S&DS 351: Stochastic Processes - Homework 1

Bryan SebaRaj

Professor Ilias Zadik

January 22, 2025

Problem 1.1

Suppose you have a matrix $X \in \mathbb{R}^{m\times n}$ and another matrix $Y \in \mathbb{R}^{n\times p}$. Let $Z = X \times Y$, i.e., the matrix multiplication of X and Y.

(a) (5 points) What are the dimensions of Z? What is the i, jth entry of Z in terms of those of the matrices X and Y? Is Z necessarily equal to $Y \times X$? If not, provide a counterexample.

The dimensions of $Z \in \mathbb{R}^{m\times p}$.
The i, jth entry of Z is given by

$$Z_{i,j} = \sum_{k=1}^{n} A_{i,k} B_{k,j}$$

Z is not necessarily equal to $Y \times X$. For example, consider the following matrices:

$$X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad Y = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

Then, $Z = X \times Y = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$, but $Y \times X = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix}$.

Therefore, there $\exists X, Y$ such that $X \times Y = Z \neq Y \times X$.

(b) (5 points) Consider the following matrix P:

$$P = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Find P^2 (that is, $P \times P$).

$$P^2 = P \times P = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{9} & \frac{5}{18} & \frac{11}{18} \\ \frac{1}{6} & \frac{5}{12} & \frac{5}{12} \end{bmatrix}$$

(c) (5 points) Find the limit of P^n as $n \to \infty$ (that is, find the limit of each entry $(P^n)_{i,j}$, $1 \le i, j \le 3$ as $n \to \infty$). You do not need to prove what the limit is; it suffices to guess correctly (using a calculator or computer is allowed).

Since P is row-stochastic,

$$\lim_{x\to\infty} P^n = \begin{bmatrix} \pi \\ \pi \\ \pi \end{bmatrix} = \mathbf{1}\pi^\top$$

where π is the stationary distribution statisfying $\pi P = \pi$ and $\pi_1 + \pi_2 + \pi_3 = 1$.
Let $\pi = (\pi_1, \pi_2, \pi_3)$.

$$(\pi P)_j = \sum_{i=1}^{3} \pi_i P_{i,j} = \pi_j$$

for each $j = 1, 2, 3$
For $j = 1$, $\pi_1 = 0\pi_1 + \frac{1}{3}\pi_2 + 0\pi_3 = \frac{1}{3}\pi_2$
So, $\pi_1 = \frac{1}{3}\pi_2$
For $j = 2$, $\pi_2 = 0\pi_1 + \frac{1}{3}\pi_2 + \frac{1}{2}\pi_3 = \frac{1}{3}\pi_2 + \frac{1}{2}\pi_3$, or $\pi_3 = \frac{4}{3}\pi_2$
For $j = 3$, $\pi_3 = \pi_1 + \frac{1}{3}\pi_2 + \frac{1}{2}\pi_3 = \pi_1 + \frac{1}{3}\pi_2 + \frac{1}{2}\pi_3$
Normalizing using $\pi_1 + \pi_2 + \pi_3 = 1$,

$$\pi_1 + \pi_2 + \pi_3 = \frac{1}{3}\pi_2 + \pi_2 + \frac{4}{3}\pi_2 = \frac{8}{3}\pi_2 = 1$$

Therefore, $\pi_2 = \frac{3}{8}$, $\pi_1 = \frac{1}{3}\pi_2 = \frac{1}{8}$, and $\pi_3 = \frac{4}{3}\pi_2 = \frac{1}{2}$.

$$\pi = \left(\frac{1}{8}, \frac{3}{8}, \frac{1}{2}\right)$$

In a 3×3 matrix, where every row is π, we get

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \\ \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \\ \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \end{bmatrix}$$

(d) (Bonus, 10 points) Prove the following statement for any $P \in \mathbb{R}^{3\times3}$. Assume the limit of P^n as $n \to \infty$ equals a matrix of the form $\mathbf{1}\pi^\top$ for some $\pi \in \mathbb{R}^{3\times1}$ and $\mathbf{1} = (1, 1, 1)^\top \in \mathbb{R}^{3\times1}$. Confirm that $\mathbf{1}^\top\pi \in \mathbb{R}^{3\times3}$. Prove that $P^\top\pi = \pi$.

Problem 1.2

Suppose that we are given two geometric random variables A_1 and A_2 with parameter p which are not necessarily independent. Let $\{B_1, B_2, \dots\}$ be a sequence of random variables independent of A_1 and A_2, such that each B_i has mean μ and variance σ^2.

(a) (5 points) Compute $\mathbb{E}[A_1 + 300A_2]$.

Given that A_1 and A_2 are geometric, even if they are not independet, the expectation of a sum of random variables is the sum of their expectations. As such

$$\mathbb{E}[A_1 + 300A_2] = \mathbb{E}[A_1] + 300\mathbb{E}[A_2] = \frac{1}{p} + \frac{300}{p} = \frac{301}{p}$$

(b) (5 points) Prove that $\mathbb{P}[A_1 + 300A_2 \geq 5000/p] \leq 0.1$.

Employing Markov's inequality, where for any non-negative random variable X and any $a > 0$,

$$\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$$

$$\mathbb{P}\left(A_1 + 300A_2 \geq \frac{5000}{p}\right) \leq \frac{\mathbb{E}[A_1 + 300A_2]}{\frac{5000}{p}} = \frac{\frac{301}{p}}{\frac{5000}{p}} = \frac{301}{5000}$$

Since $\frac{301}{5000} < 0.1$, $\mathbb{P}[A_1 + 300A_2 \geq 5000/p] \leq \frac{301}{5000} < 0.1$, proving the inequality.

(c) (10 points) Compute $\mathbb{E}[\sum_{i=1}^{A_1} B_i^2]$. (Hint: condition on A_1).

From the given information,

$$\mathbb{E}[B_i^2] = Var(B_i) + (\mathbb{E}[B_i])^2 = \sigma^2 + \mu^2$$

Using the law of total expectation and conditioning on A_1,

$$\mathbb{E}[\sum_{i=1}^{A_1} B_i^2] = \mathbb{E}[\mathbb{E}[\sum_{i=1}^{A_1} B_i^2 | A_1]] = \mathbb{E}[A_1(\sigma^2 + \mu^2)] = (\sigma^2 + \mu^2)\mathbb{E}[A_1]$$

Since $A_1 \sim \text{Geometric}(p)$,

$$\mathbb{E}[A_1] = \frac{1}{p}$$

Therefore,

$$\mathbb{E}\left[\sum_{i=1}^{A_1} B_i^2\right] = \frac{\sigma^2 + \mu^2}{p}$$

Problem 1.3

Suppose that two teams play a best of 5 series. That is, whichever team wins 3 games is the winner of the series. Suppose that each game is played independently, and for each game team A has a probability 0.7 of winning and team B has a probability 0.3.

(a) (5 points) What is the probability that team A wins the series?

The probability that A wins the best of 5 series can be denoted by $X \sim \text{Binomial}(n = 5, p - 0.7)$ Therefore,

$$\mathbb{P}(X \geq 3) = \sum_{k=3}^{5} \binom{5}{k} 0.7^k 0.3^{5-k}$$

$$\mathbb{P}(X \geq 3) = \binom{5}{3}(0.7)^3(0.3)^2 + \binom{5}{4}(0.7)^4(0.3) + \binom{5}{5}(0.7)^5$$

$$\mathbb{P}(X \geq 3) = 10(0.7)^3(0.3)^2 + 5(0.7)^4(0.3) + (0.7)^5$$

$$\mathbb{P}(X \geq 3) \approx 0.8369$$

(b) (5 points) What is the probability that team A wins the series conditioned on the fact that team B won the first game? If you had to bet, would you bet on A winning the series? Would you still bet on A winning after B won the first game?

Given that B won the first game, the series is now a best of 4 series, where A needs to win 3 games, from the perspective of team A. Let $Y \sim \text{Binomial}(n = 4, p = 0.7)$. Therefore,

$$\mathbb{P}(Y \geq 3) = \sum_{k=3}^{4} \binom{4}{k} (0.7)^k (0.3)^{4-k}$$

Therefore, $\mathbb{P}(A \text{ winning the series } | B \text{ won the first game}) = \mathbb{P}(Y \geq 3) = \binom{4}{3}(0.7)^3(0.3) + \binom{4}{4}(0.7)^4 = 0.6517$

Given those probabilities, I would bet on A winning the series, regardless if B wins the first game.

Problem 1.4

Let X, Y be two *independent* standard normal random variables. Let R be an exponential random variable with parameter 1 and let Θ be a uniform random variable taking values between $[0, 2\pi]$.

(a) (5 points) Compute $\mathbb{P}(R = 0)$ (please note that this isn't the PDF of R at 0, we are asking what is the probability that R equals 0).

Given that R is an exponential random variable with parameter 1, its distribution is continuous on $[0, \infty)$. A fundamental property of continuous distributions is that they place 0 probability mass on a single point, or

$$\mathbb{P}(R = 0) = \int_0^\infty f_R(r) \mathbf{1}r = 0 dr = 0$$

where $f_R(r) = e^{-r}\ \forall r \geq 0$. Therefore, $\mathbb{P}(R = 0) = 0$

(b) (10 points) Compute $\mathbb{P}(X^2 + Y^2 \geq t)$.

Since X, Y are independent $N(0,1)$, $X^2 + Y^2$ follows a χ^2 distribution with 2 degrees of freedom,

$$X^2 + Y^2 \sim \chi_2^2$$

which is an exponential distribution with rate $\frac{1}{2}$

$$X^2 + Y^2 \sim \text{Exp}\left(\frac{1}{2}\right)$$

Specifically,

$$\mathbb{P}(X^2 + Y^2 \leq r) = 1 - e^{-\frac{r}{2}}, \forall r \geq 0$$

Therefore,

$$\mathbb{P}(X^2 + Y^2 \geq t) = 1 - [1 - e^{-\frac{t}{2}}] = e^{-\frac{t}{2}}$$
$$\mathbb{P}(X^2 + Y^2 \geq t) = e^{-\frac{t}{2}}, \ t \geq 0$$

(c) (10 points) Assume that R and Θ are independent. Define $A = \sqrt{R}\cos(\Theta)$ and $B = \sqrt{R}\sin(\Theta)$, what is the joint PDF of A, B? What is the marginal PDF of A?

Calculating the joint PDF of A, B, we can first invert the transformation from

$$A = \sqrt{R}\text{cos}\theta \text{ and } B = \sqrt{R}\text{sin}\theta$$

to yield R and Θ in terms of A and B:

$$R = A^2 + B^2 \text{ and } \theta = \arctan(B, A)$$

This yields a bijection from $(r, \theta) \in [0, \infty) \times [0, 2\pi]$ to $(a, b) \in \mathbb{R}^2$ Computing the Jacobian of the transformation using the generation formula for transformation of PDFs:

$$f_{A.B}(a, b) = f_{R,\Theta}(r, \theta) \cdot \left|\text{det}\frac{\partial(r, \theta)}{\partial(a, b)}\right|$$

First, note that

$$f_{R,\Theta}(r, \theta) = f_R(r) f_\Theta(\theta) = (e^{-r}\mathbf{1}_{r \geq 0})(\frac{1}{2\pi}\mathbf{1}_{[0,2\pi]}\theta) = \frac{1}{2\pi}e^{-r}, (r \geq 0, 0 \leq \theta \leq 2\pi)$$

Computing the forward Jaconbian,

$$\frac{\partial a}{\partial r} = \frac{1}{2\sqrt{r}}\text{cos}\theta$$
$$\frac{\partial a}{\partial \theta} = -\sqrt{r}\text{sin}\theta$$
$$\frac{\partial b}{\partial r} = \frac{1}{2\sqrt{r}}\text{sin}\theta$$
$$\frac{\partial b}{\partial \theta} = \sqrt{r}\text{cos}\theta$$

Hence, det

$$\frac{\partial(a, b)}{\partial(r, \theta)} = \frac{1}{2\sqrt{r}}\text{cos}\theta \cdot \sqrt{r}\text{cos}\theta - (-\sqrt{r}\text{sin}\theta \cdot \frac{1}{2\sqrt{r}}\text{sin}\theta) = \frac{1}{2}\text{cos}^2\theta + \frac{1}{2}\text{sin}^2\theta = \frac{1}{2}$$

Therefore,

$$\text{det}\frac{\partial(r, \theta)}{\partial(a, b)} = \frac{1}{\frac{1}{2}} = 2$$

Substituting r, θ and the Jacobian factor into $f_{R,\Theta}(r\theta)$:

$$f_{A,B}(abs) = f_{R,\Theta}(r, \theta)(a^2 + b^2, \ \arctan(b, a)) \times 2$$

Since

Problem 1.5

Suppose that $X \sim \text{Exp}(\lambda_1)$, $Y \sim \text{Exp}(\lambda_2)$, and Y is independent of X.

(a) (5 points) Compute $\mathbb{P}(X > Y)$.

From the suppositions above,

$$f_X(x) = \lambda_1 e^{-\lambda_1 x},\ x \geq 0$$

and $\mathbb{P}(X > t) = e^{-\lambda_1 t}$ and

$$f_Y(y) = \lambda_2 e^{-\lambda_2 y},\ y \geq 0$$

and $\mathbb{P}(Y > t) = e^{-\lambda_2 t}$.

Y being independent of X implies that

$$f_{X,Y}(x,y) = f_X(x)f_Y(y), \forall x, y \geq 0$$

$$\mathbb{P}(X > Y) = \int_0^\infty \int_0^\infty \mathbf{1}_{x>y} f_{X,Y}(x,y)dx\ dy$$

$$\mathbb{P}(X > Y) = \int_0^\infty \int_{x=y}^\infty \lambda_1 e^{-\lambda_1 x} \lambda_2 e^{-\lambda_2 y} dx\ dy = \int_0^\infty e^{-\lambda_1 y} \lambda_2 e^{-\lambda_2 y}\ dy$$

$$\mathbb{P})(X > Y) = \lambda_2 \int_0^\infty x^{-(\lambda_1+\lambda_2)y} dy$$

$$\mathbb{P}(X > Y) = \lambda_2 \left(- \left. \frac{e^{-(\lambda_1+\lambda_2)y}}{\lambda_1 + \lambda_2} \right|_0^\infty \right)$$

$$\mathbb{P}(X > Y) = \frac{\lambda_2}{\lambda_1 + \lambda_2}$$

(b) (5 points) Compute $\mathbb{P}(X > (t + x) \mid X > t)$, for $t > 0$ and $x > 0$.

From the definition of conditional probability,

$$\mathbb{P}(X > (t + x) \mid X > t) = \frac{\mathbb{P}(X > (t + x) \cap X > t)}{\mathbb{P}(X > t)}$$

Since $x > 0$, $X > t + x$ is contained in $X > t$, so $\mathbb{P}(X > (t + x) \cap X > t) = \mathbb{P}(X > t + x)$

Therefore,

$$\mathbb{P}(X > (t + x) \mid X > t) = \frac{\mathbb{P}(X > t + x)}{\mathbb{P}(X > t)}$$

Since $X \sim \text{Exp}(\lambda_1)$,

$$\mathbb{P}(X > t + x | X > t) = e^{\lambda_1(t+x)},\ \text{and}\ \mathbb{P}(X > t) = e^{-\lambda_1 t}$$

Therefore,

$$\mathbb{P}(X > t + x | X > t) = \frac{e^{-\lambda_1(t+x)}}{e^{-\lambda_1 t}} = e^{-\lambda_1 x}$$

(c) (5 points) Compute $\mathbb{P}(\min(X, Y) > t)$.

$min(X, Y) > t$ can be rewritten as $X > t \cup Y > t$.

Therefore, given that X and Y are independent,

$$\mathbb{P}(\min(X, Y) > t) = \mathbb{P}(X > t)\mathbb{P}(Y > t)$$

Since $X \sim \text{Exp}(\lambda_1)$ and $Y \sim \text{Exp}(\lambda_2)$,

$$\mathbb{P}(X > t) = e^{-\lambda_1 t}\ \text{and}\ \mathbb{P}(Y > t) = e^{-\lambda_2 t}$$

Therefore,

$$\mathbb{P}(\min(X, Y) > t) = e^{-\lambda_1 t} e^{-\lambda_2 t} = e^{-(\lambda_1+\lambda_2)t}$$

Problem 1.6

Given a fair die with 8 possible sides, let T be the number of times you have to roll so that all eight sides have appeared at least once. Let N be the number of distinct sides obtained from the first eight rolls.

(a) (5 points) Find $\mathbb{E}(T)$.

Let t_i be the number of die rolls required to obtain the ith distinct side, where $i = 1, 2, ..., 8$, after $i - 1$ distinct rolls have been observed.

$$T = t_1 + t_2 + ... + t_n$$

Note that the probability of rolling a distinct side on the ith roll is $\frac{n-i+1}{n}$. Therefore, t_i has a geometric distribution with expectation

$$\frac{1}{p_i} = \mathbb{E}[t_i] = \frac{n}{n-i+1}$$

By linearity of expectations,

$$\mathbb{E}[T] = \mathbb{E}[t_1 + t_2 + ... + t_n]$$
$$\mathbb{E}[T] = \mathbb{E}[t_1] + \mathbb{E}[t_2] + ... + \mathbb{E}[t_n]$$
$$\mathbb{E}[T] = \frac{n}{n} + \frac{n}{n-1} + ... \frac{n}{1}$$
$$\mathbb{E}[T] = 8\left(\frac{1}{8} + \frac{1}{7} + \frac{1}{6} + \frac{1}{5} + \frac{1}{4} + \frac{1}{3} + \frac{1}{2} + 1\right)$$
$$\mathbb{E}[T] \approx 21.743$$

(b) (5 points) Find $\mathbb{E}(N)$.

For $1 \leq i \leq 8$, let X_i be the indicator random variable which equals 1 if ith distinct side appears at least once in the first eight rolls, and 0 otherwise.

The total number of distinct sides obtained is then

$$N = \sum_{i=1}^{8} X_i$$

By linearity of expectations,

$$\mathbb{E}[N] = \mathbb{E}[\sum_{i=1}^{8} X_i] = \sum_{i=1}^{8} \mathbb{E}[X_i]$$

Since X_i is an indicator random variable, $\mathbb{E}[X_i] = \mathbb{P}(X_i = 1)$, which is the probability that the ith distinct side appears at least once in the first eight rolls.

$$\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(\text{face } i \text{ does not appear in the first eight rolls})$$
$$\mathbb{P}(X_i = 1) = 1 - \left(\frac{7}{8}\right)^8$$

Therefore,

$$\mathbb{E}[N] = 8\left[1 - \left(\frac{7}{8}\right)^8\right]$$
$$\mathbb{E}[N] \approx 5.251$$

(c) (5 points) Find $\mathbb{E}(T \mid N = 4)$.

T can be expressed as $T = 8 + (T - 8)$. Let T' be the number of rolls required to see all missing, distinct sides, starting after roll 8 (i.e. on roll 9).

$$T' = T - 8$$

$$\mathbb{E}[T|N = 4] = \mathbb{E}[8 + (T - 8)|N = 4] = 8 + \mathbb{E}[T'|N = 4]$$

(Note: I'm going to attempt to define this as a Markov chain, but the same logic from part (a) can also be applied).

Let S be a set of relevant states, where $S = \{0, 1, 2, 3, 4\}$, where state i represents the number of missing, unobserved sides (e.g. state 0 represents the observation of all 8 distinct sides).

Let the Markov chain start in state 4 (since $8 - N = 8 - 4 = 4$) and let p_{ij} be the probability of transitioning from state i to state j, such that $j = i - 1$.

From any state i, $p(i, j) = \frac{i}{8}$. Therefore, from state i where $1 \leq i \leq 4$:

$$\mathbb{P}(i, j) = \frac{i}{8} \text{ and } \mathbb{P}(i, i) = 1 - \frac{i}{8}$$

and for state 0, we remain in state 0 with probability 1, i.e. state 0 is the absorbing state.

Let T_i be the number of rolls required to see all missing, distinct sides, starting from state i to state 0. From the markov chain, we can define the standard equation:

$$T_i = 1 + \frac{i}{8}T_{i-1} + \left(1 - \frac{i}{8}\right)T_i$$

$$T_i = \frac{8}{i} + T_{i-1}$$

This takes the closed form:

$$T_i = \sum_{k=1}^{i} \frac{8}{k} \text{ where } T_0 = 0$$

$$T_4 = 0 + \frac{8}{1} + \frac{8}{2} + \frac{8}{3} + \frac{8}{4} = \frac{50}{3}$$

Therefore,

$$\mathbb{E}[T'|N = 4] = \frac{50}{3}$$

Since $\mathbb{E}[T|N = 4] = 8 + \mathbb{E}[T'|N = 4]$,

$$\mathbb{E}[T|N = 4] = 8 + \frac{50}{3} = \frac{74}{3}$$