Department of Applied Mathematics and Statistics
The Johns Hopkins University

INTRODUCTORY EXAMINATION–SPRING SEMESTER
REAL ANALYSIS

MONDAY, JANUARY 13, 2025

Instructions: Read carefully!

1. This **closed-book** examination consists of 6 problems, each worth 5 points. Your best five scores will be used to determine the exam grade. The passing grade is 2/3 of the total points. Partial credit will be given as appropriate; each part of a problem will be given the same weight. If you are unable to prove a result asserted in one part of a problem, you may still use that result to help in answering a later part.

2. You have been provided with a syllabus indicating the scope of the exam. Our purpose is to test not only your knowledge, but also your ability to apply that knowledge, and to provide mathematical arguments presented in **clear, logically justified steps**. The grading will reflect that broader purpose.

3. The problems have not been arranged systematically by difficulty.

4. Start your answer to each problem on a NEW sheet of paper. Write only on ONE SIDE of each sheet, and please do not write very near the margins on any sheet. Arrange the sheets in order, and write your NAME and the PROBLEM NUMBER on each sheet.

5. This examination will begin at 8:30 AM and end at 11:30 AM. You may leave before then, but in that case you may not return.

6. Paper will be provided, but you should bring and use writing instruments that yield marks dark enough to be read easily.

7. **No calculators of any sort are needed or permitted.**

1. Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous. Define

$$f_n(x) = \frac{n}{2} \int_{t=x-\frac{1}{n}}^{x+\frac{1}{n}} f(t)\,\mathrm{d}t,$$

for $n = 1, 2, 3, \ldots$.

(i) Show that f_n converges pointwise to some limit function and identify the function.

(ii) Prove that the convergence is uniform in any compact interval.

Solution:

(i) The limit function is f. Fix $x \in \mathbb{R}$. By continuity of f at x, given $\epsilon > 0$ there exists $\delta > 0$ such that if $|t-x| \le \delta$ then $|f(t) - f(x)| \le \epsilon$. Choose N so that $1/N \le \delta$. Then for $n \ge N$ we have

$$\begin{aligned}
|f_n(x) - f(x)| &= \left| \frac{n}{2} \int_{t=x-\frac{1}{n}}^{x+\frac{1}{n}} f(t)\,\mathrm{d}t - f(x) \right| \\
&= \left| \frac{n}{2} \int_{t=x-\frac{1}{n}}^{x+\frac{1}{n}} (f(t) - f(x))\,\mathrm{d}t \right| \\
&\le \frac{n}{2} \int_{t=x-\frac{1}{n}}^{x+\frac{1}{n}} |f(t) - f(x)|\,\mathrm{d}t \\
&\le \frac{n}{2} \int_{t=x-\frac{1}{n}}^{x+\frac{1}{n}} \epsilon\,\mathrm{d}t = \epsilon.
\end{aligned}$$

So $f_n(x) \to f(x)$.

(ii) Fix a compact interval $[a, b]$. Since a continuous function on a compact set is uniformly continuous, f is uniformly continuous on $[a-1, b+1]$. So for any $\epsilon > 0$ there exists $\delta > 0$ such that for all $x, x' \in [a-1, b+1]$ if $|x' - x| \le \delta$, then $|f(x') - f(x)| \le \epsilon$. Choose N so that $1/N \le \delta$. Then for $n \ge N$ and $x \in [a, b]$ we have $t \in [a-1, b+1]$

whenever $t \in [x - \frac{1}{n}, x + \frac{1}{n}]$ and we have

$$\begin{aligned}
|f_n(x) - f(x)| &= \left| \frac{n}{2} \int_{t=x-\frac{1}{n}}^{x+\frac{1}{n}} f(t)\,\mathrm{d}t - f(x) \right| \\
&= \left| \frac{n}{2} \int_{t=x-\frac{1}{n}}^{x+\frac{1}{n}} (f(t) - f(x))\,\mathrm{d}t \right| \\
&\leq \frac{n}{2} \int_{t=x-\frac{1}{n}}^{x+\frac{1}{n}} |f(t) - f(x)|\,\mathrm{d}t \\
&\leq \frac{n}{2} \int_{t=x-\frac{1}{n}}^{x+\frac{1}{n}} \epsilon\,\mathrm{d}t = \epsilon.
\end{aligned}$$

So $f_n \to f$ uniformly on $[a, b]$.

---

2. What is $\lim_{x\to\infty} \left(\frac{e^x}{1+e^x}\right)^x$?

Solution: We proceed to show that the limit is 1 by showing that the limit of the reciprocal is 1. Since

$$\lim_{x\to\infty} (1 + 1/e^x)^{e^x} = \lim_{y\to\infty} (1 + 1/y)^y = e^1 = e,$$

and

$$\lim_{x\to\infty} \frac{x}{e^x} = 0,$$

we conclude that

$$\lim_{x\to\infty} \left(\frac{1+e^x}{e^x}\right)^x = \lim_{x\to\infty} (1 + 1/e^x)^x = \lim_{x\to\infty} \left[(1 + 1/e^x)^{e^x}\right]^{\frac{x}{e^x}}.$$

$$\lim_{x\to\infty} \left(\frac{1+e^x}{e^x}\right)^x = e^0 = 1.$$

---

3. Suppose X and Y are disjoint sets and $f : X \to Y$ and $g : Y \to X$ are injective (i.e., one-to-one) functions. The results of this exercise can be viewed partial building blocks for using f and g to construct a bijection between X and Y. Starting with $y_1 \in Y$ we consider infinite sequences of the form $y_1, x_1, y_2, x_2, \ldots$, with $f(x_i) = y_{i+1}$ for $i = 1, 2, \ldots$, and $g(y_i) = x_i$ for $i = 1, 2, \ldots$, which can be pictured using the following diagram:

$$y_1 \xrightarrow{g} x_1 \xrightarrow{f} y_2 \xrightarrow{g} x_2 \xrightarrow{f} y_3 \xrightarrow{g} \cdots.$$

Such a sequence will be called *initialized* if $y_1 \notin \mathrm{Im} f$, i.e., if there is no $x \in X$ such that $f(x) = y_1$.

(i) Show that an initialized sequence cannot contain repeated values, or equivalently that $x_i \neq x_j$ and $y_i \neq y_j$ whenever $i \neq j$.

(ii) Show that if $y_1, x_1, y_2, x_2, \ldots$ and $y_1', x_1', y_2', x_2', \ldots$ are initialized sequences with $y_1 \neq y_1'$, then

$$\{y_i : i = 1, 2, \ldots\} \cap \{y_i' : i = 1, 2, \ldots\} = \emptyset, \tag{a}$$

and

$$\{x_i : i = 1, 2, \ldots\} \cap \{x_i' : i = 1, 2, \ldots\} = \emptyset. \tag{b}$$

You do not have to show this, but as a consequence of (i) and (ii), if every element of X is an element in some initialized sequence then we can define a bijection $h : X \to Y$as follows. For $x \in X$, by (ii) there can be at most one initialized sequence containing x and by (i) the position of x in that sequence is unambiguous, and the map h sending x to its immediate predecessor y in that sequence is well-defined. The image of this map one-to-one since by (i) and (ii) y cannot appear in more than one initialized sequence and in that sequence it can appear at most once. For any $y \in Y$, y has some $x = g(y)$ as a successor, so $y = h(x)$ and we see that h is a surjection.

Solution: (i) The proof is by contradiction. Suppose x_i is the first repeated value in an initialized sequence. If it is x_i then we have $x_i = x_j$ for some $j > i$. Since $g(y_i) = x_i = x_j = g(y_j)$ and since g is injective we have $y_i = y_j$ giving an earlier repeated value, a contradiction. On the other hand, y_i is the first repeated value, then there are two cases to consider. If $i = 1$ then $y_1 = y_j$ for some $j > 1$. But then we would have $y_1 = f(x_{j-1})$ contradicting the assumption that $y_1 \notin \mathrm{Im}(f)$. If $i > 1$ then we would have $f(x_{i-1}) = y_i = y_j = f(x_{j-1})$ so $x_{i-1} = x_{j-1}$ by injectivity of f, again giving an earlier repeated value, a contradiction.

(ii) The proof of (a) is by contradiction. Suppose $y_i = y_j'$. If $i, j > 1$ then we have $f(x_{i-1}) = y_i = y_j' = f(x_{j-1}')$ so $x_{i-1} = x_{j-1}'$ since f is injective. Similarly, $y_{i-1} = g(x_{i-1}) = g(x_{j-1}') = y_{j-1}'$ giving $y_{i-1} = y_{j-1}'$ since g is injective. If $i = j$ we can repeat this argument $i - 1$ times to give $y_1 = y_1'$ a contradiction. If $j > i$ then $y_1 = y_j' = f(x_j')$ contradicting the assumption that the first sequence is initialized. If $j < i$ we also get a contradiction by the same argument.

The proof of (b) is by contradiction. If $x_i = x_j'$ for some i, j then we have $y_i = f(x_i) = f(x_j') = y_j$ which contradicts (a).

---

4. Suppose $(x^{(n)})_{n=1}^{\infty}$ is a sequence of points in $\mathbb{R}^d$ with the property that

$$\|x^{(n+1)} - x^{(n)}\| \leq c\|x^{(n)} - x^{(n-1)}\|$$

for $n = 2, 3, \ldots,$ and for some constant $c \in (0, 1)$. (Here $\|x\|$ denotes the Euclidean norm of $x \in \mathbb{R}^d$.) Show that the sequence converges to some finite limit.

Solution: For $m > n$ observe that by making repeated use of the given inequality we have

$$\begin{aligned}\|x^{(m)} - x^{(m-1)}\| &\le c\|x^{(m-1)} - x^{(m-2)}\| \\ &\le c^2\|x^{(m-2)} - x^{(m-3)}\| \\ &\vdots \\ &\le c^{m-n-1}\|x^{(n+1)} - x^{(n)}\|.\end{aligned}$$

We proceed to show that the sequence is Cauchy, from which the desired conclusion follows. Indeed, if $m > n \ge 1$ we can use the triangle inequality to write

$$\begin{aligned}\|x^{(m)} - x^{(n)}\| &= \left\|\sum_{i=n+1}^{m} (x^{(i)} - x^{(i-1)})\right\| \\ &\le \sum_{i=n+1}^{m} \|x^{(i)} - x^{(i-1)}\| \\ &\le \sum_{i=n+1}^{m} c^{i-n-1}\|x^{(n+1)} - x^{(n)}\| \le \frac{1}{1-c}\|x^{(n+1)} - x^{(n)}\| \\ &\le \frac{c^{n-1}}{1-c}\|x^{(2)} - x^{(1)}\|.\end{aligned}$$

---

5. Suppose $f : [0, +\infty) \to \mathbb{R}$ is a continuous function and a sequence of functions $f_n : [0, +\infty) \to \mathbb{R}$ for $n = 1, 2, \ldots$ is defined by $f_n(x) := f(x^n)$. Suppose that the functions $f_1, f_2, \ldots$ are equicontinous at 1, i.e., that for any $\epsilon > 0$ there exists $\delta > 0$ such that

$$|f_n(x) - f_n(1)| < \epsilon \text{ for all } n = 1, 2, \ldots \text{ and } x \in (1 - \delta, 1 + \delta).$$

Show that f must be a constant function.

Solution: Fix $x^* > 0$ and $\epsilon > 0$. By equicontinuity at 1, there exists $\delta > 0$ such that for all n and x satisfying $|x - 1| \le \delta$ we have $|f_n(x) - f_n(1)| \le \epsilon$. For such a value of δ, since $\lim_{n\to\infty}(x^*)^{1/n} = 1$, there exists n such that $|(x^*)^{1/n} - 1| \le \delta$, so we have

$$|f(x^*) - f(1)| = |f_n((x^*)^{1/n}) - f_n(1^{1/n})| = |f_n((x^*)^{1/n}) - f_n(1)| \le \epsilon.$$

Since ϵ is arbitrary and f is continuous we conclude that $f(x^*) = f(1)$. We have shown $f(x^*) = f(1)$ for all $x^* > 0$. Since f is continuous, we can conclude that $f(0) = f(1)$ holds as well, so f is constant.

6. Consider the power series

$$f(x) = \sum_{p \text{ prime}} x^p = x^2 + x^3 + x^5 + \cdots .$$

(i) What is the radius of convergence? Hint: Compare with other series with known radius of convergence.

(ii) Show that

$$f(x) \le \frac{x^2}{1-x} \text{ for all } 0 \le x < 1.$$

Solution: (i) We have

$$|f(x)| \le \sum_{p \text{ prime}} |x|^p \le \sum_{p=2}^{\infty} |x|^p,$$

which we know converges for $|x| < 1$ so the radius of convergence is at least 1. On the other hand, the series does not converge for $x = 1$, so the radius of convergence is 1.

(ii) For $0 \le x < 1$ the we have

$$f(x) \le \sum_{p=2}^{\infty} x^p = x^2 \sum_{p=0}^{\infty} x^p = \frac{x^2}{1-x}.$$

Department of Applied Mathematics and Statistics
The Johns Hopkins University

Probability

Tuesday, January 14, 2025

Instructions: Read carefully!

1. This **closed-book** examination consists of 6 problems, each worth 5 points. Your best five scores will be used to determine the exam grade. The passing grade is 2/3 of the total points. Partial credit will be given as appropriate; each part of a problem will be given the same weight. If you are unable to prove a result asserted in one part of a problem, you may still use that result to help in answering a later part.
2. You have been provided with a syllabus indicating the scope of the exam. Our purpose is to test not only your knowledge, but also your ability to apply that knowledge, and to provide mathematical arguments presented in **clear, logically justified steps**. The grading will reflect that broader purpose.
3. The problems have not been arranged systematically by difficulty.
4. Start your answer to each problem on a NEW sheet of paper. Write only on ONE SIDE of each sheet, and please do not write very near the margins on any sheet. Arrange the sheets in order, and write your NAME and the PROBLEM NUMBER on each sheet.
5. This examination will begin at 8:30 AM and end at 11:30 AM. You may leave before then, but in that case you may not return.
6. Paper will be provided, but you should bring and use writing instruments that yield marks dark enough to be read easily.
7. **No calculators of any sort are needed or permitted.**

1. Suppose $n > 1$. We put n balls into n boxes independently of one another so that each ball is equally likely to be put into any of the n boxes. Compute the probability that box number 1 is the only empty box.

Solution: Box 1 is the only empty box means that the n balls must go into the remaining $n-1$ boxes in such a way that exactly 2 balls go into some box other than box 1 and the remaining $n-2$ balls must go into the remaining $n-2$ boxes one in each. Let E_i be the event that box 1 is the only empty box and the 2 balls go into box i ($i = 2, 3, \ldots, n$). To count the number of configurations in E_i we select the 2 balls to go into box i from the n possible in $\binom{n}{2}$ ways, and once these are selected, the remaining $n-2$ balls can go into the remaining $n-2$ boxes in $(n-2)!$ ways to keep box 1 empty. Since there are n^n possible ways the n balls can enter the n boxes, we have

$$P(E_i) = \frac{\binom{n}{2}(n-2)!}{n^n}.$$

The E_i's are mutually exclusive, therefore,

$$P(\bigcup_{i=2}^{n} E_i) = \sum_{i=2}^{n} P(E_i) = (n-1)\frac{\binom{n}{2}(n-2)!}{n^n} = \frac{\binom{n}{2}(n-1)!}{n^n}$$

is the desired probability.

---

2. For each integer $n \geq 1$, X_n is discrete uniform on the integers from 1 through n inclusive, and let N be geometrically distributed with probability of success $0 < p < 1$ independent of all the X_n's. Compute $E(X_N)$ and simplify as much as possible. Recall: $P(N = n) = p(1-p)^{n-1}$ for $n = 1, 2, 3, \ldots$; and, $P(X_n = j) = \frac{1}{n}$ for $j = 1, 2, \ldots, n$.

Solution: We first compute $E(X_N|N = n) = E(X_n|N = n) = E(X_n) = \frac{n+1}{2}$. By the law of total expectation

$$\begin{aligned} E(X_N) &= \sum_{n=1}^{\infty} E(X_n|N=n)P(N=n) \\ &= \sum_{n=1}^{\infty} \frac{n+1}{2} \cdot p(1-p)^{n-1} \\ &= \frac{1}{2}\sum_{n=1}^{\infty} n \cdot p(1-p)^{n-1} + \frac{1}{2}\sum_{n=1}^{\infty} p(1-p)^{n-1} \\ &= \frac{1}{2}E(N) + \frac{1}{2} \cdot 1 \stackrel{*}{=} \frac{1}{2p} + \frac{1}{2} \\ &= \frac{1+p}{2p}. \end{aligned}$$

* Here we used the fact that the mean of a geometric(p) is $1/p$.

---

3. Suppose $n > 1$. We flip a fair coin until the n-th head occurs and we stop. If this nth head occurs on the $2n$-th flip of the coin, what is the probability that the $(n-1)$-st head happens on flip $2n-1$? Simplify completely.

Solution: To determine the probability that the $(n-1)$-th head occurs on the $(2n-1)$-th flip given that the n-th head occurs on the $2n$-th flip, we observe that there are $\binom{2n-1}{n-1}$ possible sequences with exactly $n-1$ heads in the first $2n-1$ flips. Among these, the number of sequences where the $(n-1)$-th head specifically occurs on the $(2n-1)$-th flip is $\binom{2n-2}{n-2}$, since the first $2n-2$ flips must contain exactly $n-2$ heads. Therefore, the desired probability is the ratio of these two quantities:

$$\frac{\binom{2n-2}{n-2}}{\binom{2n-1}{n-1}} = \frac{n-1}{2n-1}.$$

Alternatively: Given that the $2n$-th flip is the first time we see n heads, all arrangements of the the $n-1$ heads and n tails among the previous $2n-1$ flips are equally likely, so the chance that the last $(2n-1)$-th flip is heads is

$$(n-1)/(2n-1)$$

.

---

4. Let X and Y be independent and identically distributed unit exponential random variables, i.e., they each have the PDF $f(x) = e^{-x}$ for $x > 0$. Compute the conditional probability that $X < 2$ given that $X/Y > 1$.

Solution:

$$\begin{aligned}
P(X < 2|X > Y) &= \frac{P(Y < X < 2)}{P(X > Y)} \\
&\overset{**}{=} \frac{\int_0^2 \int_0^x e^{-x}e^{-y}\,dydx}{\frac{1}{2}} \\
&= 2\int_0^2 e^{-x} - e^{-2x}\,dx \\
&= 1 - 2e^{-2} + e^{-4} \quad \text{or} \quad (1-e^{-2})^2.
\end{aligned}$$

**FYI: Since X, Y are IID and continuous it follows that $P(X > Y) = P(Y > X)$ and $P(X = Y) = 0$, which implies $P(X > Y) = \frac{1}{2}$.

---

5. Let $a > 0$ be a fixed positive constant, and suppose that in a crowd of size n the probability that a particular person has a trait is a/n independent from person to person. If it's known that as n tends to ∞, the probability the crowd has nobody with the trait is $1/e^2$. Determine with justification the value of a.

Solution: If we let X_n count the number in the crowd of size n that exhibit the trait, then X_n is binomial with parameters n and $p = \frac{a}{n}$. Therefore, $P(X_n = 0) = (1 - \frac{a}{n})^n$. Now, as $n \to \infty$, we have

$$\lim_{n\to\infty} (1 - \frac{a}{n})^n = e^{-a} = \frac{1}{e^2}.$$

This implies $a = 2$.

Alternatively: X_n converges in distribution to $X \sim$ Poisson(a). Consequently, $P(X = 0) = \frac{e^{-a}a^0}{0!} = e^{-a}$. Since we're told $P(X = 0) = e^{-2}$, it clearly follows $a = 2$.

---

6. Consider rectangles whose base and height are independent and uniformly distributed on the interval $[0, 1]$. Find the PDF of the area of the rectangle.

Solution: Let X and Y, respectively, denotes the length and width of the rectangle and let $A = XY$ be the resulting area. Since $x, y \in (0,1)$ we have $a = xy \in (0,1)$, and, for such a, the CDF is

$$\begin{aligned} F_A(a) = P(XY \le a) &= 1 - P(XY > a) \\ &= 1 - \int_a^1 \int_{a/x}^1 1\, dy dx \\ &= 1 - \int_a^1 1 - \frac{a}{x}\, dx \\ &= 1 - \left\{ x - a\ln(x)|_{x=a}^{x=1} \right\} = a - a\ln(a). \end{aligned}$$

Consequently, the PDF of the area is

$$f_A(a) = \frac{d}{da} F_A(a) = -\ln(a) \quad \text{for } 0 < a < 1.$$

Alternatively: Consider the transformation $a = xy$ and $b = y$. The inverse transformation is $x = a/b$ and $y = b$. Since $0 < x, y < 1$ it follows $0 < a < b < 1$, and the Jacobian on this inverse transformation is $J = \det \begin{bmatrix} \frac{1}{b} & -\frac{a}{b^2} \\ 0 & 1 \end{bmatrix} = \frac{1}{b}$. It follows the joint PDF of A and B is

$$f_{A,B}(a,b) = \frac{1}{b} \quad \text{for } 0 < a < b < 1$$

and the marginal PDF of A is, for $0 < a < 1$,

$$f_A(a) = \int_a^1 \frac{1}{b}\, db = -\ln(a).$$

Department of Applied Mathematics and Statistics
The Johns Hopkins University

Introductory Examination–Spring Semester
Linear Algebra

Wednesday, January 15, 2025

Instructions: Read carefully!

1. This **closed-book** examination consists of 6 problems, each worth 5 points. Your best five scores will be used to determine the exam grade. The passing grade is 2/3 of the total points. Partial credit will be given as appropriate; each part of a problem will be given the same weight. If you are unable to prove a result asserted in one part of a problem, you may still use that result to help in answering a later part.

2. You have been provided with a syllabus indicating the scope of the exam. Our purpose is to test not only your knowledge, but also your ability to apply that knowledge, and to provide mathematical arguments presented in **clear, logically justified steps**. The grading will reflect that broader purpose.

3. The problems have not been arranged systematically by difficulty. If a problem directs you to use a particular method of analysis, you *must* use it in order to receive substantial credit.

4. Start your answer to each problem on a NEW sheet of paper. Write only on ONE SIDE of each sheet, and please do not write very near the margins on any sheet. Arrange the sheets in order, and write your NAME and the PROBLEM NUMBER on each sheet.

5. This examination will begin at 8:30 AM and end at 11:30 AM. You may leave before then, but in that case you may not return.

6. Paper will be provided, but you should bring and use writing instruments that yield marks dark enough to be read easily.

7. **No calculators of any sort are needed or permitted.**

1. Let $S \in \mathbb{R}^{n\times n}$ be a symmetric and invertible matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Consider a polynomial $p(x)$ of degree at most $n-1$:

$$p(x) = h_0 + h_1 x + h_2 x^2 + \ldots + h_{n-1} x^{n-1}.$$

Suppose you are given values $y_1, y_2, \ldots, y_n \in \mathbb{R}$ and you wish to determine coefficients $h_0, h_1, \ldots, h_{n-1}$ such that $p(\lambda_i) = y_i$ for all $i = 1, 2, \ldots, n$.

 (a) Write the linear system corresponding to this problem, i.e., write the problem in the form $Ab = c$, specifying the $n \times n$ matrix A and the column vectors b and c.

 (b) What conditions must S (or equivalently, its eigenvalues $\lambda_1, \ldots, \lambda_n$) satisfy for the system in (a) to have a unique solution? Explain.

Solution: (a)

$$A = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{n-1} \\ 1 & \vdots & \vdots & \ldots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \ldots & \lambda_n^{n-1} \end{bmatrix} \quad b = \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{n-1} \end{bmatrix} \quad c = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

(b) In order for this system to have a unique solution, A must be invertible. Hence, we need $\det(A) \neq 0$. Calculating this determinant, we get

$$\det(A) = \prod_{0 \leq i \leq j \leq n} \lambda_j - \lambda_i$$

hence all λ_i must be distinct or, equivalently, the matrix S non-derogatory. The matrix A is called the Vandermonde matrix.

---

2. Verify that given an invertible matrix $A \in \mathbb{R}^{n\times n}$ and $u, v \in \mathbb{R}^n$ such that $1 + v^T A^{-1} u \neq 0$, then

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1} u v^T A^{-1}}{1 + v^T A^{-1} u}.$$

Solution: Multiplying the expressions in the identity on the right it suffices to show

$$(I + A^{-1} u v^T)^{-1} = I - \frac{A^{-1} u v^T}{1 + v^T A^{-1} u}.$$

Letting $w = A^{-1}u$ this reduces to showing that

$$(I + wv^T)^{-1} = I - \frac{wv^T}{1 + v^T w}$$

(which is a somewhat standard identity). To see this, we have

$$(I + wv^T)(I - \frac{wv^T}{1 + v^T w}) = I + wv^T - \frac{wv^T}{1 + v^T w} - \frac{wv^T wv^T}{1 + v^T w}$$

$$= I + wv^T - \frac{wv^T}{1 + v^T w} - \frac{w(v^T w)v^T}{1 + v^T w}$$

$$= I + (1 - \frac{1}{1 + v^T w} - \frac{v^T w}{1 + v^T w})wv^T = I.$$

---

3. Let $A, B \in \mathbb{R}^{n \times n}$. Suppose that (i) AB is symmetric and that (ii) A and B commute. This problem has two parts:

   (a) Prove that A does not need to be symmetric by providing a counterexample, i.e., give an example of matrices A and B satisfying conditions (i) and (ii) for which A is not symmetric.

   (b) Give conditions on B such that for all A satisfying (i) and (ii) we can conclude that A is symmetric.

   Include a clear proof and explanation for both parts.

   Solution: (a) For example, if we take A to be any invertible and asymmetric matrix and $B = A^{-1}$ then $AB = BA = I$, and the product is symmetric.

   (b) Note that $B = I$ works, but we can weaken this to B symmetric and invertible. Under this assumption, if A and B commute and AB is symmetric we have

$$AB = BA = (AB)^T = B^T A^T = BA^T,$$

   and multiplying on the left by B^{-1} we get $A = A^T$.

---

4. Let $U, V \in \mathbb{C}^{n \times n}$. Assume that $UV = VU$ and that U is Hermitian. Prove that there exists a basis in which U is diagonal and V is block diagonal, with each block corresponding to an eigenspace of U.

Solution: Since U is Hermitian, by the spectral theorem, there exists a unitary matrix Q such that $UQ = Q\Lambda$ with Λ a diagonal matrix with ordered diagonal elements $\lambda_1, \dots, \lambda_n$ expressed as

$$\underbrace{\lambda'_1, \dots \lambda'_1}_{k_1}, \underbrace{\lambda'_2, \dots \lambda'_2}_{k_2}, \dots \underbrace{\lambda'_m, \dots \lambda'_m}_{k_m}.$$

with λ'_i distinct. Let the $I_j = \{p \; : \; \lambda'_p = \lambda_j\}$ for $j = 1, \dots, m$. Since Q is unitary, its columns $q^{(i)},\ i = 1, \dots, n$ form a basis of $\mathbb{C}^n$ and Λ is the matrix of Q in this basis. In addition, $q^{(i)}, i \in I_j$ forms a basis of the subspace $\{q \in \mathbb{C}^n \; : \; Uq = \lambda'_j q\}$. Observe that for any eigenvector q with eigenvalue λ_i we have

$$UVq = VUq = V\lambda_i q = \lambda Vq$$

i.e. Vq is also an eigenvector of U with the same eigenvalue. If follows that $\lambda = \lambda'_j$ for some j so we can write Vq as a linear combination of $q^{(i)}, i \in I_j$. Thus, the matrix of V in our basis has the desired block form.

---

5. Consider the following matrix:

$$T = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where $A \in \mathbb{R}^{u \times u}$, $B \in \mathbb{R}^{u \times w}$, $C \in \mathbb{R}^{w \times u}$, and $D \in \mathbb{R}^{w \times w}$.

Assuming A is invertible, define $S = D - CA^{-1}B$.

(a) Prove that T is invertible if and only if S is also invertible.

(b) Assuming invertible S, prove $\text{rank}(T) = \text{rank}(A) + \text{rank}(S)$.

Hint: Start from the factorization:

$$T = \begin{bmatrix} A & 0 \\ C & S \end{bmatrix} \begin{bmatrix} I_u & A^{-1}B \\ 0 & I_w \end{bmatrix}.$$

Solution:

(a) (i) T invertible $\Rightarrow S$ invertible

Write

$$T^{-1} = \begin{bmatrix} E & F \\ G & H \end{bmatrix}$$

where $E \in \mathbb{R}^{u \times u}$, $F \in \mathbb{R}^{u \times w}$, $G \in \mathbb{R}^{w \times u}$, $H \in \mathbb{R}^{w \times w}$.

Since by assumption T is invertible,

$$I = \begin{bmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{bmatrix}$$

which implies

$$\begin{aligned} AF + BH = 0_{u\times w} \quad &\Rightarrow \quad F = -A^{-1}BH \\ CF + DH = I_w \quad &\Rightarrow \quad (D - CA^{-1}B)H = SH = I_w \quad \Rightarrow \quad H = S^{-1} \end{aligned}$$

thus proving invertibility of S.

(ii) S invertible $\Rightarrow T$ invertible

It suffices to show that the factors in the matrix product given in the hint are both invertible. The upper block triangular matrix (right factor) is clearly invertible, as its determinant is 1. The lower block triangular matrix (left factor) has the following inverse

$$\begin{bmatrix} A & 0 \\ C & S \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & 0 \\ -S^{-1}CA^{-1} & S^{-1} \end{bmatrix}$$

which exists due to invertibility of A and S (by assumption).

(b) Let $T = UV$ with U, V the left and right factors in the hint respectively. Knowing that the rank of a block triangular matrix is lower bounded by the sum of the ranks of the block diagonals, we have $\operatorname{rank}(V) = u + w$. Hence, since V is full rank, $\operatorname{rank}(UV) = \operatorname{rank}(U)$. By the same property,

$$\operatorname{rank}(T) = \operatorname{rank}(U) \geq \operatorname{rank}(A) + \operatorname{rank}(S).$$

But since A and S are invertible, they have full rank and the above holds with equality.

---

6. Let $A \in \mathbb{R}^{m\times n}$ with singular value decomposition:

$$A = U\Sigma V^T$$

where Σ is a rectangular matrix with diagonal entries $\sigma_1, \sigma_2, \ldots, \sigma_r, 0, \ldots, 0$.

Let P_R denote the orthogonal projection matrix onto the range of A.

(a) Write P_R as a function of U.

(b) Prove $P_R = AA^\dagger$, where $A^\dagger$ is the Moore-Penrose pseudoinverse.

Solution: (a)

$$P_R = U \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} U^T$$

(b) We know

$$A^\dagger = V\Sigma^\dagger U^T = V \begin{bmatrix} \frac{1}{\sigma_1} & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & 0 & \vdots & \vdots & \dots & 0 \\ \vdots & 0 & \ddots & 0 & \vdots & \dots & 0 \\ \vdots & \vdots & 0 & \frac{1}{\sigma_r} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix} U^T$$

leading to $AA^\dagger = U\Sigma V^T V \Sigma^\dagger U^T = U\Sigma\Sigma^\dagger U^T = P_R$.
