# Department of Applied Mathematics and Statistics The Johns Hopkins University

## INTRODUCTORY EXAMINATION–SPRING SEMESTER REAL ANALYSIS

Tuesday, January 16, 2024

# 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. Let  $f:[0,1] \to \mathbb{R}$  be a Riemann integrable function satisfying

$$\int_0^1 |f(x)| \,\mathrm{d}x = 0.$$

- (a) Use the definition of continuity to show that if f is continuous then f = 0.
- (b) Give an example of a discontinuous function f that satisfies the assumptions of the problem but  $f \neq 0$ .

#### Solution:

(a) Assume  $f \neq 0$ ; then there exists  $x \in [0,1]$  such that |f(x)| = a > 0. By continuity there exists  $\delta > 0$  such that, if  $|y - x| < \delta$ , then |f(x) - f(y)| < a/2. Then |f(y)| > a/2 for all y in the open ball  $B(x, \delta)$ . Therefore

$$\int_{0}^{1} |f(x)| \mathrm{d}x = \int_{B(x,\delta)} |f(x)| \mathrm{d}x + \int_{[0,1]\setminus B(x,\delta)} |f(x)| \mathrm{d}x \ge \delta \frac{a}{2} + 0 > 0$$

This is a contradiction. Therefore f = 0.

(b) Consider the function

$$f(x) = \begin{cases} 1 & \text{if } x = 0\\ 0 & \text{if } x \in (0, 1] \end{cases}$$

This function is Riemann integrable since it is piecewise continuous with only one point of discontinuity, but it is obviously not vanishing since f(0) = 1.

2. Consider the function  $f : \mathbb{R}^2 \to \mathbb{R}$  defined as

$$f(x,y) = \begin{cases} \frac{xy}{x^2 + y} & \text{if } x^2 \neq -y \\ 0 & \text{if } x^2 = -y. \end{cases}$$

Show that all directional derivatives at (0,0) exist but f is not differentiable at (0,0).

Solution: If  $e = (e_1, e_2)$  with  $e_2 \neq 0$ , then as  $t \to 0$  with  $t \neq 0$  we have

$$\frac{1}{t}f(te_1, te_2) = \frac{1}{t}\frac{t^2e_1e_2}{t^2e_1^2 + te_2} = \frac{e_1e_2}{te_1^2 + e_2} \to e_1.$$

If instead  $e_2 = 0$ , then as  $t \to 0$  we have

$$\frac{1}{t}f(te_1, te_2) = \frac{1}{t}f(te_1, 0) = 0 \to 0.$$

Therefore, each directional derivative exists at (0,0), but f is not continuous. For instance, consider  $y = -x^2 + x^3$  with  $x \neq 0$  and  $x \to 0$ . Then  $(x, y) \to (0,0)$  but  $f(x, y) = (-x^3 + x^4)/x^3 \to -1 \neq 0 = f(0,0)$ .

Solution #2: Consider the smooth function  $g(x, z) = z - x^2$  and note that

$$F(x,z) = f(x,g(x,z)) = \begin{cases} x - \frac{x^2}{z} & z \neq 0\\ 0 & z = 0 \end{cases}$$

*F* is clearly not a differentiable or even a continuous function at (x, z) = (0, 0), but it would have to be so by the chain rule if *f* were differentiable at (x, y) = (0, 0). This proves that *f* is in fact not differentiable at (x, y) = (0, 0).

Solution #3: If f is differentiable at (0, 0), then the directional derivative in direction  $e = (e_1, e_2)$  must be a linear function of e. However,

$$D_e f = \begin{cases} e_1, & e_2 \neq 0\\ 0, & e_2 = 0 \end{cases}$$

is not a linear function of e.

- 3. Let X be a topological space and consider a function  $f : X \to X$ . (You may not assume that X is a metric space for this problem.)
  - (a) Define what it means for f to be continuous.
  - (b) Let A be a subset of X. Define what it means for A to be compact.
  - (c) Show that f(A) is compact if f is continuous and A is compact.

#### Solution:

- (a)  $f: X \to X$  is continuous iff for every open set  $U \subset X$  we have that  $f^{-1}(U)$  is open in X.
- (b)  $A \subset X$  is compact if every cover of A by open sets has a finite subcover.
- (c) Let  $\{U_i\}_{i\in I}$  be a cover of f(A) by open sets. Take  $V_i = f^{-1}(U_i)$  for  $i \in I$ . It's easy to see that  $\{V_i\}_{i\in I}$  is a cover of A, since for any  $a \in A$  the image  $f(a) \in U_i$ for at least one i and thus  $a \in V_i$  for at least that i. Clearly,  $V_i = f^{-1}(U_i)$  is open for all  $i \in I$  by part (a). Since A is compact, there is a finite subcover  $A \subset \bigcup_{k=1}^n V_{i_k}$ . Therefore  $f(A) \subset \bigcup_{k=1}^n U_{i_k}$ .

4. Give an example of a closed and bounded set (in some metric space) that is not compact. Prove all your claims.

Solution: Many solutions are possible. For instance, take  $\mathbb{R}$  with the metric  $d(x, y) = \min\{|x-y|, 1\}$ . Note that with this metric, the entire real line is closed and bounded, but it is not compact. To see this last claim, consider the collection of open sets  $(n - \frac{3}{4}, n + \frac{3}{4})$  with  $n \in \mathbb{Z}$ . This collection covers the real line, but no subcollection does.

5. Show that the sequence  $\sqrt{2}$ ,  $\sqrt{2\sqrt{2}}$ ,  $\sqrt{2\sqrt{2\sqrt{2}}}$ , ... converges and find its limit.

Solution: We can write the sequence as  $x_1 = \sqrt{2}$ ,  $x_{n+1} = \sqrt{2x_n}$ . One can check:

- $x_n < 2$  for all n by induction on n.
- $x_{n+1} > x_n$  for all *n* by the previous result.

Therefore the sequence is convergent. To find the limit x, one can solve  $x = \sqrt{2x}$ , which implies that the limit is either 0 or 2. Since the sequence is increasing and greater than  $\sqrt{2}$ , then the limit is 2.

- 6. (a) By direct calculation, determine at precisely what points  $(x_0, y_0) \in \mathbb{R}^2$  you can solve the equation  $F(x, y) = y^2 + y + 3x + 1 = 0$  for y as a unique, continuously differentiable real-valued function f defined for x in a neighborhood of  $x_0$  so that  $y_0 = f(x_0)$ .
  - (b) Check whether your answer to part (a) agrees with the answer you expect from the implicit function theorem. Compute dy/dx.

#### Solution:

(a) Solving the quadratic equation we get

$$y = \frac{-1 \pm \sqrt{1 - 4(3x + 1)}}{2} = \frac{-1 \pm \sqrt{-3(1 + 4x)}}{2}$$

Therefore y is a continuously differentiable function of x if 1 + 4x < 0, i.e., if x < -1/4. The quadratic formula with the + sign defines a smooth function for y > -1/2, while this formula with the - sign defines a smooth function for y < -1/2. When x = -1/4 we have the point (-1/4, -1/2) where y is not a differentiable function of x. For x > -1/4 no real points of the form (x, y) satisfy F(x, y) = 0. Thus, we conclude that any  $(x_0, y_0) \in \mathbb{R}^2$  satisfying  $x_0 < -1/4$  and  $y_0 \neq -1/2$  meets the conditions stated in the problem.

(b) The implicit function theorem says that if  $(x_0, y_0)$  satisfies  $F(x_0, y_0) = 0$  and  $\frac{\partial F(x,y)}{\partial y}\Big|_{(x,y)=(x_0,y_0)} \neq 0$ , then y is given by an implicit function y = f(x) for x in a neighborhood of  $x_0$  such that  $f(x_0) = y_0$ . Note that  $\partial F/\partial y = 2y + 1 \neq 0$  precisely when  $y \neq -1/2$ . On the solution set where F = 0 note further that  $x = -\frac{1}{3}(1+y+y^2) = -\frac{1}{4} - \frac{1}{3}(y+\frac{1}{2})^2 \leq -\frac{1}{4}$ , and  $x = -\frac{1}{4}$  only if  $y = -\frac{1}{2}$ . Thus, the implicit function theorem implies that points  $(x_0, y_0) \in \mathbb{R}^2$  where an implicit function exists will be precisely those satisfying  $x_0 < -1/4$  and  $y_0 \neq -1/2$ , consistent with the previous direct solution.

Note that 
$$dy/dx = -\frac{\partial F/\partial x}{\partial F/\partial y} = -\frac{3}{2y+1} = \mp \sqrt{\frac{3}{-(1+4x)}}$$
.

# Department of Applied Mathematics and Statistics The Johns Hopkins University

## INTRODUCTORY EXAMINATION–SPRING SEMESTER PROBABILITY

Wednesday, January 17, 2024

# 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. Suppose  $X_1$  and  $X_2$  are independent Poisson random variables with means  $\lambda_1$  and  $\lambda_2$  respectively. Determine the conditional distribution of  $X_1$  given that  $X_1 + X_2 = n$ .

Solution: Let  $X_3 = X_1 + X_2$ . Since  $X_3$  is a sum of two Poisson random variables, it is also Poisson, with parameter  $\lambda_1 + \lambda_2$ :

$$P(X_3 = n) = \frac{(\lambda_1 + \lambda_2)^n e^{-\lambda_1 - \lambda_2}}{n!}.$$

For  $P(X_1 = k | X_1 + X_2 = n)$ , we have:

$$P(X_1 = k \mid X_1 + X_2 = n) = \frac{P(X_1 = k, X_2 = n - k)}{P(X_3 = n)}.$$

Since  $X_1$  and  $X_2$  are independent,

$$P(X_1 = k, X_2 = n - k) = P(X_1 = k)P(X_2 = n - k) = \frac{\lambda_1^k e^{-\lambda_1}}{k!} \frac{\lambda_2^{n-k} e^{-\lambda_2}}{(n-k)!}$$

Combining the above,

$$P(X_1 = k \mid X_1 + X_2 = n) = \binom{n}{k} \frac{\lambda_1^k \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^n}.$$

Note that this is the distribution of a Binomial  $\left(n, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$  random variable.

2. Let  $S_n$  be the number of successes in n tosses of a bent coin whose success probability is p. Prove a weak law of large numbers for  $\frac{S_n}{n}$ , i.e., show that  $\frac{S_n}{n}$  converges to p in probability as  $n \to \infty$ .

Solution: Let q = 1 - p. Note that  $S_n$  is a binomial random variable with mean np and variance npq. Hence,  $\frac{S_n}{n}$  has mean p and variance  $\frac{pq}{n}$ . By Chebyshev's inequality, we have

$$P\left(\left|\frac{S_n}{n} - p\right| > \epsilon\right) \le \frac{pq}{n\epsilon^2}$$

for any  $\epsilon > 0$ . Hence, for any  $\epsilon > 0$ ,

$$\lim_{n \to \infty} P\left( \left| \frac{S_n}{n} - p \right| > \epsilon \right) = 0,$$

completing the proof.

Solution #2: Note by definition that  $S_n = \frac{1}{n} \sum_{i=1}^n X_i$ , where  $X_1, \ldots, X_n$  are independent Bernoulli random variables with mean p. Thus, each  $X_i$  has characteristic function

$$E[e^{itX_i}] = pe^{it} + (1-p) = p(e^{it} - 1) + 1,$$

and

$$E[e^{itS_n/n}] = \left[1 + p(e^{it/n} - 1)\right]^n$$
  
=  $\left[1 + \frac{ipt}{n} + O\left(\left(\frac{t}{n}\right)^2\right)\right]^n$   
 $\rightarrow e^{ipt} \text{ as } n \rightarrow \infty.$ 

Thus,  $S_n/n$  converges in distribution to a constant value p as  $n \to \infty$ , which implies convergence of  $S_n/n$  in probability to p.

3. We start with a stick of length L. We break it at a point which is chosen randomly and uniformly over its length, and keep the piece that contains the left end of the stick. We then repeat the same process with the stick that we keep. After breaking twice, what are the expected length and variance of the stick we are left with?

Solution: Let Y be the length of the stick after we break it for the first time. Let X be the length after the second time. We have E[X | Y] = Y/2, since the breakpoint is chosen uniformly over the length Y of the remaining stick. For a similar reason, we also have E[Y] = L/2. Thus,

$$E[X] = E[E[X|Y]] = E\left[\frac{Y}{2}\right] = \frac{1}{2} \cdot \frac{L}{2} = \frac{L}{4}.$$

Since Y is uniformly distributed between 0 and L,

$$\operatorname{var}(E[X|Y]) = \operatorname{var}\left(\frac{Y}{2}\right) = \frac{1}{4} \cdot \frac{L^2}{12} = \frac{L^2}{48}.$$

Since X is uniformly distributed between 0 and Y, we have

$$\operatorname{var}(X|Y) = \frac{Y^2}{12}.$$

Thus, since Y is uniformly distributed between 0 and L,

$$E[\operatorname{var}(X|Y)] = \frac{1}{12} \int_0^L \frac{1}{L} y^2 \mathrm{d}y = \frac{L^2}{36}.$$

Using the law of total variance, we obtain

$$\operatorname{var}(X) = \operatorname{var}(E[X|Y]) + E[\operatorname{var}(X|Y)] = \frac{7L^2}{144}.$$

Solution #2: Following the notations of the previous solution, the pdf of Y is

$$p_Y(y) = \begin{cases} 1/L & 0 < y < L \\ 0 & \text{otherwise} \end{cases}$$

and the conditional pdf of X given Y is

$$p_{X|Y}(x|y) = \begin{cases} 1/y & 0 < x < y \\ 0 & \text{otherwise} \end{cases}.$$

Thus, for  $x \in [0, L]$ ,

$$p_X(x) = \int p_{X|Y}(x|y)p_Y(y)dy$$
  
=  $\frac{1}{L} \int_x^L \frac{dy}{y}$   
=  $\frac{1}{L} \ln\left(\frac{L}{x}\right),$ 

and for all other x we have  $p_X(x) = 0$ . Thus,

$$E(X) = L \int_0^L \frac{x}{L} \ln\left(\frac{L}{x}\right) \frac{dx}{L} = -L \int_0^1 u \ln u \, du = -L \cdot \left(\frac{1}{2}u^2 \ln u - \frac{1}{4}u^2\right) \Big|_0^1 = \frac{1}{4}L,$$

and also

$$E(X^2) = L^2 \int_0^L \frac{x^2}{L^2} \ln\left(\frac{L}{x}\right) \frac{dx}{L} = -L^2 \int_0^1 u^2 \ln u \, du = -L^2 \cdot \left(\frac{1}{3}u^3 \ln u - \frac{1}{9}u^3\right) \Big|_0^1 = \frac{1}{9}L^2,$$

so that

$$\operatorname{var}(X) = E(X^2) - [E(X)]^2 = \frac{L^2}{9} - \frac{L^2}{16} = \frac{7L^2}{144}$$

4. Let n > 3.

(a) Suppose a fair coin is flipped n times. Let  $p_n$  be the probability that there are at most 3 heads. Find  $p_n$  as a simple function of n.

- (b) Suppose a fair coin is flipped until heads appears 3 times. Let  $q_n$  be the probability that it takes at least n flips (including the flip resulting in the 3rd head). Find  $q_n$  as a simple function of n.
- (c) Which of the following holds:  $p_n > q_n$ ,  $p_n = q_n$ ,  $p_n < q_n$ ? Justify your answer.

#### Solution:

(a) The number of heads X is a binomial random variable. Adding together the probability that X = 0, 1, 2, 3, we obtain

$$p_n = \frac{1}{2^n} \left[ 1 + n + \binom{n}{2} + \binom{n}{3} \right]$$

(b) Consider the first *n* flips of the coin. (We can imagine that we continue to flip the coin.) The specified event occurs if and only if (i) there are at most 2 heads in the first *n* flips, or (ii) there are are 3 heads in the first *n* flips with the last one occurring in the *n*th flip. The last case occurs with probability  $\frac{1}{2^{n-1}} {\binom{n-1}{2}} \cdot \frac{1}{2}$ . Thus

$$q_n = \frac{1}{2^n} \left[ 1 + n + \binom{n}{2} + \binom{n-1}{2} \right]$$

- (c) By direct comparison, noting  $\binom{n}{3} = \frac{(n-1)(n-2)}{2} \cdot \frac{n}{3} > \binom{n-1}{2}$ , we have  $p_n > q_n$ . Alternatively, the event in (b) is strictly contained in the event in (a), with the set difference being the event of positive probability that there are 3 heads in the first n-1 flips and the *n*th flip is a tail.
- 5. Let 0 and let X be a random variable such that

$$P(X = 1) = p, \quad P(X = -1) = 1 - p.$$

Let Y be a random variable whose conditional distribution given X is normal with mean X and variance 4, that is,  $Y|X \sim N(X, 4)$ .

- (a) Conditioned on Y = y, what is the distribution of X? Simplify your answer.
- (b) What is the variance of Y?

Solution:

(a) By Bayes's rule, letting  $p_{Y|X}$  be the conditional density of Y given X we have

$$P(X = 1|Y = y) = \frac{P(X = 1)p_{Y|X}(y|X = 1)}{P(X = 1)p_{Y|X}(y|X = 1) + P(X = -1)p_{Y|X}(y|X = -1)}$$
$$= \frac{pe^{-\frac{(y-1)^2}{8}}}{pe^{-\frac{(y-1)^2}{8}} + (1-p)e^{-\frac{(y+1)^2}{8}}} = \frac{p}{p+(1-p)e^{-\frac{y}{2}}}$$
$$P(X = -1|Y = y) = \frac{P(X = -1)p_{Y|X}(y|X = -1)}{P(X = 1)p_{Y|X}(y|X = 1) + P(X = -1)p_{Y|X}(y|X = -1)}$$
$$= \frac{(1-p)e^{-\frac{(y+1)^2}{8}}}{pe^{-\frac{(y-1)^2}{8}} + (1-p)e^{-\frac{(y+1)^2}{8}}} = \frac{1-p}{pe^{\frac{y}{2}} + (1-p)}.$$

(b) Note that

$$\operatorname{var}(X) = E(X^2) - (EX)^2 = 1 - (p - (1 - p))^2 = 1 - (2p - 1)^2 = 4p(1 - p)$$

and that Y has the same distribution as X + Z where Z is normal with mean 0 and variance 4, independent of X. By additivity of variance,

$$\operatorname{var}(Y) = 4p(1-p) + 4.$$

Alternatively, using the explicit formula for the density

$$p_Y(y) = \frac{1}{\sqrt{8\pi}} \left[ p e^{-\frac{(y-1)^2}{8}} + (1-p) e^{-\frac{(y+1)^2}{8}} \right],$$

one can calculate that

$$E(Y) = p - (1 - p) = 2p - 1, \quad E(Y^2) = (4 + 1)p + (4 + 1)(1 - p) = 5$$

and

$$\operatorname{var}(Y) = E(Y^2) - (EY)^2 = 5 - (4p^2 - 4p + 1) = 4 + 4p(1 - p).$$

6. Let X, Y be random variables whose joint density function is

$$p(x,y) = 2e^{x-2y}, \quad x \ge 0, \ y \ge x.$$

Let U, V be independent random variables whose distributions are uniform over (0,1). Find an explicit continuous function  $f: (0,1)^2 \to \mathbb{R}^2$  such that f(U,V) has the same distribution as (X,Y).

Solution: Rewriting

$$p(x,y) = 2e^{-x}e^{-2(y-x)}, \quad x \ge 0, \ y \ge x$$

and noting that the transformation T(x, y) = (x, y - x) is volume-preserving (the Jacobian is  $\begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$ ), we see that X and Y - X are independent exponential random variables with rates 1 and 2, respectively. The cumulative distribution function of an exponential random variable with rate  $\lambda$  is  $F_{\lambda}(x) = 1 - e^{-\lambda x}$  for  $x \ge 0$  and  $F_{\lambda}(x) = 0$  for x < 0. By inverse transform sampling, for a random variable Z with continuous increasing cdf  $F : [0, \infty] \to \mathbb{R}$ , Z has the same distribution as  $F^{-1}(U)$  where U is uniform in [0, 1]. This is because for  $z \ge 0$ ,

$$F(z) = P(U \le F(z)) = P(F^{-1}(U) \le z)$$

Note U and 1-U have the same distribution, so alternatively,  $Z \stackrel{d}{=} F^{-1}(1-U)$ . This means that X has the same distribution as  $F_1^{-1}(1-U) = -\ln U$  and Y - X has the same distribution as  $F_2^{-1}(1-V) = -\frac{1}{2}\ln V$ . Because X and Y - X are independent,

$$(X, Y - X) \stackrel{\mathrm{d}}{=} \left( -\ln U, -\frac{1}{2}\ln V \right),$$

or, equivalently,

$$(X,Y) \stackrel{\mathrm{d}}{=} \left(-\ln U, -\ln U - \frac{1}{2}\ln V\right).$$

Thus we can choose  $f(u, v) = \left(-\ln u, -\ln u - \frac{1}{2}\ln v\right)$ .

# Department of Applied Mathematics and Statistics The Johns Hopkins University

## INTRODUCTORY EXAMINATION–SPRING SEMESTER LINEAR ALGEBRA

## THURSDAY, JANUARY 18, 2024

## 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. Consider the two real  $3 \times 3$  matrices

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 5 \\ 0 & 5 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -6 & 1 \\ -6 & 0 & 0 \\ 1 & 0 & 3 \end{pmatrix}.$$

Show that there is no basis for  $\mathbb{R}^3$  which diagonalizes A and B simultaneously.

Solution: The existence of such a basis would imply that A and B commute, but

$$AB = \begin{pmatrix} 6 & -6 & 1 \\ -7 & 6 & 14 \\ -30 & 0 & 0 \end{pmatrix}, \quad BA = \begin{pmatrix} 6 & -7 & -30 \\ -6 & 6 & 0 \\ 1 & 14 & 0 \end{pmatrix}.$$

2. If A is a square matrix with all eigenvalues real, prove that the matrix  $B = I + A + \frac{1}{2}A^2$  is invertible.

Hint: Consider the Jordan canonical form of A.

Solution: The polynomial spectral theorem (an easy consequence of Jordan canonical form) implies that the eigenvalues of B are exactly of the form  $\mu = 1 + \lambda + \frac{1}{2}\lambda^2$  for eigenvalues  $\lambda$  of A. However, the polynomial  $x^2 + 2x + 2$  has only the complex roots  $-1 \pm i$ , and thus  $\mu = 1 + \lambda + \frac{1}{2}\lambda^2 \neq 0$  for all real  $\lambda$ .

Alternative statement of the solution:

The matrix A has a Jordan canonical form, i.e., it is possible to write  $SAS^{-1}$  as a matrix consisting of diagonal blocks of the form

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ 0 & 0 & \lambda & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \cdots & \lambda \end{bmatrix}$$

with the possibility of a  $1 \times 1$  block  $[\lambda]$  for each distinct eigenvalue, for some invertible matrix S.

It follows that  $S(I + A + \frac{1}{2}A^2)S^{-1} = I + SAS^{-1} + \frac{1}{2}SAS^{-1}SAS^{-1}$  which is a matrix of diagonal blocks, one for each block in the Jordan form of A, where the block corresponding to eigenvalue  $\lambda$  is upper triangular with diagonal entry  $1 + \lambda + \frac{1}{2}\lambda^2$ . It follows that  $S(I + A + \frac{1}{2}A^2)S^{-1}$  is upper triangular with these entries on the diagonal. These diagonal entries are all nonzero since  $1 + \lambda + \frac{1}{2}\lambda^2$  is non-zero for all real values of  $\lambda$ . So the matrix  $S(I + A + \frac{1}{2}A^2)S^{-1}$ , and hence also  $I + A + \frac{1}{2}A^2$ , is invertible. 3. If W is the  $n \times n$  matrix all of whose entries are ones and if I is the  $n \times n$  identity matrix, find the inverse matrix  $(I + W)^{-1}$ .

Solution: # 1: A simple calculation gives  $W^2 = nW$ . Guessing that  $(I + W)^{-1} = I + \alpha W$  for scalar  $\alpha$ , direct multiplication gives

$$(I + \alpha W)(I + W) = I + (n\alpha + \alpha + 1)W$$

so that  $\alpha = -1/(n+1)$  gives the required inverse.

Solution # 2: Note that  $W = ww^{\top}$  where w is the *n*-dimensional column vector with all components = 1. We can thus apply the Sherman–Morrison formula  $(A+uv^{\top})^{-1} = A^{-1} - \frac{A^{-1}uv^{\top}A^{-1}}{1+v^{\top}A^{-1}u}$  to obtain  $(I + ww^{\top})^{-1} = I - \frac{ww^{\top}}{1+w^{\top}w} = I - \frac{1}{1+n}W$  using  $w^{\top}w = n$ . Solution # 3: We can express W as

$$W = V \begin{bmatrix} n & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} V^T$$

where V is orthonormal, specifically

$$V = \begin{bmatrix} 1/\sqrt{n} & | & | & | \\ \vdots & v_2 & v_3 & \dots & v_n \\ 1/\sqrt{n} & | & | & | \end{bmatrix},$$

with  $\frac{1}{\sqrt{n}}\mathbf{1}, v_2, \ldots, v_n$  orthonormal basis vectors. Hence

$$I + W = V \begin{bmatrix} n+1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} V^{T},$$

which implies

$$(I+W)^{-1} = V \begin{bmatrix} 1/(n+1) & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} V^{T} = \frac{1}{(n+1)n} \mathbf{1} \mathbf{1}^{T} + v_{2} v_{2}^{T} + \dots + v_{n} v_{N}^{T}$$

Note that  $\mathbf{11}^T = W$ , and letting  $v_1 = [1/\sqrt{n} \dots 1/\sqrt{n}]^T$  we get  $v_2 v_2^T + \dots + v_n v_n^T = I - v_1 v_1^T = I - \frac{1}{n} W$ . Finally,

$$(I+W)^{-1} = \frac{1}{(n+1)n}W + I - \frac{1}{n}W = I - \frac{n}{(n+1)n}W = I - \frac{1}{n+1}W.$$

4. If A is a linear map from a vector space S to a vector space T with  $\dim(S) > \dim(T)$ , then prove that the subspace of vectors  $x \in S$  such that Ax = 0 has dimension at least  $\dim(S) - \dim(T)$ .

Solution: The subspace in question is  $\operatorname{Ker} A = (\operatorname{Ran} A^*)^{\perp}$  where  $A^* : T \to S$  is the adjoint linear map. Since  $\operatorname{dim}(\operatorname{Ran} A^*) \leq \operatorname{dim} T$ , we have  $\operatorname{dim}(\operatorname{Ker} A) = \operatorname{dim}((\operatorname{Ran} A^*)^{\perp}) = \operatorname{dim}(S) - \operatorname{dim}(\operatorname{Ran} A^*) \geq \operatorname{dim}(S) - \operatorname{dim}(T)$ .

Solution # 2: The rank–nullity theorem gives

$$\dim \ker(A) + \dim \operatorname{im}(A) = \dim(S),$$

and since im(A) is a subspace of T its dimension is at most dim(T). So

 $\dim \ker(A) = \dim(S) - \dim \operatorname{im}(A) \ge \dim(S) - \dim(T).$ 

5. Assume that u, v are two vectors in  $\mathbb{C}^n$  and consider the following possible assignments of Euclidean norms:

(i) 
$$||u+v|| = 2$$
,  $||u-v|| = 2$ ,  $||u+iv|| = 3$ ,  $||u-iv|| = 3$   
(ii)  $||u+v|| = 2$ ,  $||u-v|| = 2$ ,  $||u+iv|| = 2$ ,  $||u-iv|| = 2$ .

For each assignment, either prove that it is impossible or find an illustrative example.

Solution: From the complex polarization identity

$$\langle u, v \rangle = \frac{1}{4} \left[ \|u + v\|^2 - \|u - v\|^2 - i\|u + iv\|^2 + i\|u - iv\|^2 \right]$$

we see that the first condition (i) implies that u, v are orthogonal. However, in that case all of the given norms must be equal to  $\sqrt{\|u\|^2 + \|v\|^2}$ , so that assignment (i) is inconsistent. Condition (ii) implies as well orthogonality of u, v, but in that case any orthogonal vectors separately normalized to  $\|u\| = \|v\| = \sqrt{2}$  will illustrate (ii).

6. If A is a Hermitian, positive-definite  $n \times n$  matrix, B is an  $n \times m$  matrix for  $m \leq n$  with full rank, and  $O_m$  is the  $m \times m$  zero matrix, then prove that the  $(n+m) \times (n+m)$  matrix C defined by

$$C = \begin{pmatrix} A & B \\ B^* & O_m \end{pmatrix}$$

is Hermitian with all eigenvalues non-zero.

Solution: Elementary calculation gives  $C^* = C$ . Supposing that C has eigenvalue 0, then let  $z = (x, y)^{\top}$  be a corresponding eigenvector for *n*-dimensional x and *m*-dimensional y. Then Cz = 0 is equivalent to

$$Ax + By = 0, \quad B^*x = 0.$$

These equations give

$$x^*Ax = -x^*By = -(B^*x)^*y = 0.$$

Since A is positive-definite, the last equality implies x = 0 and 0 = Ax + By = By. However, because B has full rank (all m columns linearly independent), then By = 0implies y = 0. Since Cz = 0 implies z = 0, there is no 0 eigenvalue.