MOCK EXAM
Linear Algebra and Probability
Spring 2019

IT University of Copenhagen

May 2019

Instructions

- This is a 4 hour written exam with all written and printed aids allowed (A2)
- There are 6 problems and 4 pages.
- Each problem is divided into questions
- The point value of each problem and the distribution of points among questions is given explicitly.
- There is a total of 100 points to be earned.
- The problems are formulated in english and should be answered in english.
- Read each question completely before trying to solve it.
- **Make sure** you do all actions marked with bold font
- Please **order** and **number** the pages before handing in.
- Solutions should be hand-written
- Access to aid in the form of books, own notes, e-books, also on laptops and iPads is permitted.
- Use of internet including email and social media is not permitted.
- Use of any other hardware or software such as MatLab or pocket calculators is not permitted
- Any form of communication between students or with the outside world is not permitted.

1 Matrices (25 points)

For the first question, consider the matrices

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -3 & 0 \\ 2 & 1 & 5 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 3 & 0 & 1 & 1 \\ 2 & 1 & -1 & 3 \\ 1 & 0 & -1 & -3 \end{bmatrix}$$

a) [9 points] Compute the determinants of each of the matrices A, B, C and determine which of them are invertible.

For the next two subquestions, consider the matrix

$$D = \begin{bmatrix} 1 & 0 & 2 & 1 & 1 \\ 2 & 1 & 6 & 4 & 2 \\ 2 & 1 & 7 & 6 & 1 \\ 3 & 0 & 7 & 5 & 2 \end{bmatrix}$$

b) [9 points] Compute the rank of D, and the dimensions of the following subspaces associated to D: The column space, the row space, and the nullspace.

c) [7 points] Compute a basis for the nullspace of D.

Solution

The determinants are $|A| = 2$, $|B| = -15$, $|C| = 0$. The latter can be seen e.g. by expanding along the second row: The determinant of C is -1 times that of

$$\begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 1 \\ 1 & -1 & -3 \end{bmatrix}$$

Thus, A and B are invertible, but C is not.

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 1 \\ 2 & 1 & 6 & 4 & 2 \\ 2 & 1 & 7 & 6 & 1 \\ 3 & 0 & 7 & 5 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 2 & 2 & 0 \\ 0 & 1 & 3 & 4 & -1 \\ 0 & 0 & 1 & 2 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 2 & 2 & 0 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & 2 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 2 & 2 & 0 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & -3 & 3 \\ 0 & 1 & 0 & -2 & 2 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The rank is 3, dimension of column space and row space are 3, dim of nullspace is $5 - 3 = 2$.
The nullspace is spanned by $(3, 2, -2, 1, 0)$ and $(-3, -2, 1, 0, 1)$.

2 Eigenvectors and eigenvalues (15 points)

Consider the matrix

$$A = \begin{bmatrix} 8 & 1 \\ 2 & 9 \end{bmatrix}$$

a) [10 points] Find the eigenvalues and eigenvectors of A.

b) [5 points] Find matrices P and D such that D is diagonal and $A = PDP^{-1}$.

Solution

The characteristic polynomial is $(\lambda - 8)(\lambda - 9) - 2 = \lambda^2 - 17\lambda + 70$ with roots 7 and 10. The corresponding eigenvectors are (any non-zero scalar multiplum of) $(1, -1)$ for 7 and $(1, 2)$ for 10.

$$D = \begin{bmatrix} 10 & 0 \\ 0 & 7 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$$

3 Bases (15 points)

Consider the polynomials

$$\begin{aligned} p(x) &= x^2 + 2 \\ q(x) &= 3x^2 + x - 1 \\ r(x) &= 2x^2 + 4x + 1 \end{aligned}$$

a) [10 points] Show that $B = \{p, q, r\}$ forms a basis for the vector space P_2 of polynomials of degree at most 2.

b) [5 points] Compute the coordinates of the polynomial $s(x) = 5x^2 + 11x + 8$ relative to the basis B.

Solution

The vectors span P_2 if and only if each polynomial $c_2x^2 + c_1x + c_0$ can be written as a linear combination of p, q, r in a unique way, i.e., if there exist unique numbers a, b, c such that

$$c_2x^2 + c_1x + c_0 = ap + bq + cr$$

Since two polynomials are equal if and only if they have the same coefficients, this equation reduces to a system of linear equations:

$$\begin{array}{rcrcrcl} a & + & 3b & + & 2c & = & c_2 \\ & & b & + & 4c & = & c_1 \\ 2a & - & b & + & c & = & c_0 \end{array}$$

This has a unique solution for all c_2, c_1, c_0 if and only if the coefficient matrix

$$\begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 4 \\ 2 & -1 & 1 \end{bmatrix}$$

is invertible. The determinant is (computed using the diagonals method)

$$1 + 24 + 0 - 4 + 4 - 0 = 25 \neq 0$$

so the matrix is indeed invertible, and the system is a basis.

In the second subquestion, the coordinates are the unique vector

$$[s]_B = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

such that $s = ap + bq + cr$. These can be computed by solving the system given by the augmented matrix below.

$$\left[\begin{array}{ccc|c} 1 & 3 & 2 & 5 \\ 0 & 1 & 4 & 11 \\ 2 & -1 & 1 & 8 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 3 & 2 & 5 \\ 0 & 1 & 4 & 11 \\ 0 & -7 & -3 & -2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 3 & 2 & 5 \\ 0 & 1 & 4 & 11 \\ 0 & 0 & 25 & 75 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 3 & 2 & 5 \\ 0 & 1 & 4 & 11 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 3 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

so that

$$[s]_B = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$$

4 Conditional probability (15 points)

Among the participants of Tour de Amager bikerace, it is estimated that 40% of all the participants use the illegal substance EPO. The tour management uses an imperfect drugtest that returns a positive results (indicating that the test subject was using EPO) with a probablity of 90% if the tested rider did in fact use EPO, and 20% if the tested rider did not use EPO.

a) [8 points] The winner Ulrich Stronglegs turns in a positive test (indicating EPO was used). What is the probability that Ulrich did in fact use EPO?

b) [7 points] The tour management considers testing all participants, but is worried about the number of false positives (positive tests on riders that did not take EPO). Give a formula for the distribution of false positives among a population of 80 riders.

Solution

a) Let E be the event that Ulrich used EPO and let T be the event that the drug test is positive. The question states that

$$P(E) = .4 \qquad P(T \mid E) = .9 \qquad P(T \mid E^c) = .2$$

Thus

$$P(T) = P(T \mid E)P(E) + P(T \mid E^c)P(E^c) = .9 \cdot .4 + .2 \cdot .6 = .36 + .12 = .48$$

and so by Bayes rule

$$P(E \mid T) = \frac{P(E)P(T \mid E)}{P(T)} = \frac{.36}{.48} = \frac{3}{4}$$

which is the answer to the question.

b) The probability that a rider gives a false positive is

$$P(T \mid E^c) \times P(E^c) = .2 \cdot .6 = .12$$

Thus the PMF for the random variable X giving the number of false negatives follows a binomial distribution

$$P(X = k) = \binom{80}{k} (.12)^k (.88)^{80-k}$$

5 Normal random variables (15 points)

Suppose X and Y are independent normal random variables with mean $\mu_X = 1$ and $\mu_Y = 2$ and standard deviations $\sigma_X = 3$ and $\sigma_Y = 2$ respectively.

a) [8 points] Compute the mean and standard deviation of the random variable $Z = X - 2Y$.

b) [7 points] Give a probability density function for $X - 2Y$.

Solution

a)

$$\mathsf{E}[Z] = \mathsf{E}[X] - 2\mathsf{E}[Y] = 1 - 4 = -3$$

$$\mathsf{Var}(Z) = \mathsf{Var}(X) + \mathsf{Var}(-2Y) = \mathsf{Var}(X) + 4\mathsf{Var}(Y) = 9 + 16 = 25$$

using the independence in the first equality. So the standard deviation is 5.

b) Since Z is a linear combination of normal distributions, it is itself normally distributed, so the PDF is determined from its mean and variance as

$$f_Z(z) = \frac{1}{5\sqrt{2\pi}} e^{\frac{-(z+3)^2}{50}}$$

6 Normal distributions and the central limit theorem (15 points)

Suppose the caffeine contents in an espresso from Cafe Analog is known to follow a normal distribution with standard deviation 5, but the average is unknown. Suppose the caffeine contents of 10 cups of espresso are measured, and the result is an average of 77 mg.

a) [5 points] Assuming the average caffeine contents measured is correct, what is the probability that a given cup of espresso contains less than 70 mg of caffeine?

b) [10 points] Suppose now we want to determine the average caffeine contents of an espresso up to an error of 1 mg with confidence 95%. Use the central limit theorem to approximate the number of espressos needed to be measured.

Solution

a) By symmetry of the normal distribution around the mean, the probability of getting less than 70 mg is the same as that of getting more than 84, which is $\frac{7}{5} = 1.4$ standard deviations from the mean. Thus

$$P(X \leq 70) = P(X \geq 84) = 1 - P(X \leq 84) = 1 - \Phi(1.4) = 1 - 0.9192 = 0.0808$$

b) Let M_n be the sample mean, so $M_n = \frac{S_n}{n}$. By the central limit theorem, M_n is almost symmetric around its mean so if μ is the actual mean we get

$$\begin{aligned} P(|M_n - \mu| \geq 1) &\approx 2P(M_n - \mu \geq 1) \\ &= 2P(S_n \geq n + n\mu) \end{aligned}$$

Applying the corollary to the central limit theorem we compute

$$z = \frac{n + n\mu - n\mu}{5\sqrt{n}} = \frac{\sqrt{n}}{5}$$

and compute

$$\begin{aligned} 2P(S_n \geq n + n\mu) &= 2(1 - P(S_n \leq n + n\mu)) \\ &\approx 2(1 - \Phi(z)) \end{aligned}$$

Since this number needs to be smaller than 0.05, we should have $\Phi(z) \geq 0.975$, i.e., $z \geq 1.96$. This means

$$\sqrt{n} \geq 5 \cdot 1.96$$

i.e., $n \geq 100$.