505B Midterm 2 Solutions[1]

1. Question 1

Let $S_0, S_1, \ldots$ denote the simple random walk on $\mathbf{Z}$ with $S_0 := 0$. Let n and c be positive integers. Show that:

$$\mathbf{P}\Big(\max_{1\leq j\leq n}|S_j| \geq c\Big) \leq 2\mathbf{P}(|S_n| \geq c).$$

(Hint: is this related at all to the reflection principle?)

Solution. The event $\max_{1\leq j\leq n}|S_j| \geq c$ is equal to the event $\{T_c \leq n\} \cup \{T_{-c} \leq n\}$, where T_c is the hitting time of c. So, using symmetry and translation invariance of the random walk and also Lemma 3.70 in the notes,

$$\begin{aligned}\mathbf{P}\Big(\max_{1\leq j\leq n}|S_j| \geq c\Big) &= \mathbf{P}\Big(\{T_c \leq n\} \cup \{T_{-c} \leq n\}\Big) \leq \mathbf{P}(T_c \leq n) + \mathbf{P}(T_{-c} \leq n)\\ &= 2\mathbf{P}(T_c \leq n) = 2(1 - \mathbf{P}(T_c > n)) = 2(1 - \mathbf{P}_0(T_c > n))\\ &= 2(1 - \mathbf{P}_c(T_0 > n)) = 2(1 - \mathbf{P}_c(-c < S_n \leq c))\\ &= 2\mathbf{P}(S_n \leq -c \text{ or } S_n > c) \leq 2\mathbf{P}(|S_n| \geq c).\end{aligned}$$

2. Question 2

Let $(X_0, X_1, \ldots)$ be the simple random walk on $\mathbf{Z}$. For any $n \geq 0$, define $M_n = X_n^3 - 3nX_n$. Show that $(M_0, M_1, \ldots)$ is a martingale with respect to $(X_0, X_1, \ldots)$

Now, fix $m > 0$ and let T be the first time that the walk hits either 0 or m. Show that, for any $0 < k \leq m$,

$$\mathbf{E}_k(T \,|\, X_T = m) = \frac{m^2 - k^2}{3}.$$

(You are allowed to apply the Optional Stopping Theorem Version 2 without verifying boundedness of the martingale.)

Solution.

$$\begin{aligned}&\mathbf{E}(M_{n+1} - M_n \,|\, X_n = x_n, \ldots, X_0 = x_0, M_0 = m_0)\\ &= \mathbf{E}(([X_{n+1} - X_n] + x_n)^3 - 3(n+1)([X_{n+1} - X_n] + x_n) - x_n^3 + 3nx_n \,|\, X_n = x_n)\\ &= \frac{1}{2}\Big((1 + x_n)^3 - 3(n+1)(1 + x_n) - x_n^3 + 3nx_n\Big)\\ &\qquad\qquad + \frac{1}{2}\Big((-1 + x_n)^3 - 3(n+1)(-1 + x_n) - x_n^3 + 3nx_n\Big)\\ &= \frac{1}{2}\Big((1 + x_n)^3 - x_n^3 - 3x_n + (-1 + x_n)^3 - x_n^3 - 3x_n\Big) = \frac{1}{2}\Big(3x_n^2 + 1 - 3x_n^2 - 1\Big) = 0.\end{aligned}$$

Now, if $X_0 = k$ with $0 \leq k \leq m$, then the Optional Stopping Theorem says

$$k^3 = \mathbf{E}_k X_0^3 = \mathbf{E}_k M_0 = \mathbf{E}_k M_T = \mathbf{E}_k X_T^3 - 3\mathbf{E}_k T X_T.$$

[1]May 17, 2021,

(Note that $\mathbf{P}(T < \infty) = 1$ by Lemma 3.28 in the notes.) Let $p := \mathbf{P}_k(T = m)$. Since $T = 0$ or $T = m$, the Total Expectation Theorem says

$$\begin{aligned}\mathbf{E}_k T X_T &= \mathbf{P}_k(T = m)\mathbf{E}_k(T X_T \mid X_T = m) + \mathbf{P}_k(T = 0)\mathbf{E}_k(T X_T \mid X_T = 0)\\ &= mp\mathbf{E}_k(T \mid X_T = m).\end{aligned}$$

Also, $\mathbf{E}_k X_T^3 = m^3 p$. So, we have

$$k^3 = m^3 p - 3mp\mathbf{E}_k(T \mid X_T = m).$$

From Example 4.25 in the notes, $p = \frac{k}{m}$. So,

$$\mathbf{E}_k(T \mid X_T = m) = \frac{k^3 - m^3 p}{-3mp} = \frac{m^2 k - k^3}{3k} = \frac{m^2 - k^2}{3}.$$

3. Question 3

Let $m \geq 1$ be an integer. Let P be the $m \times m$ transition matrix of a finite (discrete-time) reversible, irreducible Markov chain. Denote the eigenvalues of P as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$. (You can take it as given that the eigenvalues are real, since we verified this in an exercise.) Show:

- $\lambda_i \leq 1$ for all $1 \leq i \leq m$.
- $|\lambda_i| \leq 1$ for all $1 \leq i \leq m$.
- $\lambda_1 > \lambda_2$.
- If additionally P is aperiodic, then $\lambda_m > -1$.

Solution. Suppose $f \in \mathbf{R}^m$ is a right eigenvector of P with eigenvalue λ. Since Ω is finite, there exists $x_0 \in \Omega$ such that $M := \max_{x \in \Omega} f(x) = f(x_0)$. Let $z \in \Omega$ with $P(x_0, z) > 0$, and assume that $f(z) < M$. Then since f is an eigenvector, $Pf = \lambda f$, i.e.

$$\lambda M = \lambda f(x_0) = P(x_0, z)f(z) + \sum_{y \in \Omega \colon y \neq z} P(x_0, y)f(y) < M \sum_{y \in \Omega} P(x_0, y) = M,$$

a contradiction, unless $f(z) = M$ and $\lambda \leq 1$.

Finally, for any $z \in \Omega$, irreducibility of P implies that there is a sequence of points $x_0, x_1, \ldots, x_k = z$ in Ω such that $P(x_i, x_{i+1}) > 0$ for every $0 \leq i < k$. So, by repeating the above argument $k - 1$ times, $M = f(x_0) = f(x_1) = \cdots = f(x_k) = f(z)$. That is, $f(z) = M$ for every $z \in \Omega$ or $\lambda \leq 1$. If f is constant, then $\lambda = 1$. So, in any case $\lambda \leq 1$.

Applying the same argument to the stochastic matrix P^2, we get $\lambda^2 \leq 1$ for all eigenvalues λ of P (since the eigenvalues of P^2 are the squares of the eigenvalues of P). The second item follows.

By Lemma 3.70 in the notes, the eigenspace of the largest eigenvalue is one-dimensional, so the third item $\lambda_1 > \lambda_2$ follows.

Finally, suppose π is the stationary distribution of an irreducible, aperiodic Markov chain and let $g \in \mathbf{R}^m$ be a left eigenvector of P with eigenvalue λ_m. Since the Markov chain is irreducible, g is not a multiple of π. If $\lambda_m = -1$, then $gP^2 = g$, so that $gP^{2k} = g$ for any $k \geq 1$. So, $\lim_{k \to \infty} gP^{2k} = g$, but this violates the convergence theorem. We conclude that $\lambda_m > -1$.

4. Question 4

Let Ω be a finite set. Let π be a probability distribution on Ω. For any $0 < p < \infty$, and for any $f \in \mathbf{R}^{\Omega}$, define

$$\|f\|_{p,\pi} := \Big(\sum_{x\in\Omega} |f(x)|^p \, \pi(x)\Big)^{1/p}, \qquad \mathbf{E}_{\pi} f := \sum_{x\in\Omega} f(x)\pi(x).$$

- Show that $\|f\|_{1,\pi} \leq \|f\|_{2,\pi}$ for any $f \in \mathbf{R}^{\Omega}$.
- Let μ be a probability distribution on Ω. If $\pi(x) > 0$ for all $x \in \Omega$, show that
$$\|\mu - \pi\|_{\mathrm{TV}} \leq \frac{1}{2}\left\|\frac{\mu(\cdot)}{\pi(\cdot)} - 1\right\|_{2,\pi}.$$
- Let $t > 0$. Prove the bound
$$\max_{x\in\Omega}\left\|\frac{H_t(x,\cdot)}{\pi(\cdot)} - 1\right\|_{2,\pi} \leq \sup_{f\in\mathbf{R}^{\Omega}\colon \|f\|_{2,\pi}\leq 1} \max_{x\in\Omega} |[(H_t - \mathbf{E}_{\pi})f](x)|.$$
(Hint: show, for any $f \in \mathbf{R}^{\Omega}$, $\|f\|_{2,\pi} = \sup_{g\in\mathbf{R}^{\Omega}\colon \|g\|_{2,\pi}\leq 1} |\langle f, g\rangle_{\pi}|$.)

Let P be the transition matrix of a finite, irreducible (discrete-time) Markov chain, let π be the unique stationary distribution of the chain, and let $H_t := e^{t(P-I)}$, $t \geq 0$, be the corresponding heat kernel. The mixing time of a continuous-time Markov chain measures how rapidly the Markov chain converges to equilibrium, providing a more quantitative estimate than the convergence theorem provides. For any $\varepsilon > 0$, we define

$$t_{\mathrm{mix}}(\varepsilon) := \inf\Big\{t \geq 0\colon \max_{x\in\Omega}\|H_t(x,\cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq \varepsilon\Big\}.$$

The mixing time of the continuous-time Markov chain is defined to be $t_{\mathrm{mix}}(1/4)$.

- Suppose we have a finite, irreducible, reversible Markov chain with spectral gap $\gamma := 1 - \lambda_2$. Prove that $t_{\mathrm{mix}} \leq \frac{1}{\gamma}\log(2/\sqrt{\min_{y\in\Omega}\pi(y)})$ by first proving
$$\max_{x\in\Omega}\|H_t(x,\cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq \frac{1}{2}\frac{e^{-\gamma t}}{\sqrt{\min_{y\in\Omega}\pi(y)}}, \qquad \forall\, t > 0$$
(Hint: if $g_1, \ldots, g_{|\Omega|} \in \mathbf{R}^{\Omega}$ are an orthonormal basis of eigenfunctions of P, and if $x \in \Omega$ and $e_x \in \mathbf{R}^{\Omega}$ satisfies $e_x(x) := 1$ and $e_x(y) := 0$ for all $y \neq x$, then show that $\pi(x) = \langle e_x, e_x\rangle_{\pi} = (\pi(x))^2 \sum_{j=1}^{|\Omega|} |g_j(x)|^2$.)

Solution. The first item follows from Jensen's inequality. The second item follows from Exercise 3.62 in the notes, since

$$\|\mu - \pi\|_{\mathrm{TV}} = \frac{1}{2}\sum_{x\in\Omega}|\mu(x) - \pi(x)| = \frac{1}{2}\sum_{x\in\Omega}\left|\frac{\mu(x)}{\pi(x)} - 1\right|\pi(x) = \frac{1}{2}\left\|\frac{\mu(\cdot)}{\pi(\cdot)} - 1\right\|_{1,\pi} \overset{(i)}{\leq} \frac{1}{2}\left\|\frac{\mu(\cdot)}{\pi(\cdot)} - 1\right\|_{2,\pi}.$$

For the third item, we have

$$\left\|\frac{H_t(x,\cdot)}{\pi(\cdot)} - 1\right\|_{2,\pi} = \sup_{g\in\mathbf{R}^{\Omega}\colon \|g\|_{2,\pi}\leq 1}\left|\Big\langle\Big[\frac{H_t(x,\cdot)}{\pi(\cdot)} - 1\Big], g\Big\rangle_{\pi}\right| = \sup_{g\in\mathbf{R}^{\Omega}\colon \|g\|_{2,\pi}\leq 1} |H_t g(x) - \mathbf{E}_{\pi} g|.$$

The supremum characterization of the L_2 norm is sometimes called the reverse Hölder inequality. Taking the maximum over x concludes this item.

For the fourth item, we combine the above items to get

$$\max_{x\in\Omega} \|H_t(x,\cdot)-\pi(\cdot)\|_{\mathrm{TV}} \overset{(ii)\wedge(iii)}{\leq} \frac{1}{2} \sup_{g\in\mathbf{R}^\Omega:\ \|g\|_{2,\pi}\leq 1} \max_{x\in\Omega} |[H_t-\mathbf{E}_\pi]g|. \qquad (*)$$

Let $g\in\mathbf{R}^\Omega$ with $\|g\|_{2,\pi}\leq 1$. If $g_1,\ldots,g_{|\Omega|}\in\mathbf{R}^\Omega$ are an orthonormal basis of eigenfunctions of P with eigenvalues $1\geq\lambda_1>\lambda_2\geq\cdots\geq\lambda_{|\Omega|}\geq -1$, we can then write

$$g=\sum_{i=1}^{|\Omega|} c_i g_i,$$

where $\sum_{i=1}^{|\Omega|} c_i^2 = \|g\|_{2,\pi}^2 \leq 1$. Moreover, g_1 is a constant vector, i.e. there exists $a\in\mathbf{R}$ such that $g_1=(a,\ldots,a)$, and $1=\langle g_1,g_1\rangle_\pi=\sum_{x\in\Omega} a^2\pi(x)=a^2$, so we may assume $a=1$, i.e. $g_1=(1,\ldots,1)$, so that $\mathbf{E}_\pi g_1=1$. Also, for any $1\leq j\leq|\Omega|$,

$$H_t g_j = e^{t(P-I)}g_j = e^{-t}\sum_{k=0}^{\infty}\frac{t^kP^k}{k!}g_j = e^{-t}\sum_{k=0}^{\infty}\frac{t^k\lambda_j^k}{k!}g_j = e^{t(\lambda_j-1)}g_j. \qquad (**)$$

So, $[H_t-\mathbf{E}_\pi]g_1=[e^0-1]g_1=0$. By orthonormality, $\forall\ j\geq 2$, $0=\langle g_1,g_j\rangle_\pi=\mathbf{E}_\pi g_j$, so

$$[H_t-\mathbf{E}_\pi]g=\sum_{i=1}^{|\Omega|}c_i[H_t-\mathbf{E}_\pi]g_i=\sum_{i=2}^{|\Omega|}c_iH_tg_i \overset{(**)}{=} \sum_{i=2}^{|\Omega|}c_ie^{t(\lambda_i-1)}g_i.$$

We apply the Cauchy-Schwarz inequality, $\sum_{i=1}^{|\Omega|}c_i^2\leq 1$ and the definition of γ to get

$$|[H_t-\mathbf{E}_\pi]g(x)| \leq (\sum_{i=2}^{|\Omega|}c_i^2e^{2t(\lambda_i-1)})^{1/2}(\sum_{i=2}^{|\Omega|}[g_i(x)]^2)^{1/2} \leq e^{-t\gamma}(\sum_{i=2}^{|\Omega|}[g_i(x)]^2)^{1/2}. \qquad (\ddagger)$$

As suggested in the hint, if $x\in\Omega$ is fixed, then using orthonormality of the basis $g_1,\ldots,g_\Omega$,

$$\begin{aligned}\pi(x)=\langle e_x,e_x\rangle_\pi &= \Big\langle \sum_{i=1}^{|\Omega|}\langle e_x,g_i\rangle_\pi g_i, \sum_{i'=1}^{|\Omega|}\langle e_x,g_{i'}\rangle_\pi g_{i'}, \Big\rangle_\pi \\ &= \Big\langle \sum_{i=1}^{|\Omega|} g_i(x)\cdot g_i, \sum_{i'=1}^{|\Omega|} g_{i'}(x)\cdot g_{i'}\Big\rangle_\pi = (\pi(x))^2\sum_{i=1}^{|\Omega|}|g_i(x)|^2.\end{aligned}$$

That is, $\sum_{i=1}^{|\Omega|}|g_i(x)|^2=1/\pi(x)$. Substituting into $(\ddagger)$ we finally get

$$|[H_t-\mathbf{E}_\pi]g(x)|\leq\frac{e^{-t\gamma}}{\sqrt{\pi(x)}}.$$

Taking the maximum over x and applying $(*)$ concludes the proof.

Finally, to get a bound on the mixing time $t_{\mathrm{mix}}(1/4)$, we need to find $t>0$ such that

$$\max_{x\in\Omega}\|H_t(x,\cdot)-\pi(\cdot)\|_{\mathrm{TV}}\leq 1/4.$$

From our proven inequality, it suffices to find $t>0$ such that

$$\frac{1}{2}\frac{e^{-\gamma t}}{\sqrt{\min_{y\in\Omega}\pi(y)}}\leq 1/4.$$

Solving for t, we get $e^{-\gamma t} \leq (1/2)\sqrt{\min_{y\in\Omega}\pi(y)}$, i.e. $-\gamma t \leq \log((1/2)\sqrt{\min_{y\in\Omega}\pi(y)})$, i.e.

$$t \geq \frac{1}{\gamma}\log(2/\sqrt{\min_{y\in\Omega}\pi(y)}).$$

The mixing time bound follows.

5. Question 5

Let n be a positive integer and denote $i := \sqrt{-1}$. Let $\tau := e^{2\pi i/n}$. Let $\Omega := \{\tau, \tau^2, \tau^3, \ldots, \tau^{n-1}, 1\}$ be the set of n^{th} roots of unity. Let P be the transition matrix such that $P(\omega, \tau\omega) = P(\omega, \tau^{-1}\omega) = 1/2$ for all $\omega \in \Omega$. That is, P is the simple random walk on the cyclic group of n elements.

- Show that the (discrete time) Markov chain with transition matrix P is reversible with stationary distribution $\pi(x) := 1/n$ for all $x \in \Omega$.
- Show that the eigenvalues of P are $\{\cos(2\pi j/n)\}_{j=0}^{n-1}$. Consequently, the spectral gap is $\gamma := 1 - \cos(2\pi/n)$.
- Bound the mixing time of the continuous-time Markov chain corresponding to P, using your result from the previous problem.

Solution. Reversibility is clear. For any $0 \leq j \leq n$, let $f_j \in \mathbf{R}^\Omega$ be the vector $f_j(k) := e^{2\pi ijk/n}$ for all $1 \leq k \leq n$. Then

$$Pf_j(k) = (1/2)f_j(k+1) + (1/2)f_j(k-1) = \frac{1}{2}f_j(k)[e^{2\pi ij/n} + e^{-2\pi ij/n}] = f_j(k)\cos(2\pi j/n).$$

That is $\{f_j\}_{j=0}^{n-1}$ are a set of eigenvectors of P. Taking the real and imaginary parts gives an orthogonal basis of real eigenvectors with the stated eigenvalues.

From the previous problem, we have (for $n \geq 3$)

$$t_{\text{mix}} \leq \frac{1}{\gamma}\log(2/\sqrt{\min_{y\in\Omega}\pi(y)}) \leq \frac{1}{1-\cos(2\pi/n)}\log(2/\sqrt{1/n}) \leq n^2(\log 2 + (1/2)\log n) \leq 2n^2\log n.$$