Problem Set 2 - Linear Algebra - Solutions
Math Camp 2025, UCSB
Instructors: Camilo Abbate and Sofía Olguín

1. Let $S$ be a sample space and let $\mathcal{B}$ be a $\sigma$-algebra on $S$. Use the properties of a $\sigma$-algebra to prove that:
   (a) $S \in \mathcal{B}$ .
   (Hint: start with that $\mathcal{B}$ should be nonempty.)

   - $S \in \mathcal{B}$

$$\begin{aligned} \mathcal{B} \neq \emptyset \quad &\Rightarrow \quad E \in \mathcal{B} && (\mathcal{B} \text{ is nonempty}) \\ &\Rightarrow \quad E^c \in \mathcal{B} && (\text{closed under complements}) \\ &\Rightarrow \quad E \cup E^c \in \mathcal{B} && (\text{closed under countable unions}) \\ &\Rightarrow \quad S \in \mathcal{B} \end{aligned}$$

   - $\emptyset \in \mathcal{B}$

$$S \in \mathcal{B} \quad \Rightarrow \quad \emptyset \in \mathcal{B} \qquad (\text{closed under complements})$$

   (b) $\mathcal{B}$ is closed under countable intersections.
   - $\mathcal{B}$ is closed under countable intersections.

$$\begin{aligned} E_1, E_2, \cdots \in \mathcal{B} \quad &\Rightarrow \quad E_1^c, E_2^c, \cdots \in \mathcal{B} && (\text{closed under complements}) \\ &\Rightarrow \quad \bigcup_{i=1}^{\infty} E_i^c \in \mathcal{B} && (\text{closed under countable unions}) \\ &\Rightarrow \quad \left( \bigcap_{i=1}^{\infty} E_i \right)^c \in \mathcal{B} && (\text{DeMorgan's Laws}) \\ &\Rightarrow \quad \bigcap_{i=1}^{\infty} E_i \in \mathcal{B} && (\text{closed under complements}) \end{aligned}$$

2. Let $\mathbb{P}$ be a probability measure on a sample space $S$ with $\sigma$-algebra $\mathcal{B}$, and let $A, B \in \mathcal{B}$. Prove the following properties:

(a) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

$$\mathbb{P}(S) = \mathbb{P}(A) + \mathbb{P}(A^c) \qquad (S = A \cup A^c \quad \wedge \quad A \cap A^c = \emptyset)$$

$$1 = \mathbb{P}(A) + \mathbb{P}(A^c) \qquad (\mathbb{P}(S) = 1)$$

(b) $\mathbb{P}(A) \leq 1$

$$1 = \mathbb{P}(A) + \mathbb{P}(A^c)$$

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$$

$$\mathbb{P}(A) \leq 1 \qquad (\mathbb{P} : \mathcal{B} \mapsto [0, \infty))$$

(c) $\mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(B \cap A)$

$$\mathbb{P}(B) = \mathbb{P}(B \cap A^c) + \mathbb{P}(B \cap A)$$
$$(B = (B \cap A^c) \cup (B \cap A) \quad \wedge \quad (B \cap A^c) \cap (B \cap A) = \emptyset)$$

(d) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c) \quad (A \cup B = A \cup (B \cap A^c) \quad \wedge \quad A \cap (B \cap A^c) = \emptyset)$$

(e) If $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$

$$\mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

$$\mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(A) \qquad (A \subseteq B)$$

$$0 \leq \mathbb{P}(B) - \mathbb{P}(A) \qquad (\mathbb{P} : \mathcal{B} \mapsto [0, \infty))$$

3. Let $S$ be a sample space with $\sigma$-algebra $\mathcal{B}$, and let $A, B \in \mathcal{B}$. Prove that if $A$ and $B$ are independent, then the following pairs of events are also independent:

(a) $A$ and $B^C$

$$\begin{aligned} \mathbb{P}(A \cap B^C) &= \mathbb{P}(A) - \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) \qquad \text{(since } A \text{ and } B \text{ are independent)} \\ &= \mathbb{P}(A)(1 - \mathbb{P}(B)) \qquad \text{(rearranging)} \\ &= \mathbb{P}(A)\mathbb{P}(B^C) \qquad \text{(by property)} \end{aligned}$$

(b) $A^C$ and $B$

$$\begin{aligned} \mathbb{P}(A^C \cap B) &= \mathbb{P}(B) - \mathbb{P}(A \cap B) \\ &= \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) \qquad \text{(since } A \text{ and } B \text{ are independent)} \\ &= \mathbb{P}(B)(1 - \mathbb{P}(A)) \qquad \text{(rearranging)} \\ &= \mathbb{P}(B)\mathbb{P}(A^C) \qquad \text{(by property)} \end{aligned}$$

(c) $A^C$ and $B^C$

$$\begin{aligned} \mathbb{P}(A^C \cap B^C) &= P((A \cup B)^C) \qquad \text{(by De Morgan's Law )} \\ &= 1 - P(A \cup B) \qquad \text{(by property)} \\ &= 1 - [\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)] \\ &= 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A)\mathbb{P}(B) \qquad \text{(since } A \text{ and } B \text{ are independent)} \\ &= 1 - \mathbb{P}(A) - (1 - \mathbb{P}(A))\mathbb{P}(B) \qquad \text{(rearranging)} \\ &= (1 - \mathbb{P}(A))(1 - \mathbb{P}(B)) \qquad \text{(rearranging)} \\ &= \mathbb{P}(A^C)\mathbb{P}(B^C) \qquad \text{(by property)} \end{aligned}$$

4. Let $\mathbb{P}$ be a probability measure on a sample space $S$ with $\sigma$-algebra $\mathcal{B}$. Let $A, B, C \in \mathcal{B}$.
   (a) Show that $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$ does not imply that the events $A, B, C$ are pairwise independent. (Hint: you only need to provide a counterexample).
   Example 1.3.10 in Casella and Berger (2002)
   Let the experiment consist of tossing two dice. The sample space is:

$$S = \{(1,1), (1,2), \ldots, (1,6), (2,1), \ldots, (2,6), \ldots, (6,1), \ldots, (6,6)\}$$

This gives us $6 \times 6 = 36$ ordered pairs.
Define the following events:

$$A = \{\text{doubles appear}\} = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$$

$$B = \{\text{the sum is between 7 and 10}\}$$

$$C = \{\text{the sum is 2 or 7 or 8}\}$$

The probabilities of these events are:

$$P(A) = \frac{6}{36} = \frac{1}{6}, \quad P(B) = \frac{18}{36} = \frac{1}{2}, \quad P(C) = \frac{12}{36} = \frac{1}{3}$$

The probability of the intersection of the three events is:

$$P(A \cap B \cap C) = \text{Probability that all three events occur}$$

This happens only when the outcome is $(4,4)$, since:

- It's a double $\Rightarrow (4,4) \in A$
- $4 + 4 = 8$, which is between 7 and 10 $\Rightarrow (4,4) \in B$
- The sum is 8 $\Rightarrow (4,4) \in C$

Thus:

$$P(A \cap B \cap C) = \frac{1}{36}$$

And interestingly:

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C) = \frac{1}{6} \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{36}$$

This could suggest that the events $A$, $B$, and $C$ are independent. However,

$$P(B \cap C) = P(\text{sum equals 7 or 8}) = \frac{11}{36} \neq P(B) \cdot P(C)$$

Similarly, it can be shown that: $P(A \cap B) \neq P(A) \cdot P(B)$

Therefore, the condition: $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ is not enough to guarantee pairwise independence.

(b) What additional conditions are needed to guarantee that $A$, $B$, and $C$ are mutually independent?
Conditions for mutual independence:

- $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$
- $\mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C)$
- $\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C)$
- $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$

5. A test is used to detect the presence of a disease. The test has the following properties:
   - If a patient has the disease, the test always returns a positive result.
   - If a patient does not have the disease, the test returns a false positive with probability 0.005.

   Suppose the probability of having the disease is 0.001.
   If a patient receives a positive test result, what is the probability that they have the disease?
   From the question we know:

$$\mathbb{P}(\text{positive}|\text{disease}) = 1$$

$$\mathbb{P}(\text{positive}|\text{no disease}) = 0.005$$

$$\mathbb{P}(\text{disease}) = 0.001$$

We are interested in the probability that a patient has the disease given that they test positive: $\mathbb{P}(\text{disease}|\text{positive})$.
Recall Bayes' rule:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\sum_{j \in I} \mathbb{P}(B|A_j)\mathbb{P}(A_j)}.$$

Assign the event of having the disease to $A$ and the event of a positive test result to $B$. We first use the Law of Total Probability to obtain the probability of having a positive result, $\mathbb{P}(B)$:

$$\mathbb{P}(\text{positive}) = \mathbb{P}(\text{disease})\mathbb{P}(\text{positive}|\text{disease}) + \mathbb{P}(\text{no disease})\mathbb{P}(\text{positive}|\text{no disease})$$

$$= 0.001 \times 1 + 0.999 \times 0.005 = 0.005995.$$

Then by Bayes' rule, we have that:

$$\mathbb{P}(\text{disease}|\text{positive}) = \frac{\mathbb{P}(\text{positive}|\text{disease})\mathbb{P}(\text{disease})}{\mathbb{P}(\text{positive})}$$

$$= \frac{1 \times 0.001}{0.005995} \approx 0.1668$$

6. A variable $X$ is lognormally distributed if $Y = \ln(X)$ is normally distributed with mean $\mu$ and variance $\sigma^2$. That is, $f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}$. Let the transformation be defined by $x = g(y) = e^y$ so that $y = g^{-1}(x) = \ln(x)$.

   (a) Derive $f_X(x)$.

$$f_X(x) = f_Y\left(g^{-1}(x)\right)\left|\frac{dg^{-1}(x)}{dx}\right| = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2}\left(\frac{1}{x}\right)$$

   (b) Derive $\mathbb{E}[X^t]$ using $M_Y(t)$. What are $\mathbb{E}[X]$ and $V(X)$?

   We know that $M_Y(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$. Then

$$\mathbb{E}[X^t] = \mathbb{E}[e^{tY}] = M_Y(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

$$\mathbb{E}[X] = e^{\mu + \frac{1}{2}\sigma^2}$$

$$\mathbb{E}[X^2] = e^{2\mu + 2\sigma^2}$$

$$V(X) = (e^{\sigma^2} - 1) \cdot e^{2\mu + \sigma^2}$$

7. Let us consider the Law of Iterated Expectations in the continuous case. Suppose that $\mathbb{E}[Y] < \infty$. Prove the following results:

   (a) $\mathbb{E}[Y] = \mathbb{E}\Big[\mathbb{E}[Y|X]\Big]$

$$\begin{aligned}
\mathbb{E}[Y] &= \int_{-\infty}^{\infty} y f(y) dy && \text{(by definition of expectation)} \\
&= \int_{-\infty}^{\infty} y \left(\int_{-\infty}^{\infty} f(y,x) dx\right) dy && \text{(by definition of marginal distribution)} \\
&= \int_{-\infty}^{\infty} y \left(\int_{-\infty}^{\infty} f(y|x) f(x) dx\right) dy && \text{(by definition of conditional distribution)} \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} y f(y|x) dy\right) f(x) dx && \text{(by property of integral)} \\
&= \int_{-\infty}^{\infty} \mathbb{E}[Y|X=x] f(x) dx && \text{(by definition of conditional expectation)} \\
&= \mathbb{E}[\mathbb{E}[Y|X]] && \text{(by definition of expectation)}
\end{aligned}$$

(b) $\mathbb{E}[Y|X] = \mathbb{E}\Big[\mathbb{E}[Y|X,Z] \mid X\Big]$

Note that

$$\mathbb{E}[Y|X=x, Z=z] = \int_{-\infty}^{\infty} y f(y|x,z) dy.$$

In addition, note that

$$f(y|x,z)f(z|x) = \frac{f(y,x,z)}{f(x,z)}\frac{f(x,z)}{f(x)} = \frac{f(y,x,z)}{f(x)} = f(y,z|x).$$

Then we find that

$$\begin{aligned}
\mathbb{E}\Big[\mathbb{E}[Y|X,Z] \mid X\Big] &= \int_{-\infty}^{\infty} \mathbb{E}[Y|X=x, Z=z]\ f(z|x)dz \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} y f(y|x,z)dy\right) f(z|x)dz \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(y|x,z)f(z|x)dydz \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(y,z|x)dydz \\
&= \int_{-\infty}^{\infty} y f(y|x)dy \\
&= \mathbb{E}[Y|X].
\end{aligned}$$

8. Assume there are $n$ volunteers elegible to receive a treatment. For each unit $i \in \{1, \cdots, n\}$, define the treatment indicator

$$D_i = \begin{cases} 1 & \text{if unit } i \text{ is treated} \\ 0 & \text{otherwise} \end{cases}$$

Let $(D_1, \cdots, D_n)$ be the vector of the treatment indicators of all units. Due to capacity constraints, only $n_1 (< n)$ units can be treated: $\sum_{i=1}^{n} D_i = n_1$.

(a) What is the total number of distinct treatment assignment vectors $(D_1, \cdots, D_n)$ we can construct?

We are choosing $n_1$ out of $n$ units, unordered, so there are $\binom{n}{n_1}$ possible ways.

We say that treatment is randomly assigned if $(D_1, \cdots, D_n)$ are random variables, and if for any vector of $n$ numbers $(d_1, \cdots, d_n) \in \{0,1\}^n$ such that $\sum_{i=1}^n d_i = n_1$,

$$P(D_1 = d_1, \cdots, D_n = d_n) = \frac{1}{\binom{n}{n_1}}$$

That is, random assignment generates uniform treatment probabilities across units. Assuming the treatment is randomly assigned, answer the following:

(b) For any unit $i \in \{1, \cdots, n\}$, what is $P(D_i = 1)$?

$$P(D_i = 1) = \frac{\binom{n-1}{n_1-1}}{\binom{n}{n_1}} = \frac{n_1}{n}$$

(c) For any units $i \neq j$, what is $P(D_i = 1 \ \wedge \ D_j = 1)$? Is it true that unit $i$ getting treated is independent from unit $j$ getting treated?

$$P(D_i = 1 \ \wedge \ D_j = 1) = \frac{\binom{n-2}{n_1-2}}{\binom{n}{n_1}} = \frac{n_1(n_1-1)}{n(n-1)}.$$

Note that $P(D_i = 1 \ \wedge \ D_j = 1) \neq P(D_i = 1) P(D_j = 1)$ and $P(D_i = 1 \mid D_j = 1) \neq P(D_i = 1)$. The intuition is that if $D_j = 1$, $D_i$ is less likely to be equal to 1 than if $D_j = 0$. If $D_j = 1$, then there are only $n_1 - 1$ treatment seats left for $n - 1$ units, while if $D_j = 0$, then there are still $n_1$ treatment seats left for $n - 1$ units.

## References

Casella, G. and Berger, R. (2002). *Statistical inference*. Chapman and Hall/CRC, 2nd edition.