SC9 Probability on Graphs and Lattices: Sheet #4

Due on December 2, 2025 at 12:00pm

Professor C. Goldschmidt and Professor J. Jorritsma

Ray Tsai

Problem 1

Some basic properties of the total variation distance.

Show that, for two probability measures μ, ν on a finite set S:

(a) $|\mu - \nu|_{\mathrm{TV}} = \frac{1}{2}\sum_{x \in S} |\mu(x) - \nu(x)|$;

Proof. Note that for $A \subseteq S$,

$$\mu(A) - \nu(A) = \sum_{x \in A} \mu(x) - \nu(x).$$

Let $A^+ = \{x \in S : \mu(x) - \nu(x) \geq 0\}$ and $A^- = \{x \in S : \mu(x) - \nu(x) < 0\}$. Then,

$$|\mu - \nu|_{\mathrm{TV}} = \max_{A \subseteq S} |\mu(A) - \nu(A)| = \max(\mu(S^+) - \nu(S^+), \nu(S^-) - \mu(S^-)).$$

But then note that

$$\mu(S^+) + \mu(S^-) = 1 = \nu(S^+) + \nu(S^-) \implies \mu(S^+) - \nu(S^+) = \nu(S^-) - \mu(S^-),$$

so

$$\sum_{x \in S} |\mu(x) - \nu(x)| = (\mu(S^+) - \nu(S^+)) + (\nu(S^-) - \mu(S^-)) = 2\max(\mu(S^+) - \nu(S^+), \nu(S^-) - \mu(S^-)).$$

The reuslt now follows. □

(b) $|\mu - \nu|_{\mathrm{TV}} = \sum_{x \in S} \max\{\mu(x) - \nu(x), 0\}$;

Proof. By the proof of (a), we have

$$|\mu - \nu|_{\mathrm{TV}} = \sum_{x \in S^+} \mu(x) - \nu(x) = \sum_{x \in S} \max\{\mu(x) - \nu(x), 0\}.$$

□

(c) the mapping $(\mu, \nu) \mapsto |\mu - \nu|_{\mathrm{TV}}$ is a distance on the set of all probability measures on S;

Proof. It is clear that $|\mu - \mu|_{\mathrm{TV}} = 0$ and $|\mu - \nu|_{\mathrm{TV}} = |\nu - \mu|_{\mathrm{TV}}$. The triangle inequality follows from the fact that

$$\begin{aligned}
|\mu - \gamma|_{\mathrm{TV}} &= \frac{1}{2}\sum_{x \in S} |\mu(x) - \nu(x)| \\
&= \frac{1}{2}\sum_{x \in S} |\mu(x) - \nu(x) + \nu(x) - \gamma(x)| \\
&\leq \frac{1}{2}\sum_{x \in S} |\mu(x) - \gamma(x)| + \frac{1}{2}\sum_{x \in S} |\nu(x) - \gamma(x)| \\
&= |\mu - \gamma|_{\mathrm{TV}} + |\nu - \gamma|_{\mathrm{TV}}.
\end{aligned}$$

□

Problem 1 continued on next page...

(d) there exists a coupling (X^*, Y^*) of μ and ν such that $\mathbb{P}(X^* \neq Y^*) = |\mu - \nu|_{\text{TV}}$.

Proof. Let $p = \sum_{x \in S} \min(\mu(x), \nu(x))$. By (b),

$$1 - p = \sum_{x \in S} \mu - \min(\mu(x), \nu(x)) = \sum_{x \in S} \max(\mu(x) - \nu(x), 0) = |\mu - \nu|_{\text{TV}}.$$

We couple X^* and Y^* as follows:

1. Let C be the bernoulli random variable with parameter p.
2. If $C = 1$, we sample a value Z with distribution

$$\mathbb{P}(Z = x) = \frac{\min(\mu(x), \nu(x))}{p},$$

and set $X^* = Y^* = Z$.

3. If $C = 0$, we sample a value Z_μ, Z_ν with distribution

$$\mathbb{P}(Z_\mu = x) = \frac{\mu(x) - \min(\mu(x), \nu(x))}{1 - p}, \quad \mathbb{P}(Z_\nu = x) = \frac{\nu(x) - \min(\mu(x), \nu(x))}{1 - p}.$$

Evidently,

$$\mathbb{P}(X^* \neq Y^*) = 1 - p = |\mu - \nu|_{\text{TV}}.$$

We also have that

$$\mathbb{P}(X^* = x) = p\mathbb{P}(Z = x) + (1 - p)\mathbb{P}(Z_\mu = x) = \min(\mu(x), \nu(x)) + \nu(x) - \min(\mu(x), \nu(x)) = \mu(x),$$

$$\mathbb{P}(Y^* = x) = p\mathbb{P}(Z = x) + (1 - p)\mathbb{P}(Z_\nu = x) = \min(\mu(x), \nu(x)) + \mu(x) - \min(\mu(x), \nu(x)) = \nu(x).$$

□

Problem 2

ρ is submultiplicative.

Using Q1(d) or otherwise, show that $\rho(t) := \max_{x,y} |\mathbb{P}_x^t - \mathbb{P}_y^t|_{\text{TV}}$ (defined in Section 4.1) satisfies $\rho(t+s) \le \rho(t)\rho(s)$ for all $t, s \in \mathbb{N}$.

Proof. For starting states x, y, there exists a coupling (X_t, Y_t) such that $X_0 = x$, $Y_0 = y$, and $\mathbb{P}(X_t \neq Y_t) = |\mathbb{P}_x^t - \mathbb{P}_y^t|_{\text{TV}}$, by Q1(d). Note that if $X_t = Y_t$, then $X_{t+s} = Y_{t+s}$ and thus

$$|\mathbb{P}_x^{t+s} - \mathbb{P}_y^{t+s}| = \mathbb{P}(X_{t+s} \neq Y_{t+s} \mid X_t = Y_t) = 0$$

On the other hand, Suppose $X_t = u$ and $Y_t = v$ for some $u \neq v$. Then

$$\mathbb{P}(X_{t+s} \neq Y_{t+s} \mid X_t = u, Y_t = v) = |\mathbb{P}_u^s - \mathbb{P}_v^s| \le \rho(s).$$

It now follows that

$$\rho(t+s) = \mathbb{P}(X_t = Y_t) \cdot 0 + \mathbb{P}(X_t \neq Y_t) \cdot \rho(s) = |\mathbb{P}_x^t - \mathbb{P}_y^t|_{\text{TV}} \cdot \rho(s) \le \rho(t)\rho(s).$$

□

Problem 3

Lazy random walk on the cycle.

Consider a lazy random walk on the n-cycle, i.e. a Markov chain on the set $[n]$ with transition probabilities $p_{i,i} = \frac{1}{2}$, $p_{i,j} = \frac{1}{4}\mathbb{1}_{j \equiv i \pm 1 \pmod{n}}$.

(a) Using a coupling and the fact that $\mathbb{E}[T] = k(n-k)$, where T is the time it takes a simple symmetric random walk started at k to hit either 0 or n, show $t_{\text{mix}} \leq n^2$ for this chain.

Proof. Fix two starting vertices x, y with $x \neq y$. Let $(X_t, Y_t)_{t \geq 0}$ be a markov chain with states $\mathbb{Z}/n\mathbb{Z}$. At each step t, we choose one of the following four events with equal probability:

- $X_{t+1} = X_t + 1, \quad Y_{t+1} = Y_t$
- $X_{t+1} = X_t - 1, \quad Y_{t+1} = Y_t$
- $X_{t+1} = X_t, \quad Y_{t+1} = Y_t + 1$
- $X_{t+1} = X_t, \quad Y_{t+1} = Y_t - 1$

Note that the marginal distributions of $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ are lazy random walks on C_n with the given rules, so (X_t, Y_t) is a coupling of $(\mathbb{P}_x^t, \mathbb{P}_y^t)$. Notice that $(D_t)_{t \neq 0} = (X_t - Y_t)_{t \geq 0}$ is a markov chain on $\mathbb{Z}/n\mathbb{Z}$ with transition probabilities $\mathbb{P}(D_{t+1} = k \mid D_t = j) = \frac{1}{4}\mathbb{1}_{j \equiv k \pm 1 \pmod{n}}$. But then D_t is a simple lazy random walk on $\mathbb{Z}/n\mathbb{Z}$. Let $T = \min\{t \geq 0 \mid D_t = 0\}$. Then

$$\max_k \mathbb{E}[T \mid D_0 = k] = \max_k k(n-k) = \frac{n^2}{4}.$$

By the Markov Inequality,

$$\mathbb{P}(X_t \neq Y_t) = \mathbb{P}(t < T) \leq \frac{\mathbb{E}[T]}{t} \leq \frac{n^2}{4t}.$$

It now follows from Lemma 4.3 and Lemma 4.4 that

$$d(t) \leq \rho(t) = \max_{x,y \in \mathbb{Z}/n\mathbb{Z}} |\mathbb{P}_x^t - \mathbb{P}_y^t|_{\text{TV}} \leq \max_{x,y \in \mathbb{Z}/n\mathbb{Z}} \mathbb{P}(X_t \neq Y_t) \leq \frac{n^2}{4t}.$$

Thus, $d(t) \leq 1/4$ when $t \geq n^2$. This completes the proof. $\square$

(b) Using the Central Limit Theorem or otherwise, find also a lower bound for the mixing of order n^2.

Proof. Consider X_t a markov chain on $\mathbb{Z}$ instead of $\mathbb{Z}/n\mathbb{Z}$. By the symmetry of C_n, we may assume X_t starts at 0. Then $X_t = \sum_{i=1}^t \xi_i$, where ξ_i are i.i.d. random variables with $\mathbb{P}(\xi_i = 1) = \mathbb{P}(\xi_i = -1) = \frac{1}{4}$ and $\mathbb{P}(\xi_i = 0) = \frac{1}{2}$. Thus, X_i has mean 0 and variance $\frac{t}{2}$. By the Central Limit Theorem,

$$X_t \approx \mathcal{N}(0, \frac{t}{2}),$$

so X_t has fluctuation of order $\sqrt{t}$. Thus, for X_t to hit n we need $t = \Omega(n^2)$. $\square$

Problem 4

Random 3-colouring of the star graph.

The *star graph* on n vertices has vertex set $\{1, 2, \dots, n\}$ and edge set $\{(1, v), 2 \le v \le n\}$ of size $n-1$. Consider Glauber dynamics for proper 3-colourings of the star graph. (At each step, a vertex v and a colour c are chosen uniformly at random. If no neighbour of v has colour c, vertex v is given colour c. Otherwise, the colour of v stays unchanged.)

Fix a time K, and let A_K be the event that for some time $t < K$, all the degree-1 vertices $2, 3, \dots, n$ have the same colour at some time $t < K$.

Show that for the chain started from equilibrium, the probability of A_K is at most $2^{-(n-2)}K$. Hence or otherwise, find a lower bound for the mixing time which grows exponentially in n.

Proof. We first note that there are $3 \cdot 2^{n-1}$ proper 3-colorings of the star graph, and so $\pi(\sigma) = \frac{1}{3 \cdot 2^{n-1}}$. Let E_t be the event that all the degree-1 vertices $2, 3, \dots, n$ have the same colour at time t. Note that there are there are $3 \cdot 2$ ways to two color the star graph, so

$$\mathbb{P}(E_t) = \frac{3 \cdot 2}{3 \cdot 2^{n-1}} = \frac{1}{2^{n-2}}.$$

Thus by the union bound,

$$\mathbb{P}(A_k) = \mathbb{P}\left(\bigcup_{t<K} E_t\right) \le \sum_{t<K} \mathbb{P}(E_t) = \frac{K}{2^{n-2}}$$

Fix $x \in \Omega$. Say x has center color 1. Notice that the center color of x changes after t steps only if event A_t doeshappens. Thus,

$$\mathbb{P}_x^t(\text{Center} = 1) \ge 1 - \mathbb{P}(A_t) \ge 1 - \frac{t}{2^{n-2}}.$$

But then for $n \ge 3$,

$$d(t) \ge |\mathbb{P}_x^t - \pi|_{\text{TV}} \ge |\mathbb{P}_x^t(\text{Center} = 1) - \pi(\text{Center} = 1)| \ge 1 - \frac{t}{2^{n-2}} - \frac{1}{3} = \frac{2}{3} - \frac{t}{2^{n-2}}.$$

Since $\frac{2}{3} - \frac{t}{2^{n-2}} \le 1/4$ for $t \ge 2^{n-3}$, we have $t_{\text{mix}} \ge 2^{n-3}$. $\square$

Problem 5

Card shuffling: top-to-random and random-to-top.

Consider a standard deck of 52 cards. A top-to-random shuffle of the deck consists in picking the card on top of the deck and re-inserting it in a uniform random position (among 52 positions, ranging from the top to the bottom one in the deck). A random-to-top shuffle consists in picking a uniform random card among the 52 and putting it on top of the deck.

(a) Using a coupling, show that $t_{mix}^{RtT}(1/4) \leq 278$ for the Markov Chain on the set of all 52! possible orderings of the deck whose steps are independent random-to-top shuffles.

Proof. Let S be the set of all 52! possible orderings of the deck. Fix $x \in S$ with $x \neq y$. Let $(X_t, Y_t)_{t \geq 0}$ be a pair markov chains on S such that $X_0 = x$ and $Y_0 = y$. At each step t, we pick a random card C_t from the 52 cards, and we locate C_t in X_t and Y_t and move them to the top. Notice

$$\mathbb{P}(X_t \neq Y_t) \leq \mathbb{P}(\{C_1, \dots, C_t\} \neq \{1, \dots, 52\}) = \mathbb{P}(T > t),$$

where $T = \min\{t \geq 0 \mid \{C_1, \dots, C_t\} = \{1, \dots, 52\}\}$. By Lemma 4.3, Lemma 4.4, and the calculation in the proof of Section 4.2,

$$d(t) \leq \mathbb{P}(T > t(\log 4, 52)) \leq \frac{1}{4}$$

when $t(1/4, 52) = \lceil 52 \log 52 + 52 \log 4 \rceil = \lceil 52 \log 208 \rceil \leq 278$. Thus, $t_{\text{mix}}^{RtT}(1/4) \leq 278$. □

(b) Let $(P^{TtR})_x^t$ and $(P^{RtT})_x^t$ be the distributions of the top-to-random and random-to-top chains started at x, respectively, after t steps; given a permutation $\sigma \in S_{52}$, let $\sigma(x)$ be the deck obtained by applying the permutation σ to x. Show that

$$(P^{TtR})_x^t(\sigma(x)) = (P^{RtT})_x^t(\sigma^{-1}(x))$$

and deduce that $|(P^{TtR})_x^t - \pi|_{\text{TV}} = |(P^{RtT})_x^t - \pi|_{\text{TV}}$, where π is the uniform distribution on S_{52}.

Proof. Let $S_{TtR} = \{(1, i) : i \in [52]\}$ and $S_{RtT} = \{(i, 1) : i \in [52]\}$. P^{TtR} is a measure on the sample space S_{TtR} and P^{RtT} is a measure on the sample space S_{RtT}. Since $S_{RtT} = \{g^{-1} : g \in S_{TtR}\}$, we have

$$P^{TtR}(g) = P^{RtT}(g^{-1}).$$

But then each permutation $\sigma \in S_{52}$ may be written as a product of transpositions of the form $(1, i)$. Thus,

$$(P^{TtR})_x^t(\sigma(x)) = (P^{RtT})_x^t(\sigma^{-1}(x)).$$

But then by Q1(a),

$$\begin{aligned}
|(P^{TtR})_x^t - \pi|_{TV} &= \frac{1}{2} \sum_{\sigma \in S_{52}} \left| (P^{TtR})_x^t(\sigma(x)) - \frac{1}{52!} \right| \\
&= \frac{1}{2} \sum_{\sigma \in S_{52}} \left| (P^{RtT})_x^t(\sigma^{-1}(x)) - \frac{1}{52!} \right| \\
&= \frac{1}{2} \sum_{\sigma \in S_{52}} \left| (P^{RtT})_x^t(\sigma(x)) - \frac{1}{52!} \right| \\
&= |(P^{RtT})_x^t - \pi|_{TV}.
\end{aligned}$$

□

Problem 5 continued on next page. . .

(c) Show that the mixing times of the top-to-random chain and the random-to-top chain are equal and provide an upper bound for the mixing time t_{mix}^{TtR} of the chain of top-to-random shuffles for a deck with n cards.

Proof. Let d^{TtR} and d^{RtT} be the mixing times of the top-to-random chain and the random-to-top chain, respectively. Since $|(P^{TtR})_x^t - \pi|_{TV} = |(P^{RtT})_x^t - \pi|_{TV}$,

$$d^{TtR}(t) = \max_{x \in S_5 2} |(P^{TtR})_x^t - \pi|_{TV} = \max_{x \in S_5 2} |(P^{RtT})_x^t - \pi|_{TV} = d^{RtT}(t),$$

and thus by (a),

$$t_{mix}^{TtR} = t_{mix}^{RtT} \leq 278.$$

□

Problem 6

The greasy ladder.

A frog jumps on a ladder with n rungs. The ground is labelled 0 and the rungs are labelled $1, 2, \ldots, n$. At each time, with probability 1/2 the frog falls back to level 0, and otherwise 1/2 the frog jumps up one level (unless he is already at level n, in which case he remains there).

That is, $p_{i,i+1} = 1/2$ for $0 \le i \le n-1$, $p_{i,0} = 1/2$ for all i, and $p_{n,n} = 1/2$.

(a) Find the mixing time of the chain.

Proof. Let $x, y \in [n]$. Let $(X_t, Y_t)_{t\ge 0}$ be a pair markov chains on $[n]$ such that $X_0 = x$ and $Y_0 = y$ and $x \ne y$. Let $(Z_t)_{t\ge 0}$ be a bernoulli random variable with parameter 1/2. At each step t, if $Z_t = 1$, then $X_{t+1} = \min(X_t + 1, n)$ and $Y_{t+1} = \min(Y_t + 1, n)$. If $Z_t = 0$, then $X_{t+1} = Y_{t+1} = 0$. Then (X_t, Y_t) is a coupling of $(\mathbb{P}_x^t, \mathbb{P}_y^t)$. Notice

$$\mathbb{P}(X_t \ne Y_t) \le \mathbb{P}(Z_0 = Z_1 = \cdots = Z_{t-1} = 1) = (1/2)^t.$$

Thus by Lemma 4.3 and Lemma 4.4,

$$d(t) \le \rho(t) = \max_{x,y\in[n]} |\mathbb{P}_x^t - \mathbb{P}_y^t|_{\mathrm{TV}} \le \max_{x,y\in[n]} \mathbb{P}(X_t \ne Y_t) \le (1/2)^t.$$

Since $(1/2)^t \le 1/4$ for $t \ge 2$, we have $t_{mix} \le 2$. ☐

(b) In Q5, a Markov chain and its time reversal were shown to have the same mixing time. This is true for the class of "random walks on groups" (for which the shuffle in Q5 is an example). However, it is not true in general! Find the transition probabilities of the time-reversal of the frog chain, and find its mixing time.

Proof. We first note that the stationary distribution of the frog chain is

$$\pi(k) = \begin{cases} \frac{1}{2^{k+1}} & \text{if } 0 \le k < n \\ \frac{1}{2^n} & \text{if } k = n \end{cases}.$$

Let $p^*_{i,j}$ denote the transition probability of the time-reversal of the frog chain. Then

$$p^*_{i,j} = \mathbb{P}(X_t = j \mid X_{t+1} = i) = \frac{\mathbb{P}(X_{t+1} = i \mid X_t = j)\mathbb{P}(X_t = j)}{\mathbb{P}(X_{t+1} = i)} = \frac{p_{j,i}\pi(j)}{\pi(i)}.$$

Thus,

$$p^*_{0,j} = \pi(j), \quad p^*_{n,n} = p^*_{n,n-1} = \frac{1}{2}, \quad p^*_{i,i-1} = 1, \text{ for } 0 < i < n.$$

Let $x, y \in [n]$ with $x \ne y$. Let $(X_t, Y_t)_{t\ge 0}$ be a pair markov chains on $[n]$ such that $X_0 = x$ and $Y_0 = y$. Let T be the mixing time of the reverse frog chain $(X_t)_{t\ge 0}$. Since the transition probabilities of the reverse frog chain starting from 0 follows the stationary distribution, the chain the mixed one step after hitting 0 for the first time. Thus

$$d(t) \le \mathbb{P}(T > t) \cdot 1 + \mathbb{P}(T \le t) \cdot 0 = \mathbb{P}(T > t).$$

Notice that once the chain hits k for some $0 < k < n$, then it will hit 0 in exactly k steps. Thus the worst case occurs when $X_0 = n$, which takes $T = X + (n-1) + 1 = X + n$ steps to mix, where X is the steps needed to exit n. Thus by Markov's Inequality,

$$d(t) \le \mathbb{P}(T > t) = \mathbb{P}(X > t - n) \le \frac{\mathbb{E}[X]}{t-n} = \frac{2}{t-n}.$$

Since $d(t) \le 1/4$ for $t \ge n + 8$, we have $t_{mix} \le n + 8$. ☐