00}}\binom{\frac12(n-k_{00}+k_{11})-1}{k_{11}}} \le n\,.$$

Proof. Define $k_{01} = k_{10} = \frac12(n-k_{00}-k_{11})$. There exist exactly $\binom{k_{00}+k_{01}-1}{k_{01}-1} = \binom{k_{00}+k_{01}-1}{k_{00}}$ partitions of the number k_{00} into k_{01} nonnegative integral summands; and similarly, $\binom{k_{11}+k_{10}-1}{k_{11}}$ partitions of k_{11} into k_{10} summands. Having such partitions $k_{00} = i_1 + \dots + i_{k_{01}}$, $k_{11} = j_1 + \dots + j_{k_{10}}$, we construct $\beta \in \{0,1\}^n$ by concatenation:

$$\beta = 0^{i_1+1}1^{j_1+1}0^{i_2+1}1^{j_2+1}\dots 0^{i_{k_{01}}+1}1^{j_{k_{10}}+1}\,.$$

Clearly, $K_{00}(\beta) = k_{00}$, $K_{11}(\beta) = k_{11}$, and $i_1,\dots,i_{k_{01}}, j_1,\dots,j_{k_{10}}$ are uniquely determined by β. We see that the product $\binom{k_{00}+k_{01}-1}{k_{00}} \cdot \binom{k_{11}+k_{10}-1}{k_{11}}$ is the number of all $\beta \in \{0,1\}^n$ such that $K_{00}(\beta) = k_{00}$, $K_{11}(\beta) = k_{11}$, $\beta_1 = 0$ and $\beta_n = 1$. The lemma follows. $\square$

The case $n - k_{00} - k_{11} = 0$ is special but harmless (think, why), we put it aside. Denote

$$x = \frac{k_{00}}{n}, \quad y = \frac{k_{11}}{n}, \quad z = 1-x-y\,, \quad \Big(= \frac{k_{01}+k_{10}}{n}\Big)$$
$$u = x + \frac z2 = \frac{1+x-y}{2}, \quad \text{(the frequency of zeros)}$$
$$v = y + \frac z2 = \frac{1-x+y}{2} = 1-u\,.$$

Using 8b1,

$$\big(N(k_{00},k_{11})\big)^{1/n} \sim \binom{nu-1}{nx}^{1/n}\binom{nv-1}{ny}^{1/n} \sim$$
$$\sim \binom{nu}{nx}^{1/n}\binom{nv}{ny}^{1/n} = \Big(\frac{(nu)!(nv)!}{(nx)!(ny)!(nz/2)!^2}\Big)^{1/n}$$

as $n \to \infty$, uniformly in k_{00}, k_{11}. However, $(na)!^{1/n} \sim (na/\mathrm{e})^a$ uniformly in $a \in [0,1]$ (recall the hint to 3a3). Thus,

$$\big(N(k_{00},k_{11})\big)^{1/n} \sim \frac{(nu/\mathrm{e})^u(nv/\mathrm{e})^v}{(nx/\mathrm{e})^x(ny/\mathrm{e})^y(nz/(2\mathrm{e}))^z} = \frac{u^uv^v}{x^xy^y(z/2)^z}\,.$$

Let β be distributed uniformly on $\{0,1\}^n$, then the pair frequencies are distributed μ''_n (recall 8a).

8b2 Exercise. $(\mu''_n)_n$ satisfies LDP with the rate function

$$I(x_{00}, x_{01}, x_{10}, x_{11}) = x \ln x + y \ln y + z \ln z - u \ln u - v \ln v + (1 - z) \ln 2\,,$$

where

$$x = x_{00}\,, \quad y = x_{11}\,, \quad z = 1 - x - y = x_{01} + x_{10}\,,$$
$$u = x + \frac{z}{2} = \frac{1 + x - y}{2}\,, \quad v = y + \frac{z}{2} = \frac{1 - x + y}{2} = 1 - u\,,$$

and $x_{00}, x_{01}, x_{10}, x_{11} \in [0, 1]$ satisfy $x_{00} + x_{01} + x_{10} + x_{11} = 1$ and $x_{01} = x_{10}$.

Prove it.

Hint: similar to 3a4.

We may write just

$$I(x, y) = x \ln x + y \ln y + (1 - x - y) \ln(1 - x - y) - \\ - \frac{1 + x - y}{2} \ln \frac{1 + x - y}{2} - \frac{1 - x + y}{2} \ln \frac{1 - x + y}{2} + (x + y) \ln 2\,. \tag{8b3}$$

By 8a4, the same holds for $(\mu'_n)_n$.

By the weak law of large numbers (and a simple trick…), μ'_n concentrate near the point $x_{00} = x_{01} = x_{10} = x_{11} = 0.25$. At this point $x = y = 0.25$ and $z = u = v = 0.5$, thus $I(0.25, 0.25) = \frac{2}{4} \ln \frac{1}{4} - \frac{1}{2} \ln \frac{1}{2} + \frac{1}{2} \ln 2 = 0$, as it should be.

8b4 Exercise. Check by elementary calculation the equality of 8a7,

$$\min_{x+y=1-z} I(x, y) = I_{0.5}(z) \quad \text{for } z \in [0, 1]\,.$$

Hint: $\frac{\partial}{\partial x} I(x, y) = \ln x - \ln z - \frac{1}{2} \ln u + \frac{1}{2} \ln v + \ln 2$, $\frac{\partial}{\partial y} I(x, y) = \ln y - \ln z + \frac{1}{2} \ln u - \frac{1}{2} \ln v + \ln 2$; take the difference; show that the minimum is reached when $x = y$.

Think about the 'proportion'

$$\frac{X}{8b2} = \frac{5a9}{3a4}\,;$$

could you find X (formulate, or even prove)?

See also [4, Sect. II.2] for more than two states.

8c Markov chains

We return to the Markov chain, assuming that the transition probabilities p_{ab} do not vanish. The pair frequencies are distributed ν_n. Recall 8a8–8a11.

8c1 Exercise. $(\nu_n)_n$ satisfies LDP with the rate function

$$\begin{aligned} J(x_{00}, x_{01}, x_{10}, x_{11}) = I(x_{00}, x_{01}, x_{10}, x_{11}) - \\ - x_{00} \ln p_{00} - x_{01} \ln p_{01} - x_{10} \ln p_{10} - x_{11} \ln p_{11} - \ln 2 \,, \end{aligned}$$

that is,

$$J(x, y) = I(x, y) - x \ln p_{00} - y \ln p_{11} - \frac{1 - x - y}{2} \big(\ln(1 - p_{00}) + \ln(1 - p_{11}) \big) - \ln 2 \,,$$

where I is given by (8b3).

Prove it.

Hint: in 2c1, $c_n = 2^n$ (since $p_{00} + p_{01} = 1$ and $p_{10} + p_{11} = 1$).

8c2 Exercise. For all $\varphi, \psi \in (0, \pi/2)$,

$$\min_{x,y \ge 0, x+y \le 1} \bigg(I(x, y) + x \ln \frac{\sin\varphi \sin\psi}{\cos^2\varphi} + y \ln \frac{\sin\varphi \sin\psi}{\cos^2\psi} \bigg) = \ln(2 \sin\varphi \sin\psi) \,.$$

Prove it.

Hint: $p_{00} = \cos^2\varphi$, $p_{11} = \cos^2\psi$; use 2a19.

An elementary derivation of 8c2 is possible but more tedious. First, we find the minimizer.

Let the function $(x, y) \mapsto I(x, y) + x \ln \frac{\sin\varphi \sin\psi}{\cos^2\varphi} + y \ln \frac{\sin\varphi \sin\psi}{\cos^2\psi}$ on the triangle $x, y \ge 0$, $x + y \le 1$ have a local minimum at (x, y). As before, $z = 1 - x - y$, $u = (1 + x - y)/2$, $v = (1 - x + y)/2$.

8c3 Exercise. (x, y) is an interior point (that is, $x, y > 0$, $x + y < 1$), and

$$\begin{aligned} 2 \tan\varphi \tan\psi \sqrt{xy} = z \,, \\ xv \cos^2\psi = yu \cos^2\varphi \,. \end{aligned}$$

Prove it.

Hint: take the sum and the difference of $\frac{\partial}{\partial x} I(x, y)$, $\frac{\partial}{\partial y} I(x, y)$ (used in 8b4).

8c4 Exercise. Prove that

$$x = \frac{u(u - v) \cos^2\varphi}{u \cos^2\varphi - v \cos^2\psi} \,, \quad y = \frac{v(u - v) \cos^2\psi}{u \cos^2\varphi - v \cos^2\psi} \,.$$

Hint: both $x - y$ and x/y can be expressed in terms of u, v.

8c5 Exercise. Prove that

$$2(u-v)\sin\varphi\sin\psi = \sqrt{1-(u-v)^2}\,(\cos^2\varphi - \cos^2\psi)\,.$$

Hint: substitute 8c4 into the first equation of 8c3 and note that $2u = 1 + (u-v)$, $2v = 1-(u-v)$.

8c6 Exercise. Prove that

$$x = \frac{\cos^2\varphi\sin^2\psi}{\sin^2\varphi+\sin^2\psi}\,, \quad y = \frac{\sin^2\varphi\cos^2\psi}{\sin^2\varphi+\sin^2\psi}\,.$$

Hint: $u-v = \frac{\cos^2\varphi-\cos^2\psi}{\sin^2\varphi+\sin^2\psi} = \frac{\sin^2\psi-\sin^2\varphi}{\sin^2\varphi+\sin^2\psi}$.

The minimizer is found, and now we calculate the minimal value.

8c7 Exercise. Prove that

$$I(x,y) + x\ln\frac{\sin\varphi\sin\psi}{\cos^2\varphi} + y\ln\frac{\sin\varphi\sin\psi}{\cos^2\psi} = \ln(2\sin\varphi\sin\psi)\,.$$

Hint: the left-hand side is $x\ln\frac{x}{\cos^2\varphi} + y\ln\frac{y}{\cos^2\psi} + z\ln\frac{z}{2\sin\varphi\sin\psi} - u\ln u - v\ln v + \ln(2\sin\varphi\sin\psi)$; also $z = \frac{2\sin^2\varphi\sin^2\psi}{\sin^2\varphi+\sin^2\psi}$ and $u = \frac{\sin^2\psi}{\sin^2\varphi+\sin^2\psi}$.

This was the elementary derivation of 8c2.

However, there exists a simple probabilistic way to the minimizer! The Markov chain has a unique stationary distribution (p_0,p_1),

$$\begin{cases} p_0p_{00} + p_1p_{10} = p_0\,, \\ p_0p_{01} + p_1p_{11} = p_1\,; \end{cases}$$

$$p_1p_{10} = p_0p_{01}\,;$$

$$p_0 = \frac{p_{10}}{p_{01}+p_{10}}\,, \quad p_1 = \frac{p_{01}}{p_{01}+p_{10}}\,,$$

and every initial distribution converges to the stationary distribution (exponentially fast, in fact). Thus, the measures ν_n converge to (an atom at) the point

$$(x_{00},x_{01},x_{10},x_{11}) = (p_0p_{00},p_0p_{01},p_1p_{10},p_1p_{11})\,.$$

Substituting $p_{00} = \cos^2\varphi$, $p_{11} = \cos^2\psi$ we get

$$x_{00} = \frac{\cos^2\varphi\sin^2\psi}{\sin^2\varphi+\sin^2\psi}\,, \quad x_{11} = \frac{\sin^2\varphi\cos^2\psi}{\sin^2\varphi+\sin^2\psi}\,;$$

just 8c6...

The rate functions examined above are of the form $(x,y) \mapsto I(x,y) + Ax + By$ where I is given by (8b3) and $A,B \in \mathbb{R}$. However, did we cover all pairs $(A,B) \in \mathbb{R}^2$? Yes, we did, as is shown below.

8c8 Exercise. For every pair $(a,b) \in (0,\infty)^2$ there exists one and only one pair $(\varphi,\psi) \in (0,\pi/2)^2$ such that

$$\frac{\sin\varphi\sin\psi}{\cos^2\varphi} = a\,, \quad \frac{\sin\varphi\sin\psi}{\cos^2\psi} = b\,.$$

Prove it.

Hint: consider the curve $\frac{\cos\varphi}{\cos\psi} = \sqrt{b/a}$ in the square $(0,\pi/2)^2$ and check that the equation $\tan\varphi\tan\psi = \sqrt{ab}$ is satisfied exactly once on the curve.

8c9 Remark. Using the equality $(1+\tan^2\varphi)\cos^2\varphi = 1$ (and the same for ψ) one can find φ,ψ explicitly. Namely, $\cos^2\varphi$ satisfies a quadratic equation...

8d Ising model (one-dimensional)

As was noted in 8a, the Ising model[1] is described by the Gibbs measure G_n on $\{-1,1\}^n$, $\mathrm{d}G_n/\mathrm{d}U_n = \mathrm{const}_n \cdot \mathrm{e}^{-\beta H_n}$, and the corresponding distribution ν_n of pair frequencies. Also, LDP for $(\mu'_n)_n$ implies LDP for $(\nu_n)_n$ with the rate function

$$J(x_{++},x_{+-},x_{-+},x_{--}) = I(x_{++},x_{+-},x_{-+},x_{--}) + \\ + \beta H(x_{++},x_{+-},x_{-+},x_{--}) + \mathrm{const}\,,$$

where

$$H(x_{++},x_{+-},x_{-+},x_{--}) = -\frac{1}{2}(x_{++}+x_{--}-x_{+-}-x_{-+}) - h(u-v)\,,$$
$$u = x_{++}+x_{+-} = x_{++}+x_{-+}\,,$$
$$v = x_{-+}+x_{--} = x_{+-}+x_{--}\,.$$

That is,

$$J_{\beta,h}(x,y) = I(x,y) + \beta H(x,y) + \mathrm{const}\,,$$
$$H(x,y) = -\frac{1}{2}(1-2z) - h(x-y)\,;$$

as before, $z = 1-x-y$, and I is given by (8b3).

Clearly, $J_{\beta,h}$ is a rate function of the form $(x,y) \mapsto I(x,y) + Ax + By$ examined in 8c2–8c9. It has a single minimizer $(x_{\beta,h},y_{\beta,h})$, and ν_n converge to (the atom at) $(x_{\beta,h},y_{\beta,h})$. The minimizer can be written out explicitly

[1] Developed in 1926 by Ernst Ising (in his PhD dissertation); the young German-Jewish scientist was barred from teaching when Hitler came to power.

by solving a quadratic equation (recall 8c9). Having the minimizer one can calculate the energy $H(x_{\beta,h}, y_{\beta,h})$ and the mean spin $x_{\beta,h} - y_{\beta,h}$.

The dependence of $x_{\beta,h}$ and $y_{\beta,h}$ on β, h is (real-) analytic everywhere, which means absence of phase transitions.

See also [5, Sect. 7.4.3].

8e Pair frequencies: linear algebra approach

Consider again the cyclic pair frequencies $K''/n = K''(\beta_1, \dots, \beta_n)/n$ and their distribution μ''_n (introduced in 8a).

8e1 Exercise. For every matrix $A = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix}$,

$$\sum_{\beta_1,\dots,\beta_n} a_{00}^{K_{00}} a_{01}^{K_{01}} a_{10}^{K_{10}} a_{11}^{K_{11}} = \operatorname{trace}(A^n)\,.$$

Prove it.

Hint: straight from definitions (of matrix multiplication and trace).

Denote by λ_1, λ_2 the eigenvalues of A, then $\lambda_1 + \lambda_2 = \operatorname{trace}(A)$, and λ_1^n, λ_2^n are the eigenvalues of A^n, therefore

$$\operatorname{trace}(A^n) = \lambda_1^n + \lambda_2^n\,.$$

Assume that $a_{00} > 0$, $a_{01} > 0$, $a_{10} > 0$, $a_{11} > 0$, then $\lambda_1 + \lambda_2 > 0$ and

$$\bigl(\operatorname{trace}(A^n)\bigr)^{1/n} \to \max(\lambda_1, \lambda_2) \quad \text{as } n \to \infty\,.$$

8e2 Exercise. If $(\mu''_n)_n$ satisfies LDP with a rate function I, then

$$\min_x \bigl(I(x_{00}, x_{01}, x_{10}, x_{11}) - x_{00} \ln a_{00} - x_{01} \ln a_{01} - x_{10} \ln a_{10} - x_{11} \ln a_{11} \bigr) = \\ = -\ln \frac{\max(\lambda_1, \lambda_2)}{2}\,.$$

Prove it (not using 8b).

Hint: consider $\int f^n \, \mathrm{d}\mu''_n$ for $f(x_{00}, x_{01}, x_{10}, x_{11}) = a_{00}^{x_{00}} a_{01}^{x_{01}} a_{10}^{x_{10}} a_{11}^{x_{11}}$.

Taking into account that $K_{01} = K_{10}$ and $K_{00} + K_{01} + K_{10} + K_{11} = n$ we may restrict ourselves to $x_{01} = x_{10}$ and $x_{00} + x_{01} + x_{10} + x_{11} = 1$. Thus we take $x = x_{00}$, $y = x_{11}$ and get $x_{01} = x_{10} = z/2$ where $z = 1 - x - y$. Using $I(x, y)$ instead of $I(x_{00}, x_{01}, x_{10}, x_{11})$ we get

$$\min_{x,y \ge 0, x+y \le 1} \bigl(I(x, y) - x \ln a_{00} - y \ln a_{11} - z \ln \sqrt{a_{01} a_{10}} \bigr) = -\ln \frac{\max(\lambda_1, \lambda_2)}{2}\,.$$

Compare it with 8c2; there, $\max(\lambda_1, \lambda_2) = 1$.

We may restrict ourselves to matrices A such that $a_{01} = a_{10}$ and moreover, $a_{01} = a_{10} = 1$. Let

$$A = \begin{pmatrix} e^u & 1 \\ 1 & e^v \end{pmatrix} ,$$

then

$$\lambda_{1,2} = \frac{e^u + e^v}{2} \pm \sqrt{\Big(\frac{e^u + e^v}{2}\Big)^2 - e^u e^v + 1} = \frac{e^u + e^v}{2} \pm \sqrt{\Big(\frac{e^u - e^v}{2}\Big)^2 + 1} \, ;$$

$$\max(\lambda_1, \lambda_2) = \frac{e^u + e^v}{2} + \sqrt{\Big(\frac{e^u - e^v}{2}\Big)^2 + 1} \, .$$

Therefore

$$\min_{x,y \ge 0, x+y \le 1} \big(I(x,y) - ux - vy \big) = - \ln \left(\frac{e^u + e^v}{4} + \frac{1}{2} \sqrt{\Big(\frac{e^u - e^v}{2}\Big)^2 + 1} \right) .$$

We get the so-called Legendre-Fenchel transform of the rate function. (See also (3c4).) Does it determine I uniquely? How to calculate I? Can we use the transform in order to prove LD-convergence (rather than assume it, as in 8e2)? These questions will be answered later (in Sect. 10).

Now, what about $\{0,1,2\}^n$ (in place of $\{0,1\}^n$)? This case is similar, but leads to matrices 3×3 and a qubic (rather than quadratic) equation for their eigenvalues. Any finite alphabet may be treated this way. Accordingly one can investigate finite Markov chains and nearest-neighbor chains of higher spins.

On the other hand, return to $\{0,1\}^n$ but consider triples $(\beta_1, \beta_2, \beta_3), (\beta_2, \beta_3, \beta_4), \dots$ (rather than pairs $(\beta_1, \beta_2), \dots$). Identifying a triple $(\beta_1, \beta_2, \beta_3)$ with the pair of pairs $((\beta_1, \beta_2), (\beta_2, \beta_3))$ we get a (special) four-state Markov chain. Longer blocks may be treated similarly.

See also [2, Sect. 3.1], [3, Sect. I.5], [1, Sect. V].

8f Dimension two

We turn to two-dimensional arrays $s \in \{-1,1\}^{n \times n}$, $s = (s_{i,j})_{i,j \in \{1,\dots,n\}}$. Blocks of size 2×2 consist of 4 numbers,

$$\begin{pmatrix} s_{i,j} & s_{i,j+1} \\ s_{i+1,j} & s_{i+1,j+1} \end{pmatrix} .$$

Their frequencies belong to $P\big(\{-1,1\}^{2\times 2}\big)$. The corresponding distributions on $P\big(\{-1,1\}^{2\times 2}\big)$ are LD-convergent (I give no proof). Can we calculate the rate function explicitly? I do not know. Probably, not. What kind of function it is? How smooth? Analytic, or not? Convex, or not? I do not know. Physically, it means a two-dimensional array of spins with a general shift-invariant four-spin interaction.

We may restrict ourselves to blocks of sizes 2×1 and 1×2,

$$\begin{pmatrix} s_{i,j} & s_{i,j+1} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} s_{i,j} \\ s_{i+1,j} \end{pmatrix} .$$

These are pairs of nearest neighbours, in other words, edges of the graph $\mathbb{Z}^2$. Treating them equally, we count the number K_{++} of pairs $(+1, +1)$ (both horizontal and vertical); the same for K_{+-}, K_{-+}, K_{--}. (The boundary may be treated in two ways that are equivalent, similarly to 8a4.) The frequencies are $x_{++} = \frac{K_{++}}{2n^2}$, $x_{+-} = \frac{K_{+-}}{2n^2}$, $x_{-+} = \frac{K_{-+}}{2n^2}$, $x_{--} = \frac{K_{--}}{2n^2}$. Still, it is too difficult, to write down the rate function.

Interestingly, the combination

$$H(s) = -\frac{1}{2}(K_{++} + K_{--} - K_{+-} - K_{-+})$$

is tractable. It is well-known as the energy of the two-dimensional Ising model[1] (without external magnetic field). You see, neighbour spins tend to agree.

A very clever two-dimensional counterpart of the linear-algebraic approach (of 8e) was found in 1944 by Lars Onsager.[2] I just formulate his result, with no proof. It gives us the Legendre-Fenchel transform of the rate function I of $x = x_{++} + x_{--} - x_{+-} - x_{-+}$, defined by $\|f\|_{L_{2n^2}(\mu_n)} \to \max(|f|e^{-I})$. Namely,

$$\min_x \Big(I(x) - \frac{1}{2}\beta x \Big) = - \lim_{n\to\infty} \frac{1}{2n^2} \ln\Big(2^{-n^2} \sum_s e^{-\beta H(s)} \Big) =$$
$$= -\frac{1}{4\pi^2} \int_0^\pi \int_0^\pi \ln\big(\cosh^2 \beta - (\cos u + \cos v) \sinh \beta \big) \, du dv \, .$$

Introducing ε by $\sinh \beta = 1 + \varepsilon$ we have $\cosh \beta = 1 + (1+\varepsilon)^2$. The integrand becomes

$$\ln\big(\varepsilon^2 + 2(1+\varepsilon)(\sin^2 \tfrac{u}{2} + \sin^2 \tfrac{v}{2}) \big) \, ;$$

we observe a singularity at $\varepsilon = 0$, $u = 0$, $v = 0$. Still, the integral converges also for $\varepsilon = 0$, that is, at the critical point $\beta = \beta_c = \ln(1 + \sqrt{2})$. However,

[1]Physicists multiply it by a constant J, but anyway, we will consider βH for an arbitrary β.

[2]A Norwegian chemist, and later Nobel laureate.

the integral is not an analytic function of ε (or β). Namely, the function

$$\Lambda(\beta) = -\min_x \Big(I(x) - \frac{1}{2}\beta x \Big)$$

near the critical point β_c satisfies

$$\Lambda(\beta_c + \Delta\beta) - \Lambda(\beta_c) = \frac{\Delta\beta}{2\sqrt{2}} + \frac{1}{2\pi}(\Delta\beta)^2 \big| \ln|\Delta\beta| \big| + O\big((\Delta\beta)^2\big) \, .$$

Accordingly, the (even) rate function I has critical points $\pm x_c$, $x_c = 1/\sqrt{2}$, and near x_c

$$I(x_c + \Delta x) - I(x_c) = \frac{1}{2}\beta_c \Delta x + \frac{\pi}{2} \frac{(\Delta x)^2}{\big| \ln|\Delta x| \big|} \big(1 + o(1)\big) \, .$$

Physically, it means a phase transition. The heat capacity diverges,

$$\frac{\mathrm{d(energy)}}{\mathrm{d(temperature)}} + \infty$$

at the critical temperature.

See also [5, Sect. 9.3].

References

[1] J.A. Bucklev, *Large deviation techniques in decision, simulation, and estimation,* Wiley, 1990.

[2] A. Dembo, O. Zeitouni, *Large deviations techniques and applications,* Jones and Bartlett publ., 1993.

[3] R.S. Ellis, *Entropy, large deviations, and statistical mechanics,* Springer, 1985.

[4] F. den Hollander, *Large deviations,* AMS, 2000.

[5] D. Yoshioka, *Statistical physics,* Springer, 2007.