Stochastic Models — Second HW problem set

Gábor Pete
http://www.math.bme.hu/~gabor

April 30, 2023

The number of dots • is the value of an exercise. **Hand in solutions for 12 points by May 18 in class.** If you have seriously tried to solve some problem, but got stuck, I will be happy to help. Also, if your final solution to a problem has some mistake but has some potential to work, then I will give it back and you can try and correct the mistake.

The first exercise would have been better in the first set, but I forgot to put it there:

▷ **Exercise 1.** A function $f : V \longrightarrow \mathbb{R}$ on the state space of a Markov chain P is called discrete harmonic if it satisfies $Pf(x) = f(x)$ for every $x \in V$. (I.e., one could call it an eigenfunction with eigenvalue 1, but we don't have a Hilbert space of functions here, so I would not call it like that.)

(a)• Show that if P is irreducible on a finite set V, then every harmonic function is constant.

(b)• Let $(X_n)_{n\geq 0}$ be simple random walk on the 3-regular infinite tree, $\mathbb{T}_3$, with Markov operator P. Take any vertex $o \in V(\mathbb{T}_3)$, and let A be one of the three connected components of $\mathbb{T}_3 \setminus \{o\}$. Show that $f(x) := \mathbf{P}_x[\,\exists n_0 : X_n \in A\ \forall n \geq n_0\,]$ is a non-constant bounded harmonic function, where, remember, $\mathbf{P}_x[\,\cdot\,]$ means that $X_0 = x$.

(c)• Consider the lamplighter graph $G = \mathbb{Z}_2 \wr \mathbb{Z}^3$, with the standard 7 generators (six for moving the marker in the city $\mathbb{Z}^3$, one for switching the lamp where the marker is). Give an example of a non-constant bounded harmonic function for SRW on G. (Hint: follow the strategy of the previous part, but with a different notion of "what happens eventually". Namely, note that the marker visits the origin of $\mathbb{Z}^3$ only a finite number of times, hence there is a "final" state of the lamp there.)

Remark: There is an amazing theorem (Avez 1972, Deriennic 1980, Kaimanovich-Vershik 1983): the existence of bounded harmonic functions on a transitive graph is equivalent to the speed of the SRW being positive (linear rate of escape). Thus $\mathbb{Z}_2 \wr \mathbb{Z}^3$ is an example where the graph is amenable, but the speed is positive — which is also not hard to prove directly.

▷ **Exercise 2.**• Let P be a reversible Markov chain on n states; that is, the random walk on a finite graph G with symmetric edge-weights. We have seen that P has eigenvalues $-1 \leq \lambda_n \leq \cdots \leq \lambda_1 = 1$. Show that $\lambda_n > -1$ if and only if every connected component of G is non-bipartite.

▷ **Exercise 3.**• Let P be any reversible finite Markov chain. Let $\bar{P}$ be its 1/2-lazy version: in each step, with probability 1/2 we stay put, while with probability 1/2 we take a step according to P. This is a usual way to get rid of periodicity. Show that the spectrum of $\bar{P}$ is contained in the interval $[0, 1]$.

▷ **Exercise 4.**•• For simple random walk on any finite or infinite d-regular graph, show that after any even number of steps the most likely position is the starting vertex.

▷ **Exercise 5.**• When the New York Times in 1990 reported on 7 riffle shuffles being enough for mixing, they wrote: "By saying that the deck is completely mixed after seven shuffles, Dr. Diaconis and Dr. Bayer mean that every arrangement of the 52 cards is equally likely or that any card is as likely to be in one place as in another." True or false: Let μ be a distribution on S_n such that when $\sigma \in S_n$ is chosen according to μ, we have $\mathbf{P}[\,\sigma(i) = j\,] = 1/n$ for every $i, j \in \{1, \ldots, n\}$. Then μ is uniform on S_n.

▷ **Exercise 6.**•• Let P be a reversible Markov chain on a finite state space V, with reversible distribution π. Recall that the chain is then just the random walk w.r.t. the symmetric edge-weights $c(x, y) := \pi(x)p(x, y)$. There is the following version of the Courant-Fisher-Rayleigh theorem (which you don't have to prove):

$$\lambda_2 = \sup\left\{ \frac{(Pf, f)_\pi}{\|f\|_\pi} : \mathbf{E}_\pi[\,f\,] := \sum_{x \in V} f(x)\pi(x) = 0 \right\}.$$

Using this, show that the **spectral gap** has the following formula:

$$1 - \lambda_2 = \inf\left\{ \frac{\frac{1}{2}\sum_{x,y} \big(f(x) - f(y)\big)^2 c(x, y)}{\mathrm{Var}_\pi[f]} : \mathrm{Var}_\pi[f] := \mathbf{E}_\pi[f^2] - (\mathbf{E}_\pi f)^2 \neq 0 \right\}.$$

Show that the numerator can be written as $\mathbf{E}_{X_0 \sim \pi}\Big[\,\mathrm{Var}\big[f(X_1) \,\big|\, X_0\big]\Big]$. Thus, this formula is the infimum ratio of the local variance to the global one.

▷ **Exercise 7.** Finding good functions in the formula of the previous exercise will give you upper bounds on the spectral gap (hence lower bounds on the relaxation time, see the next exercise) of reversible Markov chains. Using this strategy, show:

- **(a)**• On the cycle C_n, the gap is at most $O(1/n^2)$.
- **(b)**• On the hypercube $\{0, 1\}^k$, the gap is at most $O(1/k)$.
- **(c)**• On the dumbbell graph (two complete graphs K_n joined by a single edge), the gap is at most $O(1/n^2)$.
- **(d)**• What bound can you give on the following lollipop graph: a complete graph K_n, with a length n^2 path emanating from it?

▷ **Exercise 8.**•• Consider a reversible Markov chain P on a finite state space V with reversible distribution π and absolute spectral gap $g_{\mathrm{abs}} := 1 - \max\{|\lambda_2|, |\lambda_n|\}$. This exercise explains why $T_{\mathrm{relax}} = 1/g_{\mathrm{abs}}$ is called the **relaxation time**.

Show that $g_{\mathrm{abs}} > 0$ implies that $\lim_{t\to\infty} P^t f(x) = \mathbf{E}_\pi f$ for all $x \in V$. Moreover,

$$\mathrm{Var}_\pi[P^t f] \leq (1 - g_{\mathrm{abs}})^{2t}\, \mathrm{Var}_\pi[f]\,,$$

with equality at the eigenfunction corresponding to the λ_i giving $g_{\mathrm{abs}} = 1 - |\lambda_i|$. Hence T_{relax} is the time needed to reduce the standard deviation of any function to $1/e$ of its original standard deviation.

▷ **Exercise 9.**• Combining our computations of the spectrum of the cycle C_n and the spectrum of product chains, show that the spectral gap of SRW on the torus $\mathbb{Z}_n^d$ is of order $1/n^2$ for any fixed $d \geq 1$.

We have accepted that the total variation distance between probability measures can be written as

$$d_{\mathrm{TV}}(\mu, \nu) = \min\big\{\mathbf{P}[\,X \neq Y\,] : (X, Y) \text{ is a coupling of } \mu \text{ and } \nu\big\}. \tag{1}$$

Consider now any Markov chain with a unique stationary measure π, and define

$$d(t) := \sup_{x \in V} d_{\mathrm{TV}}\big(p_t(x, \cdot), \pi(\cdot)\big) \quad \text{and} \quad \bar{d}(t) := \sup_{x,y \in V} d_{\mathrm{TV}}\big(p_t(x, \cdot), p_t(y, \cdot)\big)\,.$$

Furthermore, define the **total variation mixing time** by

$$T_{\mathrm{mix}}(\epsilon) := \inf\big\{t : d(t) \leq \epsilon\big\} \qquad \text{and} \qquad T_{\mathrm{mix}} := T_{\mathrm{mix}}(1/4).$$

The following exercise explains why we introduced $\bar{d}(t)$ and why this $1/4$ definition is a good one.

▷ **Exercise 10.**

(a)• Show that $d(t) \le \bar{d}(t) \le 2d(t)$.

(b)• Using (1), show that $\bar{d}(t+s) \le \bar{d}(t)\,\bar{d}(s)$.

(c)• Conclude from the previous two parts that $T_{\rm mix}(2^{-\ell}) \le \ell\, T_{\rm mix}(1/4)$.

▷ **Exercise 11.**•• Consider simple random walk on the dumbbell graph: take two copies of the complete graph K_n, add a loop at each vertex (so that the degrees become n), except at one distinguished vertex in each copy, which will be connected to each other by an edge. Show that $d(1) = 1/2$, but $T_{\rm mix} \ge cn^2$ for some uniform $c > 0$. That is, in the definition of $T_{\rm mix}$, the $1/4$ should not be replaced by $1/2$.

▷ **Exercise 12.**•• Show that $(1 - g_{\rm abs})^t \le 2d(t)$ in any finite reversible Markov chain. Deduce that $T_{\rm relax} \le C T_{\rm mix}$ for some absolute constant $C < \infty$.

▷ **Exercise 13.**•• Consider lazy SRW on the cycle C_n. Using the Central Limit Theorem, show that for any $t > 0$ there exists $\delta_0(t), \delta_1(t) > 0$, such that, for any n, we have $\delta_0(t) < d(tn^2) < 1 - \delta_1(t)$. Moreover, show that one can achieve $\lim_{t\to 0} \delta_0(t) = 1$. This proves the lower bound $T_{\rm mix} \ge cn^2$ for some uniform $c > 0$.

▷ **Exercise 14.**

(a)•• Show that, for any transitive reversible Markov chain with eigenvalues λ_i as usual,

$$4\, d_{\rm TV}\big(p_t(x,\cdot), \pi(\cdot)\big)^2 \le \left\| \frac{p_t(x,\cdot)}{\pi(\cdot)} - 1 \right\|_2^2 = \sum_{i=2}^{n} \lambda_i^{2t}\,.$$

(b)• Deduce that the mixing time of the 1/2-lazy SRW on the cycle C_n is $O(n^2)$.

(c)• Deduce that the mixing time of the 1/2-lazy SRW on the hypercube $\{0,1\}^k$ is at most $(1/2 + o(1))\, k \log k$.

The following three exercises together give a probabilistic proof that the total variation mixing time of the 1/2-lazy random walk $X_0, X_1, \dots$ on the hypercube $\{0,1\}^k$ is $\sim \frac{1}{2} k \log k$. In particular, this improves on the upper bound $(1 + o(1))\, k \log k$ proved in class (on May 4) by coupling.

▷ **Exercise 15.**•• Let Y_t be the number of missing coupons at time t in the coupon collector's problem with k coupons. Show that, for $\alpha \in (0,1)$ fixed,

$$\mathbf{E}\, Y_{\alpha\, k \log k} \sim k^{1-\alpha} \quad \text{and} \quad \mathbb{D}\, Y_{\alpha\, k\log k} = o(k^{1-\alpha}).$$

Using Markov's and Chebyshev's inequalities, deduce that $Y_{\alpha\, k \log k}/\sqrt{k} \to 0$ or ∞ in probability, for $\alpha > 1/2$ and $< 1/2$, respectively.

▷ **Exercise 16.**•• Let $\mathsf{N}(\mu, \sigma^2)$ denote the normal distribution. Show that, for any sequence $\sigma_k \to \sigma \in (0,\infty)$, we have that $d_{\rm TV}\big(\mathsf{N}(0,\sigma^2), \mathsf{N}(\mu_k, \sigma_k^2)\big) \to 0$ or 1, for $\mu_k \to 0$ and $\mu_k \to \infty$, respectively. Using this and the local version of the de Moivre–Laplace theorem, prove that

$$d_{\rm TV}\big(\mathsf{Binom}(k, 1/2),\, \mathsf{Binom}(k - k^\beta, 1/2) + k^\beta\big) \to \begin{cases} 0 & \text{if } \beta < 1/2\,, \\ 1 & \text{if } \beta > 1/2\,. \end{cases}$$

▷ **Exercise 17.**

(a)• For $X_0 = (0,0,\dots,0) \in \{0,1\}^k$, let the distribution of X_t be μ_t. What is it, conditioned on $\|X_t\|_1 = \ell$?

(b)• What is the distribution of $\|Z\|_1$, where Z has distribution π, uniform on $\{0,1\}^k$?

(c)•• Let Y_t be the number of coordinates that have not been rerandomized by time t in X_t. Compare the distribution of $k - \|X_t\|_1$, conditioned on $Y_t \ge y$, to $\mathsf{Binom}(k-y, 1/2) + y$. Deduce from the previous parts and the previous exercises that $d_{\rm TV}\big(\mu_{\alpha\, n \log n}, \pi\big) \to 0$ or 1, for $\alpha > 1/2$ and $< 1/2$, respectively.

The L^∞**- or uniform mixing time** of a Markov chain is usually defined as

$$T_{\mathrm{mix}}^{\infty} := \inf\left\{ t : \sup_{x,y} \left| \frac{p_t(x,y)}{\pi(y)} - 1 \right| < \frac{1}{e} \right\}.$$

▷ **Exercise 18.**$^{\bullet\bullet}$ Using Exercise 15, show that the uniform mixing time of the hypercube $\{0,1\}^k$ is $\sim k \log k$.

▷ **Exercise 19.** This exercise explains why it is hard to construct large expander graphs. A *covering map* $\varphi : G' \longrightarrow G$ between graphs is a surjective graph homomorphism that is locally an isomorphism: denoting by $N_G(v)$ the subgraph induced by $v \in G$ and all its neighbours, we require that each connected component of the subgraph of G' induced by the full inverse image $\varphi^{-1}(N_G(v))$ be isomorphic to $N_G(v)$.

(a)$^\bullet$ If $G' \longrightarrow G$ is a covering map of infinite graphs, then the spectral radii satisfy $\rho(G') \leq \rho(G)$, i.e., the larger graph is more non-amenable. In particular, if G is an infinite k-regular graph, then $\rho(G) \geq \rho(\mathbb{T}_k) = \frac{2\sqrt{k-1}}{k}$. (Hint: use the return probability definition of $\rho(G)$.)

(b)$^\bullet$ If $G' \longrightarrow G$ is a covering map of finite graphs, then $\lambda_2(G') \geq \lambda_2(G)$, i.e., the larger graph is a worse expander. (Hint: eigenfunctions on G can be "lifted" to G'.

▷ **Exercise 20.**$^\bullet$ Consider the configuration model $M_{n,d}$ for a random d-regular multi-graph on n vertices, with nd even. (Given by a uniform random perfect matching on the nd half-edges). Show that if we condition this random graph to have no multiple edges and no self-loops, then we get the uniform distribution on d-regular simple graphs on n vertices.

▷ **Exercise 21.** We considered the following simple model $M_{n,n,d}$ for a random d-regular bipartite (multi-)graph: take d independent uniform random permutations $\pi_i : \{1,\dots,n\} \longrightarrow \{1,\dots,n\}$, then take all the edges $\big\{(j, n+\pi(j)) : j \in \{1,\dots,n\},\ i \in \{1,\dots,d\}\big\}$.

(a)$^\bullet$ Show that the number of multiple edges is tight in n.

(b)$^\bullet$ More generally, show that, for any $k \geq 2$, the number of k-cycles is tight in n.

We claimed in class that the bipartite configuration model $M_{n,n,d}$ is a good expander with probability tending to 1. I was trying to follow Alexander Lubotzky's *Discrete Groups, Expanding Graphs and Invariant Measures* Proposition 1.2.1, but it claims the result only for $d \geq 5$, and the proof is in fact completely wrong. Instead, here is a correct and easier proof from the Levin-Peres-Wilmer book:

▷ **Exercise 22.**

(a)$^\bullet$ Show that the probability that there exists a subset $S \subset \{1,\dots,n\}$ of size $|S| = t \in \{1,\dots,\lfloor n/2 \rfloor\}$ with neighbourhood $|N(S)| \leq \lfloor (1+\delta)t \rfloor$ is at most

$$R(t) = \binom{n}{t} \frac{\binom{n-t}{\lfloor \delta t \rfloor} \binom{\lfloor (1+\delta)t \rfloor}{\lfloor \delta t \rfloor}^{d-1}}{\binom{n}{t}^{d-1}}.$$

This is in fact the same formula that we had in class, just written in a different way.

(b)$^\bullet$ Using the bounds $(n/k)^k \leq \binom{n}{k} \leq (en/k)^k$, show that, for $d \geq 3$, and $\delta > 0$ small enough, independently of n, we have $\sum_{t=1}^{\lfloor n/2 \rfloor} R(t) \to 0$ as $n \to \infty$.