 jhermon@math.ubc.ca

Submit solutions to 3 problems among problems 1 and 4-10. You are allowed to discuss the problems with your colleagues, but you have to write your own solution, and specify in the solution who did you solve/discuss the problem with. Feb 11th.

Example Sheet 1

1. (a) Suppose X is a random walk on a graph of diameter D (that is, the maximal distance between vertices of the graph). Show that $t_{\mathrm{mix}}(\delta) \geq D/2$ for all $\delta < 1/2$.

   (Huge Hint: use $\bar{d}(t) := \sup_{x,y} \|P^t(x,\cdot) - P^t(y,\cdot)\|_{TV}$ rather than $d(t)$.)

   (b) Argue that also the lazy SRW satisfies the above, and that the rate 1 continuous-time random walk satisfies $t_{\mathrm{mix}} \gtrsim D$. (No need to submit solution for this part. Just check you understand why this is the case.)

   (c) We say that a sequence of graphs $G_n = (V_n, E_n)$ is an *expander family* if $\gamma^{(n)} \asymp 1$, where $\gamma^{(n)}$ is the spectral-gap of the transition matrix of SRW on G_n. Deduce that if G_n is a sequence of d-regular graphs for a fixed d (i.e. $\deg(x) = d$ for all $x \in V_n$ for all n) and is an expander family then the lazy SRW and the rate 1 continuous-time SRW satisfy that for all fixed $\delta \in (0,1)$ and all $\varepsilon > 0$ we have for n sufficiently large that

$$(1-\varepsilon)\log_d(\delta n) \leq t_{\mathrm{mix}}(1-\delta) \leq t^{(2)}_{\mathrm{mix}}(1/n^2) \lesssim \log n.$$

   Deduce that there is pre-cutoff.

2. Let X, X' be Poisson(λ, λ') respectively, and let μ, μ' be their respective law. By a coupling argument or otherwise, show that $\|\mu - \mu'\|_{TV} \leq |\lambda' - \lambda|$ for all λ, λ'.

   (Huge HINT: consider a Poisson process either until different times or use Poisson thinning)

3. Consider the coupon collector problem: an urn contains n white balls initially and we sample from the urn with replacement uniformly at random, each time painting the ball black (whatever its colour). Show that the time τ_n until all balls are black is concentrated near $n \log n$, i.e., $\tau_n/(n \log n) \to 1$ in probability. How many balls are black by time $(1/2) n \log n$?

4. (a) (*) Let $\mathbb{P}$ be the Binomial$(n, 1/2)$ distribution and let $\mathbb{Q}$ and μ be the Binomial$(n - m_+, \frac{1}{2})$ and Binomial$(n - m_-, \frac{1}{2})$ distributions (respectively), where $m_+ = o(n^{1/2})$ and $n^{1/2} = o(m_-)$. Show that as $n \to \infty$ we have that

$$\|\mathbb{P} - \mathbb{Q}\|_{TV} \to 0 \qquad \text{and} \qquad \|\mathbb{P} - \mu\|_{TV} \to 1.$$

   (Hint: For simplicity let us assume n is even (to avoid ceiling signs). Verify by direct calculation (by bounding the ratio) that $\max_{i:|i| \leq C\sqrt{n}} \frac{\mathbb{Q}(\frac{n}{2} \pm i)}{\mathbb{P}(\frac{n}{2} \pm i)} = 1 \pm o(1/\sqrt{n})$, for all fixed $C > 0$ (as $n \to \infty$). This may also be deduced by a quantitative version of Stirling's approximation with an explicit error term. Alternatively, this can also be derived from the local CLT (see e.g., Chapter 3 of Rick Durrett's book: probability theory and examples, or Terry Tao's blog. Lastly, there is also a more probabilistic argument for $\|\mathbb{P} - \mathbb{Q}\|_{TV} \to 0$, but it is much harder to come up with it.)

(b) Now consider lazy random walk $(Z_t : t = 0, 1, \ldots)$ on the hypercube $\{0,1\}^n$ started from $(0, \ldots, 0)$ (the mixing time is independent on the initial state - make sure you understand why this is the case!): at each step we select a coordinate uniformly at random, and flip it with probability 1/2. Show that at equilibrium, the number of coordinates equal to 1 follows the Binomial$(n, 1/2)$ distribution. Let $\tau_\pm$ be the time at which all but $m_\pm = n^{1/2}(\log n)^{\mp 1}$ coordinates have been selected. Show that $\mathbb{P}(|\frac{2\tau_\pm}{n \log n} - 1| \geq \epsilon) \to 0$ as $n \to \infty$, for any $\epsilon > 0$.

(c) Let π be the equilibrium distribution. Show that τ_+ is independent of Z_{τ_+} and that

$$\|\mathbb{P}(Z_{\tau_+} = \cdot) - \pi\|_{TV} = \|\text{Binomial}(n, 1/2) - \text{Binomial}(n - m_+, 1/2)\|_{TV} \to 0$$

and that

$$\|\mathbb{P}(Z_k = \cdot) - \pi\|_{TV} \geq \mathbb{P}[\tau_- > k] \min_{i \in (m_-, n]} \|\text{Binomial}(n, 1/2) - \text{Binomial}(n - i, 1/2)\|_{TV} - o(1)$$

$$= (1 - o(1))\mathbb{P}[\tau_- > k] - o(1).$$

(Hint, if $\mu = p\mu_1 + (1-p)\mu_2$ where $p \in (0,1)$ and μ_1,μ_2 and ν are distributions, then $\|\mu - \nu\|_{TV} \geq |p\|\mu_1 - \nu\|_{TV} - (1-p)\|\mu_2 - \nu\|_{TV}|$. Use this with $\nu = \pi$ and with μ being the law of Z_k.)

(d) Deduce that Z exhibits cutoff at time $(1/2)n \log n$. (Make sure you understand why the independence of Z_{τ_+} from τ_+ is crucial for bounding $d(t)$ in terms of $\mathbb{P}(\tau_+ > t)$.)

5. Random transpositions is the Markov chain such that at each step the left hand picks a card uniformly in the deck, the right hand too, and the two cards are switched. (Note that the two cards can be the same, in which case nothing happens). Formally,

$$P(\sigma, \sigma') = \begin{cases} 1/n & \text{if } \sigma = \sigma' \\ 2/n^2 & \text{if } \sigma' = \sigma \cdot (i,j) \text{ for some } 1 \leq i \neq j \leq n \\ 0 & \text{else} \end{cases}$$

Consider the following coupling argument for random transpositions. At each time t, select a label X_t uniformly at random in $[n] = \{1, \ldots, n\}$ and a position Y_t uniformly at random in $[n]$, independently. Then at time t, transpose the card with label X_t together with the card in position Y_t.

(a) Show that this mechanism generates the random transpositions Markov chain.

(b) By coupling show that $t_{\text{mix}} = O(n^2)$ (consider the set of cards which occupy the same positions in the two decks).

(c) Let $\sigma \in S_n$ be a random uniform permutation. Let X denote the number of fixed points of σ, i.e., the numbers of $1 \leq i \leq n$ such that $\sigma(i) = i$. Show that $\mathbb{E}(X) = 1$ and $\text{Var}(X) = 1$.

Deduce that for random transpositions, $t_{\text{mix}} \geq (1/2 - \epsilon)n \log n$ for any $\epsilon > 0$ and n sufficiently large.

6. The Random to Top shuffle (on n cards) is a random walk on S_n in which we have n distinct cards labeled $[n] := \{1, \ldots, n\}$. At each step a random card is picked and is moved to

the top of the deck. As a permutation, if $X_t(i)$ is the label of the card at position i then $X_{t+1} = X_t \circ (U_{t+1}, U_{t+1} - 1, \dots, 2, 1)$ where $U_{t+1} \sim \text{Uniform}(\{1, \dots, n\})$.

(a) Bound the mixing time using a coupling or using a strong stationary time. (Huge Hint: consider picking the same label in both decks.)

(b) Bound the mixing time using the fact that $P_{\text{RT}} = P_{\text{TR}}^T$, where P_{RT} is the transition matrix of the Random to Top shuffle and P_{TR} is that of the Top to Random shuffle. Hint: Show that for each fixed t the two chains have the same $d(t)$.

7. (a) Let P be a transition matrix on state space S with a stationary distribution π. Recall that the time-reversal is given by $P^*(x,y) = \frac{\pi(y)}{\pi(x)} P(y,x)$ for all x, y. Show that

$$\forall\, k \in \mathbb{N}, \qquad \bar{d}(k) = \max_{f: S \to \mathbb{R}: \pi(f) = 0, f \not\equiv 0} \frac{\|(P^*)^k f\|_1}{\|f\|_1},$$

where $\pi(f) := \sum_{x \in S} \pi(x) f(x)$ is the mean w.r.t. π, $\|f\|_1 := \pi(|f|)$ is the 1-norm w.r.t. π and $P^* f(x) := \sum_y P^*(x,y) f(y)$.

(Small hint: $(P^*)^k = (P^k)^*$ and so one can reduce to the case $k = 1$.)

(HUGE HINT: try $f = f_{x,y} := \frac{1_x}{\pi(x)} - \frac{1_y}{\pi(y)}$ and express general f of mean 0 as such a linear combination $f = \sum_{x \neq y} a_{x,y} f_{x,y}$ such that $\|f\|_1 = 2 \sum_{x \neq y} |a_{x,y}| = \sum_{x \neq y} |a_{x,y}| \|f_{x,y}\|_1$. One way of argueing that any f with $\pi(f) = 0$ can be written as such a linear combination is by induction on the size of the support of f, i.e. on $|\{x : f(x) \neq 0\}|$.)

(b) Deduce that $\bar{d}(t+s) \leq \bar{d}(t)\bar{d}(s)$, for all $s, t \in \mathbb{N}$.

(c) Show that P is irreducible and aperiodic iff P^* is. (No need to submit solution to this part. Just check that you understand why this is the case.)

(d) Show that λ is an eigenvalue of P^* iff it is an eigenvalue of P. (Huge hint: Write $D := \text{Diag}(\pi(x))$. Then $DP^*D^{-1} = P^T$. For f such that $P^* f = \lambda f$ consider Df.)

(e) Show that if f is an eigenfunction of P with eigenvalue $\lambda \neq 1$ then $\pi(f) = 0$.

(HINT: P^* is the dual operator of P w.r.t. $\langle \cdot, \cdot \rangle_\pi$ - i.e., $\langle Pg, h \rangle_\pi = \langle g, P^* h \rangle_\pi \; \forall g, h \in \mathbb{R}^S$. Note that $\langle f, 1 \rangle_\pi = \pi(f)$ and that $P^* 1 = 1$.)

(f) Deduce that if $\lambda \neq 1$ is an eigenvalue of P then $|\lambda|^t \leq \bar{d}(t)$ for all t. Hence $|\lambda| \leq 1$. Finaly argue that if P is irreducible and aperiodic then $|\lambda| < 1$.

8. Let P is a transition matrix of an irreducible aperiodic Markov chain with stationary distribution π. The point of this exercise is to show that even when P is not reversible, some useful spectral-decomposition can be done in order to control the rate of decay of L_2 distances from stationarity and of variances in terms of the absolute spectral-gap of PP^*. (However, in contrast with the reversible case, in general the distance from equilibrium cannot be bounded from below using the eigenvalues PP^*.)

(a) Show that $Q := P^* P$ is reversible w.r.t. π.

(b) Assume that $P(x,x) > 0$ for all x, and deduce that Q is aperiodic and irreducible.

(c) Denote the eigenvalues of Q by $1 = \lambda_1(Q) > \lambda_2(Q) \geq \cdots \lambda_{|S|}(Q) > -1$. Let $\lambda_*(Q) := \max\{|\lambda_2(Q)|, |\lambda_{|S|}(Q)|\}$. Show that for any function $f : S \to \mathbb{R}$,

$$\text{Var}_\pi(Pf) \leq \lambda_*(Q) \, \text{Var}_\pi(f),$$

(where $\mathrm{Var}_\pi(g) := \pi[(g-\pi(g))^2]$) and deduce that for all $k \in \mathbb{N}$

$$\mathrm{Var}_\pi(P^k f) \le \lambda_*(Q)^k \,\mathrm{Var}_\pi(f).$$

(d) Show that for every initial distribution μ if we denote $\mu_k := \mu P^k$ (i.e. $\mu_k(y) = \sum_z \mu(z)P^k(z,y)$ - this is the law of the discrete-time chain at time k if the initial distribution is μ) then $\|\mu_k - \pi\|_{2,\pi}^2 = \mathrm{Var}_\pi((P^*)^k f)$, for $f(x) := \frac{\mu(x)}{\pi(x)}$.

(hint: $\|\nu - \pi\|_{2,\pi}^2 := \sum_x \pi(x)(\frac{\nu(x)}{\pi(x)} - 1)^2 = \mathrm{Var}_\pi(\frac{\nu}{\pi})$).

(e) Deduce that $\|\mu_k - \pi\|_{2,\pi}^2 \le \lambda_*(PP^*)^k \|\mu - \pi\|_{2,\pi}^2$.

9. Let $P_t := e^{-t(I-P)}$. The *Poincaré constant* of P is defined as $\gamma := \min_{g:S\to\mathbb{R} \text{ non-constant}} \frac{\mathcal{E}(g,g)}{\mathrm{Var}_\pi(g)}$, where $\mathcal{E}(g,g) := \langle (I-P)g, g\rangle_\pi$ (this is simply the spectral-gap when P is reversible). The point of this exercise is to show that even if P is not reversible, γ dictates the rate of decay of L_2 distances and variances (in the continuous-time setup).

(a) Show that

$$\frac{d}{dt}\mathrm{Var}_\pi(f_t) = -2\mathcal{E}(f_t, f_t),$$

where $f_t := P_t f$. (Hint: $\frac{d}{dt}e^{tQ} = Qe^{tQ}$).

(b) Let $\gamma := \min_{g:S\to\mathbb{R} \text{ non-constant}} \frac{\mathcal{E}(g,g)}{\mathrm{Var}_\pi(g)}$, where $\mathcal{E}(g,g) := \langle (I-P)g, g\rangle_\pi$. Deduce that

$$\forall\, t \ge 0, \qquad \mathrm{Var}_\pi(f_t) \le e^{-2\gamma t}\,\mathrm{Var}_\pi(f).$$

(c) Show that for every initial distribution μ if we denote $\mu_t := \mu P_t$ (i.e. $\mu_t(y) = \sum_z \mu(z)P_t(z,y)$ - this is the law of the continuous-time chain at time t if the initial distribution is μ) then $\|\mu_t - \pi\|_{2,\pi}^2 = \mathrm{Var}_\pi(P_t^* f)$, for $f(x) := \frac{\mu(x)}{\pi(x)}$, where $P_t^* := e^{-t(I-P^*)}$.

(make sure you understand why $\pi(x)P_t^*(x,y) = \pi(y)P_t(y,x)$ for all $x, y \in S$ and all t, and thus $P_t = P_t^*$ for all t iff $P = P^*$.)

(d) Deduce that $\|\mu_t - \pi\|_{2,\pi}^2 \le e^{-2\gamma t}\|\mu - \pi\|_{2,\pi}^2$. (Hint: $\langle (I-P)g, g\rangle_\pi = \langle (I-P^*)g, g\rangle_\pi$.)

10. Consider an irreducible, continuous-time reversible Markov chain. The goal of this exercise is to show that

$$t_{\text{sep}} := \inf\{t : \max_{x,y} 1 - \frac{P_t(x,y)}{\pi(y)} \le 1/4\} \lesssim \min_x \max_y \mathbb{E}_y[T_x]. \tag{1}$$

The same holds in discrete-time if all of the eignevalues of the transition matrix P are non-negative (e.g., if P is lazy, i.e. $P(x,x) \ge 1/2$ for all x).

(a)** Let x, y, z. Let $\tau_z^{(x)}$ and $\tau_z^{(y)}$ be two independent random variables distributed like the hitting time of z started from x and y respectively. Show that for all t we have that

$$1 - P_t(x,y)/\pi(y) \le \mathbb{P}[\tau_z^{(x)} + \tau_z^{(x)} > t].$$

(HUGE Hint: use reversibility to argue that

$$P_t(x,y)/\pi(y) \le \mathbb{E}[(\mathbf{1}_{\{\tau_z^{(x)}+\tau_z^{(y)}\le t\}})P_{t-\tau_z^{(x)}+\tau_z^{(y)}}(z,z)/\pi(z)]$$

and exploit the spectral decomposition to argue that for a reversible chain $P_s(z,z) \geq \pi(z)$ (this holds also in discrete-time if the eigenvalues are non-negative). You may use the fact that (in the continuous-time setup) the law of $\tau_z^{(\cdot)}$ has a density function w.r.t. Lebesgue - make sure you understand why this is the case.) Please consider either the discrete time or continuouos-time setups, but not both.

(b) Deduce (1).