ASSIGNMENT 4 - SOLUTIONS

PSTAT 160B - SPRING 2022
DUE DATE: MONDAY, MAY 17 AT 11:59PM

Instructions for the homework: Solve all of the homework problems, and submit them on GradeScope. Your reasoning has to be comprehensible and complete.

Homework Problems

Problem 4.1. Consider a CTMC $\{X_t\}$ on $\mathcal{S} = \{1, 2, 3\}$ with generator matrix

$$Q = \begin{pmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ 2 & 1 & -3 \end{pmatrix},$$

and denote the transition function of $\{X_t\}$ by $\{P(t), t \geq 0\}$.

(a) Let $\{N_t\}$ be a Poisson process with rate $\lambda = 3$. Find the smallest value $M \in \mathbb{N}_0$ such that

$$\mathbb{P}(N_2 > M) \leq 0.01.$$

(b) Using Poisson subordination and part (a), find an estimate $\hat{P}(2)$ such that

$$\max_{x,y\in\mathcal{S}} |\hat{P}(2) - P(2)| \leq 0.01.$$

(c) Without explicitly computing the transition function of $\{X_t\}$, compute the limiting distribution of $\{X_t\}$.[1]

(d) If the chain starts at $X_0 = 1$, in the long term, what proportion of time will it spend in state 1?

(e) If the chain starts at $X_0 = 2$, in the long term, what proportion of time will it spend in state 1?

Solution 4.1.

(a) If $N \sim \text{Poisson}(3)$, then $\min\{M \in \mathbb{N}_0 : \mathbb{P}(N > M) \leq 0.01\} = 12$.

(b) If we let

$$P \doteq I + \frac{Q}{3} = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 2/3 & 1/3 & 0 \end{pmatrix},$$

then our estimate is given by, with $M = 12$,

$$\hat{P}(2) = \sum_{n=0}^{12} e^{-2\cdot 3} \frac{(2\cdot 3)^n}{n!} P^n = \begin{pmatrix} 0.30980432 & 0.43365893 & 0.24770926 \\ 0.30946886 & 0.4339944 & 0.24770926 \\ 0.31013974 & 0.43298804 & 0.24804473 \end{pmatrix}$$

[1] **Hint:** is $\{X_t\}$ irreducible?

In this case, the true value of the transition function at time $t = 2$ is

$$P(2) = \exp(Q \cdot 2) = \begin{pmatrix} 0.3125629 & 0.43752097 & 0.24991613 \\ 0.31222744 & 0.43785643 & 0.24991613 \\ 0.31289836 & 0.43685004 & 0.2502516 \end{pmatrix},$$

so we have that

$$|\hat{P}(2) - P(2)| = \begin{pmatrix} 0.00275858 & 0.00386203 & 0.00220687 \\ 0.00275858 & 0.00386203 & 0.00220687 \\ 0.00275862 & 0.003862 & 0.00220687 \end{pmatrix}.$$

(c) Since $\{X_t\}$ is irreducible, its stationary distribution is its limiting distribution, and is the unique π satisfying

$$\pi Q = \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 \end{pmatrix} \begin{pmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ 2 & 1 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} = \mathbf{0}.$$

Solving the above equations yields

$$\pi = \begin{pmatrix} 5/16 & 1/4 & 7/16 \end{pmatrix}.$$

(d) From (c), since π is the limiting distribution, regardless of the state that the chain starts in, it will spend, in the long term, 5/16 of its time in state 1.

(e) From (c), since π is the limiting distribution, regardless of the state that the chain starts in, it will spend, in the long term, 5/16 of its time in state 1.

Problem 4.2. Suppose that $\{X_t\}$ is an irreducible CTMC on $\mathcal{S} = \{1, 2, 3\}$ with stationary distribution $\pi = \begin{pmatrix} 0.3 & 0.5 & 0.2 \end{pmatrix}$.

(a) Compute $\mathbb{E}[X_t^2 | X_0 \sim \pi]$. Recall that we write

$$\mathbb{E}[\cdot | X_0 \sim \pi]$$

to denote that at time $t = 0$, the chain is distributed according to π.

(b) Note that we have not specified the transition function or generator matrix of $\{X_t\}$. Does the quantity

$$\lim_{t \to \infty} \mathbb{E}[X_t^2 | X_0 \sim \pi].$$

depend on the transition function or generator matrix of the chain? Why or why not?

(c) Does the quantity

$$\lim_{t \to \infty} \mathbb{E}[X_t^2 | X_0 = 1].$$

depend on the transition function or generator matrix of the chain? Why or why not?

Solution 4.2.

(a) Since π is the stationary distribution, for all $t \geq 0$ and $x \in \mathcal{S}$, $\mathbb{P}(X_t = x | X_0 \sim \pi) = \pi_x$. Thus,

$$\mathbb{E}[X_t^2 | X_0 \sim \pi] = 1^2 \pi_1 + 2^2 \pi_2 + 3^2 \pi_3 = 4.1.$$

(b) From part (a), we see that

$$\lim_{t \to \infty} \mathbb{E}[X_t^2 | X_0 \sim \pi] = \pi_1 + 4\pi_2 + 9\pi_3 = 4.1.$$

This clearly depends only on the stationary distribution, and not the specific generator or transition function.

(c) Since the chain is irreducible, we know that $\lim_{t\to\infty} \mathbb{P}(X_t = x|X_0 = 1) = \pi_x$, for each $x \in \mathcal{S}$. Furthermore,

$$\mathbb{E}[X_t^2|X_0 = 1] = 1^2\mathbb{P}(X_t = 1|X_0 = 1) + 2^2\mathbb{P}(X_t = 2|X_0 = 1) + 3^2\mathbb{P}(X_t = 3|X_0 = 1),$$

so it follows that

$$\begin{aligned}\lim_{t\to\infty} \mathbb{E}[X_t^2|X_0 = 1] &= \lim_{t\to\infty} \left(1^2\mathbb{P}(X_t = 1|X_0 = 1) + 2^2\mathbb{P}(X_t = 2|X_0 = 1) + 3^2\mathbb{P}(X_t = 3|X_0 = 1)\right) \\ &= 1^2\pi_1 + 2^2\pi_2 + 3^2\pi_3 \\ &= 4.1.\end{aligned}$$

Again, this does depends only on the stationary distribution, and not the specific generator or transition function.

Problem 4.3. Jobs arrive at a computer server according to a Poisson process with a rate of 5 jobs per hour. The server can complete one job at a time, and the time that it takes for each job to be completed follows an Exponential distribution with a mean of 10 minutes. If the computer server is actively completing a job, then any additional jobs that arrive are added to a queue.

(a) On average, in the long term, how many jobs are there in the queue?
(b) Compute the probability, in the long term, that there at least 3 jobs in the queue.

Solution 4.3.

(a) If we let X_t denote the number of jobs in the queue (including the one being completed) at time t, then $\{X_t\}$ is a birth and death process with constant birth rate of $\lambda = \frac{1}{12}$, and constant death rate of $\mu = \frac{1}{10}$. Since $\lambda < \mu$, $\{X_t\}$ has a unique stationary (and limiting) distribution π, which is the geometric distribution on $\{0, 1, \dots\}$ with parameter $p \doteq \frac{\mu-\lambda}{\mu} = \frac{1}{6}$. Thus, in the long term, on average, there are $p^{-1} - 1 = 6 - 1 = 5$ jobs in the queue.
(b) From (a), if we let $X \sim \pi$, where π is the geometric distribution on $\{0, 1, \dots\}$ with parameter $p = 1/6$, the probability that, in the long term, there are at least 3 jobs in the queue is given by

$$\mathbb{P}(X \geq 3) = 1 - \sum_{k=0}^{2} \mathbb{P}(X = k) = 1 - (1-p)^0 p - (1-p)^1 p - (1-p)^2 p = \frac{125}{216}.$$

Problem 4.4. Let $\boldsymbol{X} = (X_1, \dots, X_d)^T \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$ for some $\boldsymbol{\mu} \in \mathbb{R}^d$ and $d \times d$ matrix Σ, and let A be a deterministic $n \times d$ matrix. Note that $A\boldsymbol{X}$ is a (random) vector in $\mathbb{R}^n$.

(a) Fix $\boldsymbol{a} \in \mathbb{R}^n$. What is the probability distribution of $\boldsymbol{a}^T A\boldsymbol{X}$?
(b) For $1 \leq i \leq n$, compute $\mathbb{E}((A\boldsymbol{X})_i)$.
(c) For $1 \leq i, j \leq n$, compute $\mathrm{Cov}((A\boldsymbol{X})_i, (A\boldsymbol{X})_j)$.
(d) Using (a), (b), and (c), determine the probability distribution of $A\boldsymbol{X}$.

Solution 4.4.

(a) For notational convenience, let $\boldsymbol{v}^T \doteq \boldsymbol{a}^T A$, so that,

$$\boldsymbol{v}^T = \begin{pmatrix} v_1 & \dots & v_d \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n a_i A_{i,1} & \dots & \sum_{i=1}^n a_i A_{i,d} \end{pmatrix} \in \mathbb{R}^d,$$

Then, since $\boldsymbol{X}$ follows a multivariate normal distribution, we know that

$$\boldsymbol{v}^T \boldsymbol{X} = \sum_{j=1}^{d} v_j X_j,$$

follows a (univariate) normal distribution. Thus, it suffices to calculate the mean and variance of $\boldsymbol{v}\boldsymbol{X}$. Using the linearity of expectation, we see that the mean is given by

$$\mathbb{E}[\boldsymbol{v}^T \boldsymbol{X}] = \mathbb{E}\left[\sum_{j=1}^{d} v_j X_j\right] = \sum_{j=1}^{d} v_j \mathbb{E}[X_j] = \sum_{j=1}^{d} v_j \mu_j = \boldsymbol{v}^T \boldsymbol{\mu}.$$

Additionally, using the bilinearity of covariance,

$$\begin{aligned}\mathrm{Var}(\boldsymbol{v}^T \boldsymbol{X}) = \mathrm{Cov}(\boldsymbol{v}^T \boldsymbol{X}, \boldsymbol{v}^T \boldsymbol{X}) = \mathrm{Cov}\left(\sum_{i=1}^{d} v_i X_i, \sum_{j=1}^{d} v_j X_j\right) &= \sum_{i=1}^{d}\sum_{j=1}^{d} v_i \mathrm{Cov}(X_i, X_j) v_j \\ &= \sum_{i=1}^{d}\sum_{j=1}^{d} v_i \Sigma_{i,j} v_j = \sum_{i=1}^{d} v_i \sum_{j=1}^{d} \Sigma_{i,j} v_j = \sum_{i=1}^{d} v_i (\Sigma \boldsymbol{v})_i = \boldsymbol{v}^T \Sigma \boldsymbol{v}.\end{aligned}$$

Recalling that $\boldsymbol{v}^T = \boldsymbol{a}^T A$, we have $\boldsymbol{v} = A^T \boldsymbol{a}$, so

$$\boldsymbol{a}^T A \boldsymbol{X} \sim \mathcal{N}\left(\boldsymbol{a}^T A \boldsymbol{\mu}, \boldsymbol{a}^T A \Sigma A^T \boldsymbol{a}\right).$$

(b) If we let $\boldsymbol{a}(i)$ denote the i-th unit vector in $\mathbb{R}^n$, so that

$$\boldsymbol{a}(i)_j = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases},$$

then, $\boldsymbol{a}(i)^T A \boldsymbol{X} = (A\boldsymbol{X})_i$. Thus, we can apply our result from part (a) with this choice of $\boldsymbol{a} = \boldsymbol{a}(i)$ to see that, for $1 \le i \le n$,

$$\mathbb{E}((A\boldsymbol{X})_i) = \mathbb{E}(\boldsymbol{a}(i)^T A \boldsymbol{X}) = \boldsymbol{a}(i)^T A \boldsymbol{\mu} = \sum_{j=1}^{d} A_{i,j} \mu_j. = (A\boldsymbol{\mu})_i.$$

This says that $\mathbb{E}(A\boldsymbol{X}) = A\mathbb{E}(\boldsymbol{X}) = A\boldsymbol{\mu}$, and an analogous calculation shows that $\mathbb{E}(\boldsymbol{X}A^T) = \mathbb{E}(\boldsymbol{X})A^T = \boldsymbol{\mu}A^T$.

(c) Note that

$$\begin{aligned}\mathrm{Cov}((A\boldsymbol{X})_i, (A\boldsymbol{X})_j) &= \mathrm{Cov}\left(\sum_{k=1}^{d} A_{i,k} X_k, \sum_{l=1}^{d} A_{j,l} X_l\right) \\ &= \sum_{k=1}^{d}\sum_{l=1}^{d} A_{i,k} \mathrm{Cov}(X_k, X_l) A_{j,l} \\ &= \sum_{k=1}^{d} A_{i,k} \sum_{l=1}^{d} \Sigma_{k,l} A_{j,l} \\ &= \sum_{k=1}^{d} A_{i,k} (\Sigma A^T)_{k,j} \\ &= (A\Sigma A^T)_{i,j}.\end{aligned}$$

Note that this says that the covariance matrix of $A\boldsymbol{X}$ is given by $\mathrm{Cov}(A\boldsymbol{X}) = A\Sigma A^T$.

(d) From (a), (b), and (c), we see that $A\boldsymbol{X} \sim \mathcal{N}(A\boldsymbol{\mu}, A\Sigma A^T)$.

Problem 4.5. Let $X, Y \overset{iid}{\sim} \mathcal{N}(0,1)$, and let W be a random variable with distribution

$$\mathbb{P}(W=1) = \mathbb{P}(W=-1) = \frac{1}{2},$$

that is independent of X and Y. Define $Z = WX + Y$. Determine whether (X, Z) is jointly normal.
Solution 4.5. To show that (X, Z) is not jointly normal, it suffices to show that $X+Z$ does not follow a normal distribution. To determine the probability distribution of $X+Z$, we calculate its mgf:

$$m(t) \doteq \mathbb{E}[e^{t(X+Z)}] = \mathbb{E}[e^{t(X+WX+Y)}] = \mathbb{E}[e^{t(1+W)X}e^{tY}] = \mathbb{E}[e^{t(1+W)X}]\mathbb{E}[e^{tY}].$$

If we let $\bar{m}(t) \doteq \mathbb{E}[e^{t(1+W)X}]$ and $\tilde{m}(t) \doteq \mathbb{E}[e^{tY}]$, then, recalling the moment generating function of the standard normal distribution, we know

$$\tilde{m}(t) = \exp\left(\frac{t^2}{2}\right),$$

and, from class (or using the law of total expectation), we know that

$$\bar{m}(t) = \mathbb{E}\left[\mathbb{E}[e^{t(1+W)X}|W]\right] = \frac{1}{2}\left(1 + \exp\left(\frac{4t^2}{2}\right)\right),$$

so

$$m(t) = \bar{m}(t)\tilde{m}(t) = \frac{1}{2}\exp\left(\frac{t^2}{2}\right)\left(1 + \exp\left(\frac{4t^2}{2}\right)\right).$$

Noting that $U \sim \mathcal{N}(\mu, \sigma^2)$ if and only if

$$\mathbb{E}[e^{tU}] = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right),$$

we see that $m(t)$ is not the mgf of a normally distributed random variable, meaning that $X+Z$ does not follow a normal distribution. It follows that (X, Z) is not jointly normal.

Problem 4.6. Let $\{S_n\}$ be a discrete-time random walk of the form

$$S_n = \sum_{i=1}^{n} X_i,$$

where $\{X_i\}$ are iid random variables with distribution

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}.$$

(a) Calculate [2]

$$\mathbb{P}\left(\frac{S_{35}}{\sqrt{50}} \geq 1\right).$$

(b) Using the central limit theorem, estimate the probability from part (a).
Solution 4.6.

[2]**Hint:** it may be helpful to review random walks from PSTAT 160A.

(a) We have

$$\begin{aligned}
\mathbb{P}\left(\frac{S_{35}}{\sqrt{50}} \geq 1\right) &= \mathbb{P}(S_{35} \geq \sqrt{50}) \\
&= \mathbb{P}(S_{35} \geq 8) \\
&= \sum_{i=0}^{13} \mathbb{P}(S_{35} = 9 + 2i) \\
&= \sum_{i=0}^{13} \binom{35}{9+2i} (1/2)^{35-9+2i} (1/2)^{9+2i} \\
&\approx 0.0877.
\end{aligned}$$

(b) The central limit theorem says that $\frac{S_{35}}{\sqrt{35}} \approx \mathcal{N}(0,1)$, so if we let $Z \sim \mathcal{N}(0,1)$, then

$$\mathbb{P}\left(\frac{S_{35}}{\sqrt{50}} \geq 1\right) = \mathbb{P}\left(\frac{S_{35}}{\sqrt{35}} \geq \frac{\sqrt{50}}{\sqrt{35}}\right) \approx \mathbb{P}(Z \geq 1.195) = 0.116.$$