DIAGNOSTIC IN-CLASS QUIZ: DUE WEDNESDAY OCTOBER 2: VECTORS (BASIC STUFF)

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____

PLEASE DO *NOT* DISCUSS ANY QUESTIONS.

Many of you are familiar with vectors, either from Math 195 or some exposure to vectors in high school (or perhaps both). This quiz is to help gauge your level of understanding coming in. We will not get to start using the ideas in their full depth until a few weeks later.

For the benefit of those who haven't seen vectors at all, the definitions are briefly provided.

There are many ways of describing the vector in $\mathbb{R}^n$ with coordinates a_1, a_2, $\dots$, a_n. You may have seen the vector described using angled braces as $\langle a_1, a_2, \dots, a_n \rangle$. In this linear algebra course, we will customarily write the vector as a *column* vector, i.e., the coordinates will be written in a vertical column. For instance, the vector $\langle 2, 3, 7 \rangle$ will be written as the column vector $\begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix}$.

Two vectors in $\mathbb{R}^n$ can be added with each other (note that both vectors need to be in the *same* $\mathbb{R}^n$ in order to be added). The addition is coordinate-wise:

$$\begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ \cdot \\ \cdot \\ \cdot \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n + w_n \end{bmatrix}$$

Also, given any real number λ (called a *scalar* to distinguish from a vector) and a vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{bmatrix}$, we can define:

$$\lambda \vec{v} = \lambda \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{bmatrix} := \begin{bmatrix} \lambda v_1 \\ \lambda v_2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda v_n \end{bmatrix}$$

We can identify the set of n-dimensional vectors with the set of points in $\mathbb{R}^n$. The vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_n \end{bmatrix}$ in this case corresponds to the point with coordinates $(v_1, v_2, \dots, v_n)$.

(1) *Do not discuss this!*: For a n-dimensional vector $\vec{v}$, the *set of scalar multiples* of $\vec{v}$ is the set of vectors that can be expressed in the form $\lambda\vec{v}$, $\lambda \in \mathbb{R}$. Assume that $\vec{v}$ is a nonzero vector. What can we say geometrically about the set of points in $\mathbb{R}^n$ that correspond to the scalar multiples of $\vec{v}$?

(A) It is a straight line in $\mathbb{R}^n$ that passes through the origin.
(B) It is a straight line in $\mathbb{R}^n$. However, it need not pass through the origin.
(C) It is a straight half-line in $\mathbb{R}^n$ with the endpoint at the origin.
(D) It is a straight half-line in $\mathbb{R}^n$, but the endpoint need not be at the origin.
(E) It is a line segment in $\mathbb{R}^n$.

Your answer: ____________________________

(2) *Do not discuss this!*: Given two n-dimensional vectors $\vec{v}$ and $\vec{w}$, the *set of linear combinations* of $\vec{v}$ and $\vec{w}$ is the set of all vectors that can be written in the form $\lambda\vec{v} + \mu\vec{w}$ where $\lambda, \mu \in \mathbb{R}$ (note that λ and μ can take arbitrary real values, and are allowed to be equal to each other). In other words, you can take scalar multiples, and you can then add these scalar multiples.

The set of linear combinations of $\vec{v}$ and $\vec{w}$ is sometimes also called the *span* of $\vec{v}$ and $\vec{w}$.

What is the span of the vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in $\mathbb{R}^2$?

(A) The zero vector only, because that is the only vector that can be expressed both as a multiple of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and as a multiple of $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
(B) The set of vectors that can be expressed as a scalar multiple either of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ or of $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
(C) The set of vectors that can be expressed as a scalar multiple of at least one of these three vectors: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
(D) All vectors in the first quadrant of $\mathbb{R}^2$, including the bounding half-lines. In other words, the set of vectors of the form $\begin{bmatrix} x \\ y \end{bmatrix}$ where $x \geq 0$ and $y \geq 0$.
(E) All vectors in $\mathbb{R}^2$.

Your answer: ____________________________

(3) *Do not discuss this!*: Consider the transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ that interchanges the coordinates of a vector. Explicitly, the transformation is given as:

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} y \\ x \end{bmatrix}$$

Which of the following describes the transformation geometrically, with $\mathbb{R}^2$ viewed as the xy-plane?

(A) It is a reflection about the x-axis in $\mathbb{R}^2$, i.e., the axis for the first coordinate.
(B) It is a reflection about the y-axis in $\mathbb{R}^2$, i.e., the axis for the second coordinate.
(C) It is a reflection about the line $y = x$ in $\mathbb{R}^2$, i.e., the line of vectors where both coordinates are equal.
(D) It is a reflection about the line $y = -x$ in $\mathbb{R}^2$, i.e., the line of vectors where the coordinates are negatives of each other.

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE FRIDAY OCTOBER 4: LINEAR FUNCTIONS AND EQUATION-SOLVING (PART 1)

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____

PLEASE DO *NOT* DISCUSS ANY QUESTIONS EXCEPT THE STARRED OR DOUBLE-STARRED QUESTIONS. YOU CAN DISCUSS THE STARRED AND DOUBLE-STARRED QUESTIONS.

This quiz covers some basics involving linear functions and equation-solving (notes at `Linear functions: a primer` and `Equation-solving with a special focus on the linear case`). The quiz tests for the following:

- What it means to be (affine) linear, and in particular, the significance of the intercept as an additional parameter to track.
- The distinction between behavior relative to the variables (the inputs) and behavior relative to the parameters.
- Using the linear paradigm to study functional forms that are not themselves linear.
- A small taste of dealing with measurement uncertainty to obtain upper and lower bounds (not covered in the notes, so this is where your famed ability to think out of the box should manifest).
- Solving "triangular" systems of equations.

(1) (*) A function f of 3 variables x, y, z defined everywhere is (affine) linear in the variables. (The "affine" is to indicate that the intercept may be nonzero). Based on the above information and some input-output pairs for f, we would like to determine f uniquely. What is the minimum number of input-output pairs that we would need in order to achieve this?
(A) 1
(B) 2
(C) 3
(D) 4
(E) 5

Your answer: ____________________________

(2) *Do not discuss this!*: Which of the following gives an example of a function F of three variables x, y, z whose third-order mixed partial derivative F_{xyz} is zero everywhere, but for which none of the second-order mixed partial derivatives F_{xy}, F_{xz}, F_{yz} is zero everywhere?
(A) $\sin(xy) - z^2$
(B) $\cos(x^2 + y^2) - \sin(y^2 + z^2)$
(C) $e^{xy} + (y - z)^2 + 3xz$
(D) $x^2 + y^2 + z^2$
(E) xyz

Your answer: ____________________________

(3) *Do not discuss this!*: Consider a function of the form $F(x,y) := Ca^x b^y$ where C, a, b are all positive reals that serve as parameters and x, y are restricted to the positive reals. We wish to study F using the paradigm of linear functions. What is the best way of doing this?
(A) Express $\ln(F(x,y))$ in terms of $\ln x$ and $\ln y$
(B) Express $\ln(F(x,y))$ in terms of x and y
(C) Express $F(x,y)$ in terms of $\ln x$ and $\ln y$
(D) Express $\ln(F(x,y))$ in terms of a^x and b^y
(E) Express $F(x,y)$ in terms of a^x and b^y

Your answer: ____________________________

(4) *Do not discuss this!*: Consider a function of the form $F(x,y) := Cx^a y^b$ where C, a, b are all positive reals that serve as parameters and x, y are restricted to the positive reals. We wish to study F using the paradigm of linear functions. What is the best way of doing this?
(A) Express $\ln(F(x,y))$ in terms of $\ln x$ and $\ln y$
(B) Express $\ln(F(x,y))$ in terms of x and y
(C) Express $F(x,y)$ in terms of $\ln x$ and $\ln y$
(D) Express $\ln(F(x,y))$ in terms of x^a and y^b
(E) Express $F(x,y)$ in terms of x^a and y^b

Your answer: ____________________________

(5) (**) *This is a hard question!* The population in the island of Andrognesia as a function of time is believed to be an exponential function. On January 1, 1984, the population was measured to be $3 * 10^5$ with a measurement error of up to 10^5 on either side, i.e., the population was measured to be between $2 * 10^5$ and $4 * 10^5$. On January 1, 1998, the population was measured to be $1.2 * 10^6$ with a measurememt error of up to $4 * 10^5$ on either side, i.e., the population was measured to be between $8 * 10^5$ and $1.6 * 10^6$. If the population is an exponential function of time (i.e., the increment in population per year is a fixed proportion of the population that year), what is the **range of possible values** of the population measured on January 1, 2012? *Hint: Think of the umbral versus penumbral region for an eclipse*
(A) Between $3.2 * 10^6$ and $6.4 * 10^6$
(B) Between $3.2 * 10^6$ and $1.28 * 10^7$
(C) Between $1.6 * 10^6$ and $3.2 * 10^6$
(D) Between $1.6 * 10^6$ and $6.4 * 10^6$
(E) Between $1.6 * 10^6$ and $1.28 * 10^7$

Your answer: ____________________________

(6) *Do not discuss this!*: Suppose, according to our model, a particular function $f(x,y)$ is of the form $f(x,y) = a_1 + a_2x + a_3y + a_4x^2y^2$ where a_1, a_2, a_3, a_4 are parameters. Our goal is to determine the values of the parameters a_1, a_2, a_3, a_4. We do this by collecting a number of (input,output) pairs for the function f and then setting up equations in terms of the parameters using the (input,output) pairs. What can we say about the nature of f and the nature of the system of equations that we will need to solve? *Note that "nonlinear" as used here simply means that the expression is not guaranteed*

to be linear, though it may turn out to be linear in some cases. Similarly, "non-polynomial" means not guaranteed to be polynomial, though it may turn out to be polynomial in some cases.

(A) f is a linear function of x and y, hence we need to solve a linear system of equations to determine the parameters a_1, a_2, a_3, a_4.

(B) f is a nonlinear polynomial function of x and y, hence we need to solve a nonlinear polynomial system of equations to determine the parameters a_1, a_2, a_3, a_4.

(C) f is a linear function of x and y. However, we need to solve a nonlinear polynomial system of equations to determine the parameters a_1, a_2, a_3, a_4.

(D) f is a nonlinear polynomial function of x and y. However, we need to solve a linear system of equations to determine the parameters a_1, a_2, a_3, a_4.

(E) f is a nonlinear polynomial function of x and y. However, we need to solve a non-polynomial system of equations to determine the parameters a_1, a_2, a_3, a_4.

Your answer: ____________________

(7) *Do not discuss this!*: Consider the system of equations:

$$\begin{aligned} x^2 - x &= 2 \\ y^2 + xy &= x + 13 \end{aligned}$$

What is the number of solutions to this system for real x and y?

(A) 0
(B) 2
(C) 4
(D) 6
(E) 8

Your answer: ____________________

(8) *Do not discuss this!*: Consider the system of equations:

$$\begin{aligned} x^2 - x &= 2 \\ y^2 + xy &= x + 13 \\ z^2 &= xy \end{aligned}$$

What is the number of solutions to this system for real x, y, and z?

(A) 0
(B) 2
(C) 4
(D) 6
(E) 8

Your answer: ____________________

(9) *Do not discuss this!*: Consider the system of equations:

$$\begin{aligned} x^2 - x &= 2 \\ y^2 + xy &= x + 13 \\ z^2 &= x^2 - y^2 \end{aligned}$$

What is the number of solutions to this system for real x, y, and z?

(A) 0
(B) 2
(C) 4
(D) 6
(E) 8

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE MONDAY OCTOBER 7: LINEAR FUNCTIONS AND EQUATION-SOLVING (PART 2)

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ______________________________________

PLEASE DO *NOT* DISCUSS ANY QUESTIONS EXCEPT THE STARRED OR DOUBLE-STARRED QUESTIONS.

This quiz covers some basics involving linear functions and equation-solving (notes at `Linear functions: a primer` and `Equation-solving with a special focus on the linear case`). The quiz tests for the following:

- The distinction between behavior relative to the variables (the inputs) and behavior relative to the parameters.
- Counting the number of parameters by creating the explicit general functional form from a verbal description (with a special focus on polynomial functional forms).
- Figuring out how to "ask the right questions" with respect to input choices, so that the answers provide meaningful information. This builds towards the ideas of hypothesis testing, rank, and overdetermination that we will see in the future.

(1) *Do not discuss this!*: Suppose f is a polynomial function of x of degree at most a *known number* n. What is the minimum number of (input,output) pairs that we need in order to determine f uniquely? *Extra information: Somewhat surprisingly, in this case, we do not need to be judicious about our input choices. Any set of distinct inputs of the required number will do. This has something to do with the "Vandermonde matrix" and "Vandermonde determinant" and is also related to the Lagrange interpolation formula.*
(A) $n-1$
(B) n
(C) $n+1$
(D) $2n$
(E) n^2

Your answer: ____________________________

(2) *Do not discuss this!*: f is a polynomial function of two variables x and y of total degree at most 2. In other words, for each monomial occurring in f, the total of the degrees of x and y in that monomial is at most 2. No other information is given about f. What is the minimum number of judiciously chosen (input,output) pairs we need in order to determine f uniquely?
(A) 2
(B) 3
(C) 4
(D) 6
(E) 7

Your answer: ____________________________

(3) *Do not discuss this!*: f is a polynomial function of two variables x and y of total degree at most 3. In other words, for each monomial occurring in f, the total of the degrees of x and y in that monomial is at most 3. No other information is given about f. What is the minimum number of judiciously chosen (input,output) pairs we need in order to determine f uniquely?
(A) 3
(B) 6
(C) 8
(D) 9
(E) 10

Your answer: ____________________________

(4) (*) *The perils of overfitting; see also Occam's Razor*: Suppose we are trying to model a function that we expect to behave in a polynomial-like manner, though we don't really have a good reason to believe this. Additionally, there is a possibility for measurement error in our observations. Our goal is to find the parameters so that we can both predict unmeasured values and do a qualitative analysis of the nature of the function and its derivatives and integrals.

We have a large number of observations (say, several thousands). We could attempt to "fit" the function using a polynomial of degree n for some fixed n using all those data points, and we will get a certain "best fit" that minimizes the deviation between the curve used for fitting and the function being fit. For instance, for $n = 1$, we are trying to find the best fit by a straight line function. For $n = 2$, we are trying to find the best fit by a polynomial of degree at most 2. We could try fitting using different values of n. Which of the following is true?

If you are interested in more on this, look up "overfitting". A revealing quote is by mathematician and computer scientist John von Neumann: "With four parameters I can fit an elephant. And with five I can make him wiggle his trunk." Another is by prediction guru Nate Silver: "The wide array of statistical methods available to researchers enables them to be no less fanciful and no more scientific than a child finding animal patterns in clouds."
(A) Larger values of n give better fits, therefore the larger the value of n we use, the better.
(B) Smaller values of n give better fits, therefore the smaller the value of n we use, the better.
(C) Larger values of n give better fits, therefore the larger the value of n we use, the less impressive a good fit (i.e., low deviation between the polynomial and the actual set of observations) should be.
(D) Smaller values of n give better fits, therefore the smaller the value of n we use, the less impressive a good fit (i.e., low deviation between the polynomial and the actual set of observations) should be.
(E) The value of n we use for trying to get a good fit is irrelevant. A good fit is a good fit, regardless of the type of function used.

Your answer: ____________________________

(5) (*) F is an affine linear function of two variables x and y, i.e., it has the form $F(x, y) := ax + by + c$ with a, b, and c real numbers. We want to determine the values of the parameters a, b, and c by using input-output pairs. It is, however, costly to find input-output pairs. We have already found $F(1, 3)$ and $F(3, 7)$. We want to find F for one other pair of inputs to determine a, b, and c. Which of these will *not* be a good choice?
(A) $F(2, 2)$, i.e., the input $x = 2$, $y = 2$
(B) $F(2, 3)$, i.e., the input $x = 2$, $y = 3$

(C) $F(2, 4)$, i.e., the input $x = 2$, $y = 4$
(D) $F(2, 5)$, i.e., the input $x = 2$, $y = 5$
(E) $F(2, 6)$, i.e., the input $x = 2$, $y = 6$

Your answer: ____________________________

DIAGNOSTIC IN-CLASS QUIZ: DUE FRIDAY OCTOBER 11: GAUSS-JORDAN ELIMINATION (ORIGINALLY DUE WEDNESDAY OCTOBER 9, BUT POSTPONED)

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____________________________________

PLEASE DO *NOT* DISCUSS ANY QUESTIONS

The quiz covers basics related to Gauss-Jordan elimination (notes titled `Gauss-Jordan elimination`, corresponding section in the book Section 1.2). Explicitly, the quiz covers:

- Setting up linear systems and interpreting the coefficient matrix in terms of the setup.
- Knowledge of the permissible rules for manipulating linear systems.
- Metacognition of the process of Gauss-Jordan elimination and its eventual result, the reduced row-echelon form, as well as the interpretation in terms of the solution set.

The questions are fairly easy questions if you understand the material. But it's important that you be able to answer them, otherwise what we study later will not make much sense.

(1) *Do not discuss this!*: The row operations that we can perform on the augmented matrix of a linear system include adding or subtracting another row. However, they do not include multiplying another row. In other words, suppose we start with:

$$\begin{bmatrix} 1 & 2 & | & 5 \\ 2 & 7 & | & 6 \end{bmatrix}$$

What we're not allowed to do is multiply row 2 by row 1 and obtain:

$$\begin{bmatrix} 1 & 2 & | & 5 \\ 2 & 14 & | & 30 \end{bmatrix}$$

What's the most compelling reason for our not being allowed to perform this operation?

(A) The row operations arise from the corresponding operations on equations. For the "multiplication of rows" operation to be legitimate, it must correspond to multiplication of the corresponding equations, and multiplying equations is not a legitimate operation.

(B) The row operations arise from the corresponding operations on equations. However, the "multiplication of rows" operation does not correspond to any legitimate operation on equations. Note that it does not correspond to multiplying the equations, because that is not how multiplication of linear polynomials work (in fact, if we multiplied the equations, we would end up with an equation that is not linear).

Your answer: ____________________________

(2) *Do not discuss this!*: Consider a model where the functional form is linear in the parameters (though not necessarily in the inputs). We can use (input, output) pairs to set up a system of linear equations in the parameters. Given enough such equations, we can determine the values of the parameters.

What is the relation between the coefficient matrix and the parameters and (input, output) pairs?

(A) The columns of the coefficient matrix correspond to the (input, output) pairs and the rows correspond to the parameters.

(B) The rows of the coefficient matrix correspond to the (input, output) pairs and the columns correspond to the parameters.

Your answer: ____________________________

(3) *Do not discuss this!*: Consider a model where the functional form is linear in the parameters (though not necessarily in the inputs). We can use (input, output) pairs to set up a system of linear equations in the parameters. Given enough such equations, we can determine the values of the parameters.

What is the relation between the inputs, the outputs, the coefficient matrix, and the augmenting column?

(A) The inputs correspond to the coefficient matrix and the outputs correspond to the augmenting column. In other words, knowing the values of the inputs allows us to write down the coefficient matrix. Knowing the values of the outputs allows us to write down the augmenting column.

(B) The outputs correspond to the coefficient matrix and the inputs correspond to the augment column. In other words, knowing the values of the outputs allows us to write down the coefficient matrix. Knowing the values of the inputs allows us to write down the augmenting column.

Your answer: ____________________________

(4) *Do not discuss this!*: Consider the following rule to check for consistency using the augmented matrix: the system is inconsistent if and only if there is a zero row of the coefficient matrix with a nonzero value for that row in the augmenting column. In what sense does this rule work?

(A) The rule can be applied to the augmented matrix directly in both the *if* and the *only if* direction.

(B) The rule can be applied to the augmented matrix only in the *if* direction in general. In the *only if* direction, the rule can be applied to the augmented matrix *after* we have reduced the system to a situation where the coefficient matrix is in row-echelon form (note: it's not necessary to reach reduced row-echelon form).

(C) The rule can be applied to the augmented matrix only in the *only if* direction in general. In the *if* direction, the rule can be applied to the augmented matrix *after* we have reduced the system to a situation where the coefficient matrix is in row-echelon form (note: it's not necessary to reach reduced row-echelon form).

(D) The rule can be applied in either direction only *after* we have reduced the system to a situation where the coefficient matrix is in row-echelon form (note: it's not necessary to reach reduced row-echelon form).

Your answer: ____________________________

(5) *Do not discuss this!*: Which of the following is *not* a possibility for the number of solutions to a system of simultaneous linear equations? Please see Options (D) and (E) before answering.

(A) 0
(B) 1
(C) 2
(D) All of the above, i.e., none of them is a possibility
(E) None of the above, i.e., they are all possibilities

Your answer: ____________________________

(6) *Do not discuss this!*: Which of the following describes the situation for a consistent system of simultaneous linear equations?

(A) The leading variables are the parameters used to describe the general solution, and the number of leading variables equals the number of nonzero equations in the reduced row-echelon form (here nonzero equation makes an equation that does not have a zero row in the augmented matrix).

(B) The non-leading variables are the parameters used to describe the general solution, and the number of non-leading variables equals the number of nonzero equations in the reduced row-echelon form (here nonzero equation makes an equation that does not have a zero row in the augmented matrix).

(C) The leading variables are the parameters used to describe the general solution, and the number of leading variables equals the value (number of variables) - (number of nonzero equations in the reduced row-echelon form) (here nonzero equation makes an equation that does not have a zero row in the augmented matrix).

(D) The non-leading variables are the parameters used to describe the general solution, and the number of non-leading variables equals the value (number of variables) - (number of nonzero equations in the reduced row-echelon form) (here nonzero equation makes an equation that does not have a zero row in the augmented matrix).

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE FRIDAY OCTOBER 11: LINEAR SYSTEMS

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____________________________________

YOU MAY DISCUSS ALL QUESTIONS, BUT PLEASE ENTER FINAL ANSWER OPTIONS THAT YOU ARE PERSONALLY MOST CONVINCED OF. BEWARE OF GROUPTHINK!

The quiz questions here, although not hard *per se*, are conceptually demanding: answering them requires a clear understanding of multiple concepts and an ability to execute them conjunctively. Even if you feel that you've understood the material as presented in class, you will need to think through each question carefully. Some of the questions are related to similar homework problems (Homeworks 1 and 2), and they test a conceptual understanding of the solutions to these problems. You might want to view them in conjunction with the homework problems. Other questions sow the seeds of ideas we will explore later. The quiz should seem relatively easier when you review it later, assuming that you work hard on attempting the questions right now and read the solutions once they're put up.

(1) (*) Rashid and Riena are trying to study a function f of two variables x and y. Rashid is convinced that the function is linear (i.e., it is of the form $f(x, y) := ax + by + c$) but has no idea what a, b, and c are. Riena thinks a linear model is completely out-of-place. Rashid is eager to find a, b, and c, whereas Riena is eager to disprove Rashid's linear model. Unfortunately, all they have is a black box that will output the value of the function for a given input pair (x, y), and that black box can only be called three times. What should Rashid and Riena try for?
(A) Rashid and Riena would both like to provide three input pairs that are non-collinear as points in the xy-plane
(B) Rashid would like to provide three input pairs that are non-collinear, while Riena would like to provide three input pairs that are collinear as points in the xy-plane.
(C) Rashid and Riena would both like to provide three input pairs that are collinear as points in the xy-plane.
(D) Rashid would like to provide three input pairs that are collinear, while Riena would like to provide three input pairs that are non-collinear as points in the xy-plane.
(E) Both Rashid and Riena are indifferent regarding how the three input pairs are picked.

Your answer: ____________________________

(2) (*) Let m and n be natural numbers with $m \geq 3$. We are given a bunch of numbers $x_1 < x_2 < \dots < x_m$ and another bunch of numbers $y_1, y_2, \dots, y_m$. We want to find a continuous function f on $[x_1, x_m]$, such that $f(x_i) = y_i$ for all $1 \leq i \leq m$, and such that the restriction of f to any interval of the form $[x_i, x_{i+1}]$ (for $1 \leq i \leq m - 1$) is a polynomial of degree $\leq n$. What is the smallest value of n for which we are guaranteed to be able to find such a function f?
(A) 1
(B) 2
(C) 3
(D) 4
(E) 5

Your answer: ____________________________

(3) (*) Let m and n be natural numbers with $m \geq 3$. We are given a bunch of numbers $x_1 < x_2 < \dots < x_m$ and another bunch of numbers $y_1, y_2, \dots, y_m$. We want to find a continuous function f on $[x_1, x_m]$, such that $f(x_i) = y_i$ for all $1 \leq i \leq m$, and such that the restriction of f to any interval of

the form $[x_i, x_{i+1}]$ (for $1 \leq i \leq m-1$) is a polynomial of degree $\leq n$. In addition, we want to make sure that f is differentiable on the open interval (x_1, x_m). What is the smallest value of n for which we are guaranteed to be able to find such a function f?
(A) 1
(B) 2
(C) 3
(D) 4
(E) 5

Your answer: ____________________________

(4) (*) Let m and n be natural numbers with $m \geq 3$. We are given a bunch of numbers $x_1 < x_2 < \cdots < x_m$ and another bunch of numbers $y_1, y_2, \ldots, y_m$. We want to find a continuous function f on $[x_1, x_m]$, such that $f(x_i) = y_i$ for all $1 \leq i \leq m$, and such that the restriction of f to any interval of the form $[x_i, x_{i+1}]$ (for $1 \leq i \leq m-1$) is a polynomial of degree $\leq n$. In addition, we want to make sure that f is differentiable on the open interval (x_1, x_m). In addition, we are told the value of the right hand derivative of f at x_1 and the left hand derivative of f at x_m. What is the smallest value of n for which we are guaranteed to be able to find such a function f?
(A) 1
(B) 2
(C) 3
(D) 4
(E) 5

Your answer: ____________________________

(5) (*) Let k, m, and n be natural numbers with $m \geq 3$. We are given a bunch of numbers $x_1 < x_2 < \cdots < x_m$ and another bunch of numbers $y_1, y_2, \ldots, y_m$. We want to find a continuous function f on $[x_1, x_m]$, such that $f(x_i) = y_i$ for all $1 \leq i \leq m$, and such that the restriction of f to any interval of the form $[x_i, x_{i+1}]$ (for $1 \leq i \leq m-1$) is a polynomial of degree $\leq n$. In addition, we want to make sure that f is at least k times differentiable on the open interval (x_1, x_m). What is the smallest value of n for which we are guaranteed to be able to find such a function f?
(A) $k-2$
(B) $k-1$
(C) k
(D) $k+1$
(E) $k+2$

Your answer: ____________________________

The next few questions are framed deterministically, though similar real-world applications would be probabilistic, with some square roots floating around. Unfortunately, we do not have the tools yet to deal with the probabilistic versions of the questions.

(6) (*) A function f of one variable is known to be linear. We know that $f(1) = 1.5 \pm 0.5$ and $f(2) = 2.5 \pm 0.5$. Assume these are the full ranges, without any probability distribution known. Assuming nothing is known about how the measurement errors for f at different points are related, what can we say about $f(3)$?
(A) $f(3) = 3.5$ (exactly)
(B) $f(3) = 3.5 \pm 0.5$
(C) $f(3) = 3.5 \pm 1$
(D) $f(3) = 3.5 \pm 1.5$
(E) $f(3) = 3.5 \pm 2.5$

Your answer: ____________________________

(7) (*) A function f of one variable is known to be linear. We know that $f(1) = 1.5 \pm 0.5$ and $f(2) = 2.5 \pm 0.5$. Assume these are the full ranges, without any probability distribution known. Assume also that the measurement error for f at all points is the same in magnitude and sign. What can we say about $f(3)$?
(A) $f(3) = 3.5$ (exactly)
(B) $f(3) = 3.5 \pm 0.5$
(C) $f(3) = 3.5 \pm 1$
(D) $f(3) = 3.5 \pm 1.5$
(E) $f(3) = 3.5 \pm 2.5$

Your answer: ____________________________

(8) (*) Suppose f is a linear function on a bounded interval $[a, b]$ but our measurement of outputs for given inputs has some measurement error (with the range of measurement error the same regardless of the input, and no known correlation between the magnitude of measurement error at different points). Assume we can get the outputs for any two specified inputs we desire, and we will then fit a line through the (input,output) pairs to get the graph of f. How should we choose our inputs?
(A) Choose the inputs as far as possible from each other, i.e., choose them as a and b.
(B) Choose the inputs to be as close to each other as possible, i.e., choose them to be nearby points but not equal to each other.
(C) It does not matter. Any choice of two distinct inputs is good enough.

Your answer: ____________________________

(9) (*) f is a function of one variable defined on an interval $[a, b]$. You are trying to find an explicit function that fits f well. You initially try a straight line fit that works at the points a and b. It turns out that this fit systematically overestimates f for points in between (i.e., the actual function f is below the linear function) with the maximum magnitude of discrepancy occurring at the midpoint $(a + b)/2$. Based on this information, what kind of fit should you try to look for?
(A) Try to fit f using a logarithmic function
(B) Try to fit f using an exponential function
(C) Try to fit f using a quadratic function
(D) Try to fit f using a polynomial of degree at most 3
(E) Try to fit f using the reciprocal of a linear function

Your answer: ____________________________

(10) (*) Recall the Leontief input-output model. Recall that the GDP is defined as the total money value of all the *final* goods and services produced in the economy, which in this case means only those that go into meeting consumer demand, not interindustry demand (note that we are assuming away the existence of investment and government spending, which complicate the GDP calculation). Assuming that the unit prices of the goods are constant (a very unrealistic assumption given that price itself responds to supply and demand, but fortunately it does not affect the conclusion we draw here) what might be a way of increasing GDP while keeping the magnitude of output of each industry the same?
(A) Increase interindustry dependence, i.e., increase the amount needed from each industry that is necessary to produce a given amount in another industry.
(B) Reduce interindustry dependence, i.e., reduce the amount needed from each industry that is necessary to produce a given amount in another industry.
(C) Changes in interindustry dependence have no effect.

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE MONDAY OCTOBER 14: MATRIX COMPUTATIONS

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ______________________________________

PLEASE DO *NOT* DISCUSS ANY QUESTIONS EXCEPT THE STARRED OR DOUBLE-STARRED QUESTIONS.

This quiz has a few questions on the mechanics of the computational execution of Gauss-Jordan elimination, and it has one question on setting up a linear system.

Suppose f is a function on the positive integers that takes positive integer values. Suppose n is a parameter related to the input size of an algorithm. We say that the running time of an algorithm (respectively, the space requirement of the algorithm) is:

- $O(f(n))$ if, for large enough n, it can be bounded from above by a positive constant times $f(n)$.
- $\Omega(f(n))$ if, for large enough n, it can be bounded from below by a positive constant times $f(n)$.
- $\Theta(f(n))$ if it is both $O(f(n))$ and $\Omega(f(n))$.

You can read more at:
`http://en.wikipedia.org/wiki/Big_O_notation`

(1) (*) If you treat each arithmetic operation (addition, subtraction, multiplication, division) of numbers as taking constant time, and all entry rewrites and changes as again taking constant time per entry, what would be the best description of the worst-case running time of the algorithm to convert a $n \times n$ matrix to reduced row-echelon form? (Note that this complexity is termed *arithmetic complexity* and can be distinguished from the *bit complexity* of the algorithm, which could be considerably higher).
(A) $\Theta(n)$
(B) $\Theta(n^2)$
(C) $\Theta(n^3)$
(D) $\Theta(n^4)$
(E) $\Theta(n^5)$

Your answer: ____________________________

(2) (*) If you treat each arithmetic operation (addition, subtraction, multiplication, division) of numbers as taking constant space, and all matrix entries as taking constant space, what would be the best description of the worst-case space requirement of the algorithm to convert a $n \times n$ matrix to reduced row-echelon form? Assume that space is reusable, i.e., it is possible to rewrite over existing space used.
(A) $\Theta(n)$
(B) $\Theta(n^2)$
(C) $\Theta(n^3)$
(D) $\Theta(n^4)$
(E) $\Theta(n^5)$

Your answer: ____________________________

(3) (*) Suppose the coefficient matrix of a linear system with n variables and n equations is known in advance, and we can spend as much time processing it as we desire in advance (this time will not count towards the running time of the algorithm). In other words, we can use Gauss-Jordan elimination to row-reduce the coefficient matrix in advance. However, we do not have the output column with us in advance. What is the worst-case running time of the part of the algorithm that runs after the output column is known?

(A) $\Theta(n)$
(B) $\Theta(n^2)$
(C) $\Theta(n^3)$
(D) $\Theta(n^4)$
(E) $\Theta(n^5)$

Your answer: ____________________________

(4) *Do not discuss this!*: Which of the following matrices does *not* have the identity matrix as its reduced row-echelon form?

(A)

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(B)

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 5 \\ 0 & 0 & 7 \end{bmatrix}$$

(C)

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

(D)

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

(E)

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 7 & 11 \end{bmatrix}$$

Your answer: ____________________________

(5) *Do not discuss this!*: A number of different consumer price indices have been constructed. All of them use the market prices for an existing collection of commodities (though not all of them use every commodity in the collection) and take a different "weighted" linear combination of those. For instance, one price index might be 3 times (the price per ton of wheat on the Chicago wheat market) + 4 times (the price of 1 gallon of unleaded gasoline at a particular gas station) + 17 times (the price of Burt's chapstick). Another price index might use 30 times (the price of Transcend's 32 GB flash drive) + 14 times (the price of 1 gallon of gasoline at a particular gas station).

What is a good way of modeling these?

(A) The prices of the various goods are stored in a matrix, the different weightings used in various indices are stored in a vector, and the consumer price indices arise as the output vector of the matrix-vector product.
(B) The different weightings used in various indices are stored in a matrix, the prices of the various goods are stored in a vector, and the consumer price indices arise as the output vector of the matrix-vector product.
(C) The prices of the various goods are stored in a matrix, the consumer price indices are stored as a vector, and the weightings used in the indices arise as the output vector of the matrix-vector product.
(D) The different weightings used in various indices are stored in a matrix, the consumer price indices are stored in a vector, and the prices of the various goods arise as the output vector of the matrix-vector product.

(E) The consumer price indices are stored in a matrix, the prices of the various goods are stored in a vector, and the weightings used in the indices arise as the output vector of the matrix-vector product.

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE FRIDAY OCTOBER 18: LINEAR SYSTEMS: RANK AND DIMENSION CONSIDERATIONS

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ______________________________

PLEASE DO *NOT* DISCUSS ANY QUESTIONS EXCEPT THE STARRED OR DOUBLE-STARRED QUESTIONS.

The questions here consider a wide range of theoretical and practical settings where linear systems appear, and prompt you to think about the notion of rank and its relationship with whether we can uniquely acquire the information that we want. It relates approximately with the material in the `Linear systems and matrix algebra` notes (the corresponding section in the book is Section 1.3).

(1) (*) Let m and n be positive integers. It turns out that *almost all* $m \times n$ matrices over the real numbers have a particular rank. What is that rank? (Unfortunately, it is beyond our current scope to define "almost all").
(A) m (regardless of whether m or n is bigger)
(B) n (regardless of whether m or n is bigger)
(C) $(m+n)/2$
(D) $\max\{m,n\}$
(E) $\min\{m,n\}$

Your answer: ______________________

(2) (*) A container has a mix of two known gases that do not react with each other. The temperature and pressure of the container are known. Assume that $PV = nRT$. The volume of the container is also known, and so is the total mass of the gases in the container. Under what conditions can we predict the amount (say, in the form of the number of moles) of each gas that is present from this information?
(A) It is possible if both gases have the same molecular mass, because in that case, the coefficient matrix of the linear system has full rank 2.
(B) It is possible if both gases have different molecular masses, because in that case, the coefficient matrix of the linear system has full rank 2.
(C) It is possible if both gases have the same molecular mass, because in that case, the coefficient matrix of the linear system has rank 1.
(D) It is possible if both gases have different molecular masses, because in that case, the coefficient matrix of the linear system has rank 1.
(E) It is not possible to deduce the amount of each gas from the given information.

Your answer: ______________________

(3) (*) A container has a mix of three known gases with no reactions between the gases. The temperature and pressure of the container are known. Assume that $PV = nRT$. The volume of the container is also known, and so is the total mass of the gases in the container. Under what conditions can we predict the amount (say, in the form of the number of moles) of each gas that is present from this information?
(A) It is possible if all three gases have the same molecular mass, because in that case, the coefficient matrix of the linear system has full rank 3.
(B) It is possible if all three gases have different molecular masses, because in that case, the coefficient matrix of the linear system has full rank 3.

(C) It is possible if all three gases have the same molecular mass, because in that case, the coefficient matrix of the linear system has rank 2.
(D) It is possible if all three gases have different molecular masses, because in that case, the coefficient matrix of the linear system has rank 2.
(E) It is not possible to deduce the amount of each gas from the given information.

Your answer: ____________________________

The branch of chemistry called quantitative analysis has historically used stoichiometric methods to determine the proportions of various chemicals present in a given mix. The idea is to use information about the amounts needed and produced in various reactions to estimate the quantities of chemicals present (the possible chemicals are first identified via "qualitative analysis" techniques). We generally find that these conditions give linear systems, and the coefficient matrices of these systems have (or can be written in a manner as to have) small integer entries.

(4) (*) Consider a situation where we have a material that is a mix (in fixed proportion) of three known chemicals X, Y, and Z. Our goal is to find the amount of X, Y, and Z present. Suppose we want to set up a collection of experiments so that the coefficient matrix is diagonal, i.e., we are effectively solving a diagonal system of equations and can recover the quantities of each of X, Y, and Z. Which of the following is the best approach? Assume that we can measure, for each reagent, the amount of the reagent that gets used up for the reaction(s) to proceed to completion, but cannot isolate or separate the outputs from each other.
(A) Choose a single reagent that reacts with all of X, Y, and Z.
(B) Choose a single reagent that reacts with only one of X, Y, and Z.
(C) Choose three separate reagents, each of which reacts with *all* of X, Y, and Z.
(D) Choose three separate reagents, each of which reacts only with X.
(E) Choose three separate reagents, one of which reacts only with X, one of which reacts only with Y, and one of which reacts only with Z.

Your answer: ____________________________

(5) (*) Suppose we are given an aqueous solution with two known dissolved substances. There are two different types of reactions. One is an acid-base reaction and the other is a redox reaction. For both reactions, we can use titrations (separately) to deduce the quantity of reagent needed. What type of system should we expect to get if only one of the solutes participates in the redox reaction but both participate in the acid-base reaction?
(A) A diagonal system, i.e., the coefficient matrix is a diagonal matrix.
(B) A triangular system, i.e., the coefficient matrix is a triangular matrix (whether it is upper or lower triangular depends on the order in which we write the rows).
(C) A system of rank one, i.e., the coefficient matrix has rank one.

Your answer: ____________________________

(6) (*) Suppose we are given an aqueous solution with two known dissolved substances. There are two different types of reactions. One is an acid-base reaction and the other is a redox reaction. For both reactions, we can use titrations (separately) to deduce the quantity of reagent needed.Suppose we are given an aqueous solution with two known dissolved substances. Suppose both solutes participate in both reactions. What should we desire if we want to use the data from the two titrations to determine the amounts of each of the substances?
(A) The proportions in which the two substances react should be the same for the two reactions.
(B) The proportions in which the two substances react should differ for the two reactions.
(C) It does not matter; we will be able to determine the amounts of each of the substances in both cases.
(D) It does not matter; we will not be able to determine the amounts of each of the substances in either case.

Your answer: ____________________________

(7) *Do not discuss this!*: A consumer price index is obtained from a "goods basket" by multiplying the price of each good in the basket by a fixed weight, and then adding up all the price X weight products. The weights are kept fixed, but the prices vary from year to year. Thus, the consumer price index value itself fluctuates from year to year.

What is a good way of modeling this?

(A) The prices of the various goods in various years are stored in a matrix, the weights used in the index are stored in a vector, and the consumer price index values arise as the output vector of the matrix-vector product.

(B) The weights used in the index are stored in a matrix, the prices of the various goods in various years are stored in a vector, and the consumer price index values arise as the output vector of the matrix-vector product.

(C) The prices of the various goods in various years are stored in a matrix, the consumer price index values are stored as a vector, and the weights used in the index arise as the output vector of the matrix-vector product.

(D) The weights used in the index are stored in a matrix, the consumer price index values are stored in a vector, and the prices of the various goods in various years arise as the output vector of the matrix-vector product.

(E) The consumer price index values are stored in a matrix, the prices of the various goods in various years are stored in a vector, and the weights used in the index arise as the output vector of the matrix-vector product.

Your answer: ____________________________

(8) *Do not discuss this!*: Amelia wants to choose a healthy balanced diet. She has access to 30 different types of foods. There are 400 different nutrients that she wants a good amount of. Each of the foods that Amelia consumes offers a positive amount of each nutrient per unit foodstuff. Amelia is interested in meeting the daily value requirements for all nutrients. For some nutrients, her daily value requirements specify only a minimum. For some nutrients, both a minimum and a maximum are specified. Assume that the total amount of any nutrient can be obtained by adding up the amounts obtained from each of the foodstuffs Amelia consumes. Amelia wants to determine how much of each foodstuff she should consume. How should she model the situation?

(A) The matrix with information on the nutritional contents of the foodstuffs is a 400×400 matrix, and the vector of amounts of each foodstuff consumed is a 400×1 column vector.

(B) The matrix with information on the nutritional contents of the foodstuffs is a 30×30 matrix, and the vector of amounts of each foodstuff consumed is a 30×1 column vector.

(C) The matrix with information on the nutritional contents of the foodstuffs is a 400×30 matrix, and the vector of amounts of each foodstuff consumed is a 30×1 column vector.

(D) The matrix with information on the nutritional contents of the foodstuffs is a 30×400 matrix, and the vector of amounts of each foodstuff consumed is a 400×1 column vector.

(E) The matrix with information on the nutritional contents of the foodstuffs is a 400×400 matrix, and the vector of amounts of each foodstuff consumed is a 30×1 column vector.

Your answer: ____________________________

DIAGNOSTIC IN-CLASS QUIZ: DUE FRIDAY OCTOBER 18: LINEAR TRANSFORMATIONS

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____________________

PLEASE DO *NOT* DISCUSS ANY QUESTIONS.

The quiz covers basics related to linear transformations (notes titled `Linear transformations`, corresponding section in the book Section 2.1). Explicitly, the quiz covers:

- Representation of a linear transformation using a matrix, and identifying the domain and co-domain in terms of the row and column counts of the matrix.
- Relationship between injectivity, surjectivity, rank, row count, and column count.
- Relationship between the entries of the matrix and the images of the standard basis vectors under the corresponding linear transformation.

The question are fairly easy if you understand the material. But it's important that you be able to answer them, otherwise what we study later will not make much sense.

(1) *Do not discuss this!*: Which of the following correctly describes a $m \times n$ matrix?
  (A) There are m rows, and each row gives a vector with m coordinates. There are n columns, and each column gives a vector with n coordinates.
  (B) There are m rows, and each row gives a vector with n coordinates. There are n columns, and each column gives a vector with m coordinates.
  (C) There are n rows, and each row gives a vector with m coordinates. There are m columns, and each column gives a vector with n coordinates.
  (D) There are n rows, and each row gives a vector with n coordinates. There are m columns, and each column gives a vector with m coordinates.

  Your answer: ____________________

(2) *Do not discuss this!*: For a $p \times q$ matrix A, we can define a linear transformation T_A by $T_A(\vec{x}) := A\vec{x}$. What type of linear transformation is T_A?
  (A) T_A is a linear transformation from $\mathbb{R}^p$ to $\mathbb{R}^q$
  (B) T_A is a linear trnasformation from $\mathbb{R}^q$ to $\mathbb{R}^p$
  (C) T_A is a linear transformation from $\mathbb{R}^{\max\{p,q\}}$ to $\mathbb{R}^{\min\{p,q\}}$
  (D) T_A is a linear transformation from $\mathbb{R}^{\min\{p,q\}}$ to $\mathbb{R}^{\max\{p,q\}}$

  Your answer: ____________________

(3) *Do not discuss this!*: With the same notation as for the preceding question, which of the following is true?
  (A) If $p < q$, T_A must be injective
  (B) If $p > q$, T_A must be injective
  (C) If $p = q$, T_A must be injective
  (D) If $p < q$, T_A cannot be injective
  (E) If $p > q$, T_A cannot be injective

  Your answer: ____________________

(4) *Do not discuss this!*: With the same notation as for the previous two questions, which of the following is true?
  (A) If $p < q$, T_A must be surjective

(B) If $p > q$, T_A must be surjective
(C) If $p = q$, T_A must be surjective
(D) If $p < q$, T_A cannot be surjective
(E) If $p > q$, T_A cannot be surjective

Your answer: ____________________________

(5) *Do not discuss this!*: With the same notation as for the last three questions, which of the following is true?
(A) The rows of A are the images under T_A of the standard basis vectors of $\mathbb{R}^p$.
(B) The columns of A are the images under T_A of the standard basis vectors of $\mathbb{R}^p$.
(C) The rows of A are the images under T_A of the standard basis vectors of $\mathbb{R}^q$.
(D) The columns of A are the images under T_A of the standard basis vectors of $\mathbb{R}^q$.

Your answer: ____________________________

DIAGNOSTIC IN-CLASS QUIZ: DUE FRIDAY OCTOBER 25: MATRIX MULTIPLICATION (BASIC)

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____________________

PLEASE DO *NOT* DISCUSS ANY QUESTIONS

This quiz tests for basic comprehension of the setup for matrix multiplication. It corresponds to the material from Sections 1-6 (excluding Section 4) of the `Matrix multiplication and inversion` notes, and also to Section 2.3 of the book.

(1) *Do not discuss this!*: Suppose A and B are (not necessarily square) matrices. Then, which of the following describes correctly the relationship between the existence and value of the (alleged) matrix product AB and the existence and value of the (alleged) matrix product BA?
   (A) AB is defined if and only if BA is defined, and if so, they are equal.
   (B) AB is defined if and only if BA is defined, but they need not be equal.
   (C) If AB and BA are both defined, then $AB = BA$. However, it is possible for one of AB and BA to be defined and the other to not be defined.
   (D) It is possible for only one of AB and BA to be defined. It is also possible for both AB and BA to be defined, but to not be equal to each other.

   Your answer: ____________________

(2) *Do not discuss this!*: Suppose A and B are matrices such that both AB and BA are defined. Which of the following correctly describes what we know about AB and BA?
   (A) Both AB and BA are square matrices and have the same dimensions, i.e., in both AB and BA, the number of rows equals the number of columns, and further, the number of rows of AB equals the number of rows of BA.
   (B) Both AB and BA are square matrices (the number of rows equals the number of columns) but they may not have the same dimensions: the number of rows in AB need not equal the number of rows in BA.
   (C) AB and BA need not be square matrices but both must have the same dimensions: the number of rows in AB equals the number of rows in BA, and the number of columns in AB equals the number of columns in BA.
   (D) AB and BA need not be square matrices and they need not have the same row count or the same column count, i.e., the number of rows in AB need not equal the number of rows in BA, and the number of columns in AB need not equal the number of columns in BA.

   Your answer: ____________________

(3) *Do not discuss this!*: Suppose A and B are matrices such that both AB and $A + B$ are defined. Which of the following correctly describes what we know about A and B?
   (A) Both A and B are square matrices and have the same dimensions, i.e., in both A and B, the number of rows equals the number of columns, and further, the number of rows of A equals the number of rows of B.
   (B) Both A and B are square matrices (the number of rows equals the number of columns) but they may not have the same dimensions: the number of rows in A need not equal the number of rows in B.
   (C) A and B need not be square matrices but both must have the same dimensions: the number of rows in A equals the number of rows in B, and the number of columns in A equals the number of columns in B.

(D) A and B need not be square matrices and they need not have the same row count or the same column count, i.e., the number of rows in A need not equal the number of rows in B, and the number of columns in A need not equal the number of columns in B.

Your answer: ____________________________

(4) *Do not discuss this!*: Suppose A is a $p \times q$ matrix and B is a $q \times r$ matrix. The product matrix AB is a $p \times r$ matrix. Using the convention of matrices as linear transformations via their action by multiplication on column vectors, what is the appropriate interpretation of the matrix product in terms of composing linear transformations?

(A) A corresponds to a linear transformation T_A from $\mathbb{R}^p$ to $\mathbb{R}^q$, and B corresponds to a linear transformation T_B from $\mathbb{R}^q$ to $\mathbb{R}^r$. The product AB therefore corresponds to a linear transformation from $\mathbb{R}^p$ to $\mathbb{R}^r$ that is the composite of the two linear transformations, with T_A applied first (to the domain) and then T_B (T_B being applied to the intermediate space obtained after applying T_A).

(B) A corresponds to a linear transformation T_A from $\mathbb{R}^p$ to $\mathbb{R}^q$, and B corresponds to a linear transformation T_B from $\mathbb{R}^q$ to $\mathbb{R}^r$. The product AB therefore corresponds to a linear transformation from $\mathbb{R}^p$ to $\mathbb{R}^r$ that is the composite of the two linear transformations, with T_B applied first (to the domain) and then T_A (T_A being applied to the intermediate space obtained after applying T_B).

(C) A corresponds to a linear transformation T_A from $\mathbb{R}^q$ to $\mathbb{R}^p$, and B corresponds to a linear transformation T_B from $\mathbb{R}^r$ to $\mathbb{R}^q$. The product AB therefore corresponds to a linear transformation from $\mathbb{R}^r$ to $\mathbb{R}^p$ that is the composite of the two linear transformations, with T_A applied first (to the domain) and then T_B (T_B being applied to the intermediate space obtained after applying T_A).

(D) A corresponds to a linear transformation T_A from $\mathbb{R}^q$ to $\mathbb{R}^p$, and B corresponds to a linear transformation T_B from $\mathbb{R}^r$ to $\mathbb{R}^q$. The product AB therefore corresponds to a linear transformation from $\mathbb{R}^r$ to $\mathbb{R}^p$ that is the composite of the two linear transformations, with T_B applied first (to the domain) and then T_A (T_A being applied to the intermediate space obtained after applying T_B).

Your answer: ____________________________

(5) *Do not discuss this!*: Suppose A, B, and C are matrices. Which of the following is true?

(A) If ABC is defined, then so are BCA and CAB.

(B) If ABC and BCA are both defined, then so is CAB. However, it is possible to have a situation where ABC is defined but BCA and CAB are not defined.

(C) It is possible to have a situation where ABC and BCA are both defined but CAB is not defined.

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE MONDAY OCTOBER 28: MATRIX MULTIPLICATION AND INVERSION AS COMPUTATIONAL PROBLEMS

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____

This quiz tests for a strong *conceptualization* (i.e., a metacognition) of the processes used for matrix multiplication and inversion. It is based on part of the `Matrix multiplication and inversion` notes and is related to Sections 2.3 and 2.4. It does not, however, test all aspects of that material.

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

(1) How many arithmetic operations are needed for naive matrix multiplication of a $m \times n$ matrix and a $n \times p$ matrix?
(A) $O(mnp)$ additions and $O(mnp)$ multiplications
(B) $O(m + n + p)$ additions and $O(mnp)$ multiplications
(C) $O(mn)$ additions and $O(np)$ multiplications
(D) $O(mn + mp)$ additions and $O(mnp)$ multiplications
(E) $O(m + n + p)$ additions and $O(m + n + p)$ multiplications

Your answer: ____________________________

(2) What is the arithmetic complexity (in terms of total number of arithmetic operations needed) for naive matrix multiplication of two generic $n \times n$ matrices?
(A) $\Theta(n)$
(B) $\Theta(n^2)$
(C) $\Theta(n^3)$
(D) $\Theta(n^4)$
(E) $\Theta(n^5)$

Your answer: ____________________________

(3) Which of the following is the tightest "obvious" lower bound on the possible arithmetic complexity of any generic algorithm for multiplying two $n \times n$ matrices? We use Ω to denote *at least that order*.
(A) $\Omega(n)$
(B) $\Omega(n^2)$
(C) $\Omega(n^3)$
(D) $\Omega(n^4)$
(E) $\Omega(n^5)$

Your answer: ____________________________

(4) What is the minimum number of arithmetic operations needed to compute the product of two generic diagonal $n \times n$ matrices?
(A) n
(B) $n + 1$
(C) $2n - 1$
(D) $2n$
(E) n^2

Your answer: ____________________________

(5) What is the minimum number of arithmetic operations needed to compute the product of a generic $1 \times n$ matrix and a generic $n \times 1$ matrix?

(A) n
(B) $n+1$
(C) $2n-1$
(D) $2n$
(E) n^2

Your answer: ____________________________

(6) What is the minimum number of arithmetic operations needed to compute the product of a generic $n \times 1$ matrix and a generic $1 \times n$ matrix?
(A) n
(B) $n+1$
(C) $2n-1$
(D) $2n$
(E) n^2

Your answer: ____________________________

(7) In order terms, what is the minimum number of arithmetic operations needed to compute the product of a generic $n \times n$ diagonal matrix and a generic $n \times n$ upper triangular matrix? The upper triangular matrix has zero entries below the diagonal. The entries on or above the diagonal may be nonzero (and generically, they will be nonzero).
(A) $n(n-1)/2$
(B) $n(n+1)/2$
(C) $n(n-1)$
(D) n^2
(E) $n(n+1)$

Your answer: ____________________________

Adding n numbers to each other requires $n-1$ addition operations. In a non-parallel setting, there is no way of improving this.

However, using the associativity of addition, we can write a faster parallelizable algorithm. A simple parallelization is to split the list being added into two sublists of length about $n/2$ each. Delegate the task of adding up within each sublist to different processors running in parallel. Then, add up the numbers obtained. This takes about half the time, with a little overhead (of collecting and adding up). This type of strategy is called a *divide and conquer* strategy. Using a divide and conquer strategy repeatedly, we can demonstrate that the parallelized arithmetic complexity of this approach is $\Theta(\log_2 n)$.

(8) Suppose A is a $1 \times n$ matrix and B is a $n \times 1$ matrix. Assume an unlimited number of processors that all have free read access to both A and B, free write access to the product matrix, and a shared workspace where they can store intermediate results. What is the arithmetic complexity in this context (i.e., the parallelized arithmetic complexity) for computing AB? What we mean here is: what is the smallest depth of a computational tree to compute AB?
(A) $\Theta(1)$
(B) $\Theta(\log_2 n)$
(C) $\Theta(n \log_2 n)$
(D) $\Theta(n^2)$
(E) $\Theta(n^2 \log_2 n)$

Your answer: ____________________________

(9) Suppose A is a $n \times 1$ matrix and B is a $1 \times n$ matrix. Assume an unlimited number of processors that all have free read access to both A and B, free write access to the product matrix, and a shared workspace where they can store intermediate results. What is the arithmetic complexity in this

context (i.e., the parallelized arithmetic complexity) for computing AB? What we mean here is: what is the smallest depth of a computational tree to compute AB?
(A) $\Theta(1)$
(B) $\Theta(\log_2 n)$
(C) $\Theta(n \log_2 n)$
(D) $\Theta(n^2)$
(E) $\Theta(n^2 \log_2 n)$

Your answer: ____________________________

(10) Suppose A and B are two $n \times n$ matrices. Assume an unlimited number of processors that all have free read access to both A and B, free write access to the product matrix, and a shared workspace where they can store intermediate results. What is the arithmetic complexity in this context (i.e., the parallelized arithmetic complexity) for computing AB? What we mean here is: what is the smallest depth of a computational tree to compute AB? Use naive matrix multiplication and speed it up using the parallelized processes discused here.
(A) $\Theta(1)$
(B) $\Theta(\log_2 n)$
(C) $\Theta(n \log_2 n)$
(D) $\Theta(n^2)$
(E) $\Theta(n^2 \log_2 n)$

Your answer: ____________________________

We are given a $n \times n$ matrix A and we want to use *repeated squaring* to calculate powers of A. For instance, to calculate A^4, we can simply calculate $(A^2)^2$, which requires two multiplications. To calculate A^5, we calculate $(A^2)^2 A$, which requires three multiplications. Assume that we can store any number of intermediate matrices, i.e., storage space is not a constraint.

(11) What is the smallest number of matrix multiplications needed to calculate A^7 using repeated squaring?
(A) 3
(B) 4
(C) 5
(D) 6
(E) 7

Your answer: ____________________________

(12) What is the smallest number of matrix multiplications needed to calculate A^8 using repeated squaring?
(A) 3
(B) 4
(C) 5
(D) 6
(E) 7

Your answer: ____________________________

(13) What is the smallest number of matrix multiplications needed to calculate A^{21} using repeated squaring?
(A) 3
(B) 4
(C) 5
(D) 6
(E) 7

Your answer: ____________________________

Suppose A is an *invertible* $n \times n$ matrix. It is possible to invert A using $\Theta(n^3)$ (worst-case) arithmetic operations via Gauss-Jordan elimination. We can thus add computation of the inverse to our toolkit when calculating powers. It is helpful even when calculating positive powers.

Count each matrix multiplication and each matrix inversion as one "matrix operation."

(14) What is the smallest positive r where we can achieve a saving on the total number of matrix operations to calculate A^r by also computing A^{-1}, rather than just using repeated squaring?

(A) 3
(B) 7
(C) 15
(D) 23
(E) 31

Your answer: ____________________________

(15) *Strassen's algorithm* is a *fast matrix multiplication* algorithm that can multiply two $n \times n$ matrices using $O(n^{\log_2 7})$ arithmetic operations. In practice, however, a lot of existing computer code for matrix multiplication, written long after Strassen's algorithm was discovered, uses naive matrix multiplication. Which of the following reasons explain this? Please see Options (D) and (E) before answering.

(A) Strassen's algorithm becomes faster than naive matrix multiplication only for very large matrix sizes.
(B) Strassen's algorithm is more complicated to code.
(C) Strassen's algorithm is not as easily parallelizable as naive matrix multiplication.
(D) All of the above.
(E) None of the above.

Your answer: ____________________________

There exist even faster algorithms for matrix multiplication than Strassen's algorithm. The best known algorithm currently is the *Coppersmith-Winograd algorithm*, which can multiply two $n \times n$ matrices in time $O(n^{2.3727})$. However, the Coppersmith-Winograd algorithm is even more rarely implemented than Strassen's for practical matrix multiplication code (according to some sources, Coppersmith-Winograd has *never* been implemented). The same reasons as those cited above for the reluctance to use Strassen's algorithm apply. There are some additional obstacles to practical implementations of the Coppersmith-Winograd algorithm that make it even more difficult to use.

(16) Suppose A and B are $n \times n$ matrices. What is the minimum number of matrix multiplications needed generically to compute the product $ABABABABA$?

(A) 4
(B) 5
(C) 6
(D) 7
(E) 8

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE WEDNESDAY OCTOBER 30: LINEAR TRANSFORMATIONS AND FINITE STATE AUTOMATA

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

The purpose of this quiz is to explore in greater depth particular types of matrices, the corresponding linear transformations, and the relationship between operations on sets and similar operations on vector spaces. The material covered in the quiz will also prove to be a fertile source of *examples* and *counterexamples* for later content: in the future, when you are asked to come up with matrices that satisfy some very loosely stated conditions, the matrices of the type described here can be a place to begin your search.

Let n be a natural number greater than 1. Suppose $f : \{0, 1, 2, \dots, n\} \to \{0, 1, 2, \dots, n\}$ is a function satisfying $f(0) = 0$. Let T_f denote the linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ satisfying the following for all $i \in \{1, 2, \dots, n\}$:

$$T_f(\vec{e}_i) = \begin{cases} \vec{e}_{f(i)}, & f(i) \neq 0 \\ \vec{0}, & f(i) = 0 \end{cases}$$

Let M_f denote the matrix for the linear transformation T_f. M_f can be described explicitly as follows: the i^{th} column of M_f is $\vec{0}$ if $f(i) = 0$ and is $\vec{e}_{f(i)}$ if $f(i) \neq 0$.

Note that if $f, g : \{0, 1, 2, \dots, n\} \to \{0, 1, 2, \dots, n\}$ are functions with $f(0) = g(0) = 0$, then $M_{f \circ g} = M_f M_g$ and $T_{f \circ g} = T_f \circ T_g$.

We will also use the following terminology:

- A $n \times n$ matrix A is termed *idempotent* if $A^2 = A$.
- A $n \times n$ matrix A is termed *nilpotent* if there exists a positive integer r such that $A^r = 0$.
- A $n \times n$ matrix A is termed a *permutation matrix* if every row contains one 1 and all other entries 0, *and* every column contains one 1 and all other entries 0.

(1) What condition on a function $f : \{0, 1, 2, \dots, n\} \to \{0, 1, 2, \dots, n\}$ (satisfying $f(0) = 0$) is equivalent to requiring M_f to be idempotent?
   (A) $(f(x))^2 = x$ for all $x \in \{0, 1, 2, \dots, n\}$
   (B) $f(x^2) = x$ for all $x \in \{0, 1, 2, \dots, n\}$
   (C) $(f(x))^2 = f(x)$ for all $x \in \{0, 1, 2, \dots, n\}$
   (D) $f(f(x)) = x$ for all $x \in \{0, 1, 2, \dots, n\}$
   (E) $f(f(x)) = f(x)$ for all $x \in \{0, 1, 2, \dots, n\}$

   Your answer: ____________________________

(2) What condition on a function $f : \{0, 1, 2, \dots, n\} \to \{0, 1, 2, \dots, n\}$ (satisfying $f(0) = 0$) is equivalent to requiring M_f to be nilpotent?
   (A) Composing f enough times with itself gives the identity function (i.e., the function that sends everything to itself).
   (B) Composing f enough times with itself gives the function that sends everything to 0.
   (C) Composing f enough times with itself gives the function that sends everything to 1.
   (D) Multiplying f enough times with itself gives the identity function (i.e., the function that sends everything to itself).
   (E) Multiplying f enough times with itself gives the function that sends everything to 0.

   Your answer: ____________________________

(3) What condition on a function $f : \{0, 1, 2, \ldots, n\} \to \{0, 1, 2, \ldots, n\}$ (satisfying $f(0) = 0$) is equivalent to requiring M_f to be a permutation matrix?

(A) Composing f enough times with itself gives the identity function (i.e., the function that sends everything to itself).
(B) Composing f enough times with itself gives the function that sends everything to 0.
(C) Composing f enough times with itself gives the function that sends everything to 1.
(D) Multiplying f enough times with itself gives the identity function (i.e., the function that sends everything to itself).
(E) Multiplying f enough times with itself gives the function that sends everything to 0.

Your answer: ____________________________

(4) Consider a function $f : \{0, 1, 2, \ldots, n\} \to \{0, 1, 2, \ldots, n\}$ with the property that $f(0) = 0$ and, for each $i \in \{1, 2, \ldots, n\}$, $f(i)$ is either i or 0. Note that the behavior may be different for different values of i (so some of them may go to themselves, and others may go to 0). What can we say M_f must be?

(A) M_f must be the identity matrix.
(B) M_f must be the zero matrix.
(C) M_f must be an idempotent matrix.
(D) M_f must be a nilpotent matrix.
(E) M_f must be a permutation matrix.

Your answer: ____________________________

(5) Which of the following pairs of candidates for $f, g : \{0, 1, 2\} \to \{0, 1, 2\}$ satisfies the condition that $M_f M_g = 0$ but $M_g M_f \neq 0$?

(A) $f(0) = 0$, $f(1) = 1$, $f(2) = 2$, whereas $g(0) = 0$, $g(1) = 2$, $g(2) = 1$
(B) $f(0) = 0$, $f(1) = 0$, $f(2) = 1$, whereas $g(0) = 0$, $g(1) = 2$, $g(2) = 0$
(C) $f(0) = 0$, $f(1) = 1$, $f(2) = 0$, whereas $g(0) = 0$, $g(1) = 0$, $g(2) = 2$
(D) $f(0) = 0$, $f(1) = 0$, $f(2) = 1$, whereas $g(0) = 0$, $g(1) = 1$, $g(2) = 0$
(E) $f(0) = 0$, $f(1) = 1$, $f(2) = 0$, whereas $g(0) = 0$, $g(1) = 0$, $g(2) = 1$

Your answer: ____________________________

(6) Which of the following pairs of candidates for $f, g : \{0, 1, 2\} \to \{0, 1, 2\}$ satisfies the condition that M_f and M_g are both nilpotent but $M_f M_g$ is not nilpotent?

(A) $f(0) = 0$, $f(1) = 1$, $f(2) = 2$, whereas $g(0) = 0$, $g(1) = 2$, $g(2) = 1$
(B) $f(0) = 0$, $f(1) = 0$, $f(2) = 1$, whereas $g(0) = 0$, $g(1) = 2$, $g(2) = 0$
(C) $f(0) = 0$, $f(1) = 1$, $f(2) = 0$, whereas $g(0) = 0$, $g(1) = 0$, $g(2) = 2$
(D) $f(0) = 0$, $f(1) = 0$, $f(2) = 1$, whereas $g(0) = 0$, $g(1) = 1$, $g(2) = 0$
(E) $f(0) = 0$, $f(1) = 1$, $f(2) = 0$, whereas $g(0) = 0$, $g(1) = 0$, $g(2) = 1$

Your answer: ____________________________

(7) Which of the following pairs of candidates for $f, g : \{0, 1, 2\} \to \{0, 1, 2\}$ satisfies the condition that neither M_f and M_g is nilpotent but $M_f M_g$ is nilpotent?

(A) $f(0) = 0$, $f(1) = 1$, $f(2) = 2$, whereas $g(0) = 0$, $g(1) = 2$, $g(2) = 1$
(B) $f(0) = 0$, $f(1) = 0$, $f(2) = 1$, whereas $g(0) = 0$, $g(1) = 2$, $g(2) = 0$
(C) $f(0) = 0$, $f(1) = 1$, $f(2) = 0$, whereas $g(0) = 0$, $g(1) = 0$, $g(2) = 2$
(D) $f(0) = 0$, $f(1) = 0$, $f(2) = 1$, whereas $g(0) = 0$, $g(1) = 1$, $g(2) = 0$
(E) $f(0) = 0$, $f(1) = 1$, $f(2) = 0$, whereas $g(0) = 0$, $g(1) = 0$, $g(2) = 1$

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE FRIDAY NOVEMBER 1: MATRIX MULTIPLICATION AND INVERSION: ABSTRACT BEHAVIOR PREDICTION

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

This quiz tests for *abstract behavior prediction* related to the structure of matrices defined based on the operations of matrix multiplication and inversion. It is based on part of the `Matrix multiplication and inversion` notes and is related to Sections 2.3 and 2.4. It does not, however, test all aspects of that material.

To understand this abstract behavior, we will consider *nilpotent*, *invertible*, and *idempotent* matrices.

(1) Suppose A and B are $n \times n$ matrices such that B is invertible. Suppose r is a positive integer. What can we say that $(BAB^{-1})^r$ definitely equals?
(A) A^r
(B) BA^rB^{-1}
(C) $B^rA^rB^{-r}$
(D) B^rAB^{-r}
(E) BAB^{-1-r}

Your answer: ____________________________

(2) Suppose A and B are $n \times n$ matrices (n not too small) such that $(AB)^2 = 0$. What is the smallest r for which we can conclude that $(BA)^r$ is definitely 0?
(A) 1
(B) 2
(C) 3
(D) 4
(E) 5

Your answer: ____________________________

(3) Suppose $n > 1$. A $n \times n$ matrix A is termed *nilpotent* if there exists a positive integer r such that A^r is the zero matrix. It turns out that if A is nilpotent, then $A^n = 0$. Which of the following describes correctly the relationship between being invertible and being nilpotent for $n \times n$ matrices?
(A) A matrix is nilpotent if and only if it is invertible.
(B) Every nilpotent matrix is invertible, but not every invertible matrix is nilpotent.
(C) Every invertible matrix is nilpotent, but not every nilpotent matrix is invertible.
(D) An invertible matrix may or may not be nilpotent, and a nilpotent matrix may or may not be invertible.
(E) A matrix cannot be both nilpotent and invertible.

Your answer: ____________________________

(4) Suppose A and B are $n \times n$ matrices. Which of the following is true? Please see Option (E) before answering.
(A) AB is nilpotent if and only if A and B are both nilpotent.
(B) AB is nilpotent if and only if at least one of A and B is nilpotent.
(C) If both A and B are nilpotent, then AB is nilpotent, but AB being nilpotent does not imply that both A and B are nilpotent.
(D) If AB is nilpotent, then both A and B are nilpotent. However, both A and B being nilpotent does not imply that AB is nilpotent.

(E) None of the above.

Your answer: ____________________________

(5) Suppose A and B are $n \times n$ matrices. Which of the following is true? Please see Option (E) before answering.

(A) AB is invertible if and only if A and B are both invertible.

(B) AB is invertible if and only if at least one of A and B is invertible.

(C) If both A and B are invertible, then AB is invertible, but AB being invertible does not imply that both A and B are invertible.

(D) If AB is invertible, then both A and B are invertible. However, both A and B being invertible does not imply that AB is invertible.

(E) None of the above.

Your answer: ____________________________

(6) Suppose A and B are $n \times n$ matrices. Which of the following is true? We call a $n \times n$ matrix *idempotent* if it equals its own square. Please see Option (E) before answering.

(A) AB is idempotent if and only if A and B are both idempotent.

(B) AB is idempotent if and only if at least one of A and B is idempotent.

(C) If both A and B are idempotent, then AB is idempotent, but AB being idempotent does not imply that both A and B are idempotent.

(D) If AB is idempotent, then both A and B are idempotent. However, both A and B being idempotent does not imply that AB is idempotent.

(E) None of the above.

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE WEDNESDAY NOVEMBER 6: GEOMETRY OF LINEAR TRANSFORMATIONS (ABSTRACT)

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ______________________________

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

This quiz tests for a deep abstract understanding of linear transformations and their geometry. It is related to Section 2.2 of the book and also to the `Geometry of linear transformations` lecture notes.

For the questions here, please use the following terminology.

Suppose n is a fixed natural number greater than 1. For ease of geometric visualization, you can take $n = 2$ for the discussion.

- A *linear automorphism* of $\mathbb{R}^n$ is defined as a bijective linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$.
- An *affine linear automorphism* of $\mathbb{R}^n$ is defined as a bijective function from $\mathbb{R}^n$ to itself that preserves collinearity, i.e., it sends lines to lines. In addition, it preserves the ratios of lengths within each line. This can be included as part of the definition or deduced from the fact that collinearity is preserved for $n > 1$.
- A *self-isometry* of $\mathbb{R}^n$ is defined as a bijective function from $\mathbb{R}^n$ to itself that preserves Euclidean distance: for all pairs of points $\vec{x}, \vec{y} \in \mathbb{R}^n$, the Euclidean distance between $\vec{x}$ and $\vec{y}$ equals the Euclidean distance between $T(\vec{x})$ and $T(\vec{y})$.
- A *self-homothety* (or *similitude transformation* or *similarity transformation*) of $\mathbb{R}^n$ is defined as a bijective function from $\mathbb{R}^n$ to itself that multiplies all distances by a fixed number called the *factor of similitude* (dependent on the transformation): if the factor of similitude is λ, then for all pairs of points $\vec{x}, \vec{y} \in \mathbb{R}^n$, the distance between $T(\vec{x})$ and $T(\vec{y})$ equals λ times the distance between $\vec{x}$ and $\vec{y}$.

(1) What is the relationship between linear automorphisms and affine linear automorphisms of $\mathbb{R}^n$?
 (A) Being a linear automorphism is precisely the same as being an affine linear automorphism.
 (B) Every linear automorphism is an affine linear automorphism, but not every affine linear automorphism is a linear automorphism.
 (C) Every affine linear automorphism is a linear automorphism, but not every linear automorphism is an affine linear automorphism.
 (D) A linear automorphism need not be affine linear, and an affine linear automorphism need not be linear.

 Your answer: ______________________

(2) What is the relationship between self-homotheties and self-isometries of $\mathbb{R}^n$?
 (A) Being a self-homothety is precisely the same as being a self-isometry.
 (B) Every self-homothety is a self-isometry, but not every self-isometry is a self-homothety.
 (C) Every self-isometry is a self-homothety, but not every self-homothety is a self-isometry.
 (D) A self-homothety need not be a self-isometry, and a self-isometry need not be a self-homothety.

 Your answer: ______________________

(3) What is the relationship between self-homotheties and affine linear automorphisms of $\mathbb{R}^n$?
 (A) Being an affine linear automorphism is precisely the same as being a self-homothety.
 (B) Every affine linear automorphism is a self-homothety, but not every self-homothety is an affine linear automorphism.
 (C) Every self-homothety is an affine linear automorphism, but not every affine linear automorphism is a self-homothety.

(D) An affine linear automorphism need not be a self-homothety, and a self-homothety need not be affine linear.

Your answer: ____________________________

(4) What is the relationship between affine linear automorphisms and self-isometries of $\mathbb{R}^n$?
(A) Being an affine linear automorphism is precisely the same as being a self-isometry.
(B) Every affine linear automorphism is a self-isometry, but not every self-isometry is an affine linear automorphism.
(C) Every self-isometry is an affine linear automorphism, but not every affine linear automorphism is a self-isometry.
(D) An affine linear automorphism need not be a self-isometry, and a self-isometry need not be affine linear.

Your answer: ____________________________

(5) There is a special kind of bijection from $\mathbb{R}^n$ to $\mathbb{R}^n$ called a *translation*. A translation with translation vector $\vec{v}$ is defined as the bijection $\vec{x} \mapsto \vec{x} + \vec{v}$. A *nontrivial* translation is a translation whose translation vector is not the zero vector. Which of the following is an automorphism type that nontrivial translations are *not*? Please see Option (E) before answering.
(A) Linear automorphism
(B) Affine linear automorphism
(C) Self-isometry
(D) Self-homothety
(E) None of the above, i.e., nontrivial translations are of all these types

Your answer: ____________________________

(6) A collection of bijections from $\mathbb{R}^n$ to itself is said to form a *group* if it satisfies all these three conditions:
- The composite of any two (possibly equal, possibly distinct) bijections in the collection is also in the collection.
- The identity bijection (i.e., the map sending every vector to itself) is in the collection.
- For every bijection in the collection, the inverse bijection is also in the collection.

For fixed n, which of the following collections of bijections from $\mathbb{R}^n$ to itself does *not* form a group? Please see Option (E) before answering.
(A) The collection of all linear automorphisms of $\mathbb{R}^n$
(B) The collection of all affine linear automorphisms of $\mathbb{R}^n$
(C) The collection of all self-isometries of $\mathbb{R}^n$
(D) The collection of all self-homotheties of $\mathbb{R}^n$
(E) None of the above, i.e., each of them is a group

Your answer: ____________________________

For the remaining questions, we deal with the case $n = 2$.

We consider two special types of bijections from $\mathbb{R}^2$ to $\mathbb{R}^2$: *rotations* (a rotation is specified by the center of rotation and the angle of rotation) and *reflections* (a reflection is specified by the line of reflection).

The identity map (i.e., the map sending every point to itself) is considered both a translation and a rotation. It is the translation by the zero vector. It can be viewed as a rotation about any point by the zero angle.

Note that for a rotation, the angle of rotation is determined uniquely up to additive multiples of 2π. The center of rotation is determined uniquely for all nontrivial rotations.

(7) What is the composite of two rotations centered at the same point in $\mathbb{R}^2$? Assume for simplicity that the composite is not the identity, i.e., the two rotations do not cancel each other. Note that the rotations must commute, so the order of operation does not matter.

(A) It must be a reflection about a line passing through that center point.
(B) It must be a reflection about a line *not* passing through that center point.
(C) It must be a rotation centered at the same point
(D) It must be a rotation but it need not be centered at the same point.
(E) It must be a translation.

Your answer: ___________________________

(8) What is the composite of two reflections about lines in $\mathbb{R}^2$, if the two lines of reflection are known to be parallel but distinct? Although the two reflections do not commute, the *type* of their composite does not depend upon the order in which we compose them.
(A) It must be a reflection about a third line which is parallel to both the lines and is equidistant from them.
(B) It must be a reflection about a third line which is perpendicular to both the lines.
(C) It must be a rotation about a point that is equidistant from both lines.
(D) It must be a translation by a vector parallel to the lines about which we are reflecting.
(E) It must be a translation by a vector perpendicular to the lines about which we are reflecting.

Your answer: ___________________________

(9) What is the composite of two reflections about lines in $\mathbb{R}^2$, if the two lines of reflection are distinct and intersect? Once again, the reflections do not in general commute, but the *type* of the composite does not depend on the order of composition.
(A) It must be a reflection about a third line which passes through the point of intersection of the two lines of reflection.
(B) It must be a reflection about a third line which does not pass through the point of intersection of the two lines of reflection.
(C) It must be a rotation about the point of intersection.
(D) It must be a translation by a vector that makes equal angles with both the lines.
(E) It need not be a translation, rotation, or reflection.

Your answer: ___________________________

(10) What is the composite of a nontrivial rotation in $\mathbb{R}^2$ (i.e., the angle of rotation is not a multiple of 2π) and a nontrivial translation?
(A) It must be a rotation with the same center of rotation but with a different angle of rotation.
(B) It must be a rotation with the same angle of rotation but with a different center of rotation.
(C) It must be a reflection about a line passing through the center of rotation.
(D) It must be a reflection about a line *not* passing through the center of rotation.
(E) It must be a translation.

Your answer: ___________________________

(11) An affine linear automorphism of $\mathbb{R}^2$ is termed *area-preserving* if it preserves areas, i.e., the area of the image of any triangle under the automorphism is the same as the area of the original triangle. What is the relation between being a self-isometry and being an area-preserving affine linear automorphism of $\mathbb{R}^2$?
(A) Being a self-isometry is precisely the same as being an area-preserving affine linear automorphism.
(B) Every self-isometry is area-preserving, but not every area-preserving affine linear automorphism is a self-isometry.
(C) Every area-preserving affine linear automorphism is a self-isometry, but not every self-isometry is area-preserving.
(D) A self-isometry need not be an area-preserving affine linear automorphism, and an area-preserving affine linear automorphism need not be a self-isometry.

Your answer: ____________________________

(12) An affine linear automorphism of $\mathbb{R}^2$ is termed *orientation-preserving* if it preserves orientation, i.e., it does not interchange left with right. An affine linear automorphism of $\mathbb{R}^2$ is termed *orientation-reversing* if it reverses orientation, i.e., it interchanges the roles of left and right. Obviously, the composite of two orientation-preserving affine linear automorphisms is orientation-preserving. What can we say about the composite of two orientation-reversing affine linear automorphisms?
(A) It must be orientation-preserving
(B) It must be orientation-reversing
(C) It may be orientation-preserving or orientation-reversing

Your answer: ____________________________

(13) The linear automorphism of $\mathbb{R}^2$ with matrix:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

is an example of a *shear automorphism*. Which of the following is this automorphism *not*? Please see options (D) and (E) before answering.
(A) Area-preserving
(B) Orientation-preserving
(C) Self-isometry
(D) None of the above, i.e., it is area-preserving, orientation-preserving, and a self-isometry
(E) All of the above, i.e., it is not area-preserving, not orientation-preserving, and not a self-isometry of $\mathbb{R}^2$.

Your answer: ____________________________

(14) Which of the following is guaranteed to send any triangle in $\mathbb{R}^2$ to a similar triangle? Please see Options (D) and (E) before answering.
(A) Linear automorphism
(B) Affine linear automorphism
(C) Self-homothety
(D) All of the above
(E) None of the above

Your answer: ____________________________

DIAGNOSTIC IN-CLASS QUIZ: DUE WEDNESDAY NOVEMBER 6: IMAGE AND KERNEL (BASIC)

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____

PLEASE DO NOT DISCUSS *ANY* QUESTIONS.

The questions here test for a very rudimentary understanding of the ideas covered in the lectures notes titled `Image and kernel of a linear transformation`. The corresponding section of the book is Section 3.1.

(1) *Do not discuss this!*: For a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$, the kernel of T is defined as the set of vectors $\vec{x} \in \mathbb{R}^m$ satisfying the condition that $T(\vec{x}) = \vec{0}$. Which of the following correctly describes the type of subset of $\mathbb{R}^m$ that the kernel must be? Note that, as usual, we identify a set of vectors with the set of corresponding points.
(A) The kernel is a line segment in $\mathbb{R}^m$.
(B) The kernel is a linear subspace of $\mathbb{R}^m$, i.e., it passes through the origin and, for any two points in the kernel, the line joining them is completely inside the kernel.
(C) The kernel is an affine linear subspace of $\mathbb{R}^m$ but it need not be linear, i.e., it is non-empty and the line joining any two points in it is also in it, but it need not contain the origin.
(D) The kernel is a curve in $\mathbb{R}^m$ with a parametric description.

Your answer: ____________________________

(2) *Do not discuss this!*: For a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$, the kernel of T is defined as the set of vectors $\vec{x} \in \mathbb{R}^m$ satisfying the condition that $T(\vec{x}) = \vec{0}$. Given a vector $\vec{y} \in \mathbb{R}^n$, the set of solutions to $T(\vec{x}) = \vec{y}$ is either empty, or it bears some relation with the kernel. What relation does it bear to the kernel if it is nonempty?
(A) The solution set is an affine linear subspace of $\mathbb{R}^m$ (see definition in Option (C) of Q1) that is a translate of the kernel, i.e., there is a vector $\vec{v}$ such that the vectors in the solution set are precisely the vectors expressible as ($\vec{v}$ plus a vector in the kernel).
(B) The solution set coincides precisely with the kernel.
(C) The solution set comprises a single point (i.e., a single vector) that is not in the kernel.

Your answer: ____________________________

(3) *Do not discuss this!*: Given a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$, recall that we say that T is *injective* if for every $\vec{y} \in \mathbb{R}^n$, there exists *at most* one $\vec{x} \in \mathbb{R}^m$ satisfying $T(\vec{x}) = \vec{y}$. Another way of formulating this is that if A is the $n \times m$ matrix for T, then the linear system $A\vec{x} = \vec{y}$ has at most one solution for $\vec{x}$ for each fixed value of $\vec{y}$. We had earlier worked out that this condition is equivalent to full column rank (recall: all the variables need to be leading variables), which in this case means rank m.

What is the relationship between the injectivity of T and the kernel of T?
(A) T is injective if and only if the kernel of T is empty.
(B) If T is injective, then the kernel of T is empty. However, the converse is not in general true.
(C) T is injective if and only if the kernel of T comprises only the zero vector.
(D) If T is injective, then the kernel of T comprises only the zero vector. However, the converse is not in general true.
(E) If the kernel of T comprises only the zero vector, then T is injective. However, the converse is not in general true.

Your answer: ____________________________

(4) *Do not discuss this!*: For a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$, the image of T is defined as the set of vectors $\vec{y} \in \mathbb{R}^n$ satisfying the condition that there exists a vector $\vec{x} \in \mathbb{R}^m$ satisfying $T(\vec{x}) = \vec{y}$. In other words, the image of T equals the range of T as a function. Which of the following correctly describes the type of subset of $\mathbb{R}^n$ that the image must be? Note that, as usual, we identify a set of vectors with the set of corresponding points.
(A) The image is a line segment in $\mathbb{R}^n$.
(B) The image is a linear subspace of $\mathbb{R}^n$, i.e., it passes through the origin and, for any two points in the image, the line joining them is completely inside the image.
(C) The image is an affine linear subspace of $\mathbb{R}^n$ but it need not be linear, i.e., it is non-empty and the line joining any two points in it is also in it, but it need not contain the origin.
(D) The image is a curve in $\mathbb{R}^n$ with a parametric description.

Your answer: ____________________________

(5) *Do not discuss this!*: Given a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$, recall that we say that T is *surjective* if for every $\vec{y} \in \mathbb{R}^n$, there exists *at least* one $\vec{x} \in \mathbb{R}^m$ satisfying $T(\vec{x}) = \vec{y}$. Another way of formulating this is that if A is the $n \times m$ matrix for T, then the linear system $A\vec{x} = \vec{y}$ has at least one solution for $\vec{x}$ for each fixed value of $\vec{y}$. We had earlier worked out that this condition is equivalent to full row rank (recall: we need all rows in the rref to be nonzero in order to avoid the potential for inconsistency), which in this case means rank n.

What is the relationship between the surjectivity of T and the image of T?
(A) T is surjective if and only if the image of T is empty.
(B) T is surjective if and only if the image of T comprises only the zero vector.
(C) T is surjective if and only if the image of T is all of $\mathbb{R}^n$.

Your answer: ____________________________

DIAGNOSTIC IN-CLASS QUIZ: DUE FRIDAY NOVEMBER 8: IMAGE AND KERNEL (COMPUTATIONAL)

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____

PLEASE DO NOT DISCUSS *ANY* QUESTIONS.

The questions here test for an understanding of the ideas covered in the lecture notes titled `Image and kernel of a linear transformation`. However, the format of presentation of the questions in the quiz differs somewhat from that used in typical linear algebra problems, so you need to think a bit before plugging and chugging. The corresponding section of the book is Section 3.1.

All these questions can be solved without using any part of the "toolkit" of linear algebra, but they can be understood better and more deeply using the ideas and methods of linear algebra.

(1) *Do not discuss this!*: Consider the linear transformation $\text{Avg} : \mathbb{R}^2 \to \mathbb{R}^2$ defined as:

$$\text{Avg} = \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} (x+y)/2 \\ (x+y)/2 \end{bmatrix}$$

What can we say about the kernel and image of Avg? Note that in our descriptions of the kernel and the image below, we use x to denote the first coordinate of the vector and y to denote the second coordinate of the vector.

Note: One way you can do that is to write the matrix for Avg, but in this particular situation, it's easiest to just do things directly.

(A) The kernel is the zero subspace and the image is all of $\mathbb{R}^2$
(B) The kernel is the line $y = x$ and the image is also the line $y = x$
(C) The kernel is the line $y = x$ and the image is the line $y = -x$
(D) The kernel is the line $y = -x$ and the image is also the line $y = -x$
(E) The kernel is the line $y = -x$ and the image is the line $y = x$

Your answer: ____________________________

(2) *Do not discuss this!*: Consider the *average of other two* linear transformation $\nu : \mathbb{R}^3 \to \mathbb{R}^3$ given as follows:

$$\nu = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} (y+z)/2 \\ (z+x)/2 \\ (x+y)/2 \end{bmatrix}$$

What can we say about the kernel and image of ν?

Note that in our descriptions of the kernel and the image below, we use x to denote the first coordinate of the vector, y to denote the second coordinate of the vector, and z to denote the third coordinate of the vector.

Note: This can both be reasoned directly (without any knowledge of linear algebra) or alternatively it can be done by writing the matrix of ν and computing its rank, image, and kernel.

(A) The kernel is the zero subspace and the image is all of $\mathbb{R}^3$
(B) The kernel is the line $x = y = z$ (one-dimensional) and the image is the plane $x + y + z = 0$ (two-dimensional)
(C) The kernel is the plane $x + y + z = 0$ (two-dimensional) and the image is the line $x = y = z$ (one-dimensional)
(D) The kernel is the plane $x = y = z$ (two-dimensional) and the image is the line $x + y + z = 0$ (one-dimensional)

(E) The kernel is the line $x + y + z = 0$ (one-dimensional) and the image is the plane $x = y = z$ (two-dimensional)

Your answer: ____________________________

(3) *Do not discuss this!*: Consider the *difference of other two* linear transformation $\mu : \mathbb{R}^3 \to \mathbb{R}^3$ given by:

$$\mu = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} y - z \\ z - x \\ x - y \end{bmatrix}$$

What can we say about the kernel and image of μ?

Note that in our descriptions of the kernel and the image below, we use x to denote the first coordinate of the vector, y to denote the second coordinate of the vector, and z to denote the third coordinate of the vector.

Note: This can both be reasoned directly (without any knowledge of linear algebra) or alternatively it can be done by writing the matrix of μ and computing its rank, image, and kernel.

(A) The kernel is the zero subspace and the image is all of $\mathbb{R}^3$

(B) The kernel is the line $x = y = z$ (one-dimensional) and the image is the plane $x + y + z = 0$ (two-dimensional)

(C) The kernel is the plane $x + y + z = 0$ (two-dimensional) and the image is the line $x = y = z$ (one-dimensional)

(D) The kernel is the plane $x = y = z$ (two-dimensional) and the image is the line $x + y + z = 0$ (one-dimensional)

(E) The kernel is the line $x + y + z = 0$ (one-dimensional) and the image is the plane $x = y = z$ (two-dimensional)

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE FRIDAY NOVEMBER 8: LINEAR TRANSFORMATIONS: SEEDS FOR REAPING LATER

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

In this quiz, we will sow the seeds of ideas that we will reap later. There are two broad classes of ideas that we touch upon here:

- Conjugation, similarity transformations, and products of matrices: This will be of relevance later when we discuss change of coordinates. We cover change of coordinates in more detail in Section 3.4 of the text.
- Kernel and image for linear transformations arising from calculus, typically for infinite-dimensional spaces: This will be helpful in understanding linear transformations in an *abstract* sense, a topic that we cover in more detail in Chapter 4 of the text.

(1) Suppose A and B are (possibly equal, possibly distinct) $n \times n$ matrices for some $n > 1$. Recall that the *trace* of a matrix is defined as the sum of its diagonal entries. Suppose $C = AB$ and $D = BA$. Which of the following is true?
  (A) It must be the case that $C = D$
  (B) The *set* of entry values in C is the same as the set of entry values in D, but they may appear in a different order.
  (C) C and D need not be equal, but the sum of all the matrix entries of C must equal the sum of all the matrix entries of D.
  (D) C and D need not be equal, but they have the same diagonal, i.e., every diagonal entry of C equals the corresponding diagonal entry of D.
  (E) C and D need not be equal and they need not even have the same diagonal. However, they must have the same trace, i.e., the sum of the diagonal entries of C equals the sum of the diagonal entries of D.

Your answer: ____________________________

Suppose A is an invertible $n \times n$ matrix. The *conjugation operation* corresponding to A is the map that sends any $n \times n$ matrix X to AXA^{-1}. We can verify that the following hold for any two (possibly equal, possibly distinct) $n \times n$ matrices X and Y:

$$\begin{aligned} A(X+Y)A^{-1} &= AXA^{-1} + AYA^{-1} \\ A(XY)A^{-1} &= (AXA^{-1})(AYA^{-1}) \\ AX^rA^{-1} &= (AXA^{-1})^r \end{aligned}$$

The conceptual significance of this will (hopefully!) become clearer as we proceed.

(2) Which of the following is guaranteed to be the same for X and AXA^{-1}?
  (A) The sum of all entries
  (B) The sum of squares of all entries
  (C) The product of all entries
  (D) The sum of all diagonal entries (i.e., the trace)
  (E) The sum of squares of all diagonal entries

Your answer: ____________________________

(3) A and X are $n \times n$ matrices, with A invertible. Which of the following is/are true? Please see Options (D) and (E) before answering, and select a single option that best reflects your view.
(A) X is invertible if and only if AXA^{-1} is invertible.
(B) X is nilpotent if and only if AXA^{-1} is nilpotent.
(C) X is idempotent if and only if AXA^{-1} is idempotent.
(D) All of the above.
(E) None of the above.

Your answer: ____________________________

(4) A and X are $n \times n$ matrices, with A invertible. Which of the following is equivalent to the condition that $AXA^{-1} = X$?
(A) $A + X = X + A$
(B) $A - X = X - A$
(C) $AX = XA$
(D) $XA^{-1} = AX^{-1}$
(E) None of the above

Your answer: ____________________________

Let's look at a computational application of matrix conjugation.

One computational application is power computation. Suppose we have a $n \times n$ matrix B and we need to compute B^r for a very large r. This requires $O(\log_2 r)$ multiplications, but note that each multiplication, if done naively, takes time $O(n^3)$ for a generic matrix. Suppose, however, that there exists a matrix A such that the matrix $C = ABA^{-1}$ is diagonal. If we can find A (and hence C) efficiently, then we can compute $C^r = (ABA^{-1})^r = AB^rA^{-1}$, and therefore $B^r = A^{-1}C^rA$. Note that each multiplication of diagonal matrices takes $O(n)$ multiplications, so this reduces the overall arithmetic complexity from $O(n^3 \log_2 r)$ to $O(n \log_2 r)$. Note, however, that this is contingent on our being able to find the matrices A and C first. We will later see a method for finding A and C. Unfortunately, this method relies on finding the set of solutions to a polynomial equation of degree n, which requires operations that go beyond ordinary arithmetic operations of addition, subtraction, multiplication, and division. Even in the case $n = 2$, it requires solving a quadratic equation. We do have the formula for that.

(5) Consider the following example of the above general setup with $n = 2$:

$$B = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$$

We can choose:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

The matrix $C = ABA^{-1}$ is a diagonal matrix. What diagonal matrix is it?
(A) $\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$
(B) $\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$
(C) $\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$
(D) $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$
(E) $\begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}$

Your answer: ____________________________

(6) With A, B, and C as in the preceding question, what is the value of B^8? Use that $2^8 = 256$.

(A) $\begin{bmatrix} 1 & -1 \\ 0 & 256 \end{bmatrix}$

(B) $\begin{bmatrix} 1 & -255 \\ 0 & 256 \end{bmatrix}$

(C) $\begin{bmatrix} 1 & 253 \\ 0 & 256 \end{bmatrix}$

(D) $\begin{bmatrix} 1 & 253 \\ 254 & 256 \end{bmatrix}$

(E) $\begin{bmatrix} 16 & -8 \\ 0 & 256 \end{bmatrix}$

Your answer: ______________________________

(7) Suppose $n > 1$. Let A be a $n \times n$ matrix such that the linear transformation corresponding to A is a self-isometry of $\mathbb{R}^n$, i.e., it preserves distances. Which of the following must necessarily be true? You can use the case $n = 2$ and the example of rotations to guide your thinking.

(A) The trace of A (i.e., the sum of the diagonal entries of A) must be equal to 0
(B) The trace of A (i.e., the sum of the diagonal entries of A) must be equal to 1
(C) The sum of the entries in each column of A must be equal to 1
(D) The sum of the absolute values of the entries in each column of A must be equal to 1
(E) The sum of the squares of the entries in each column of A must be equal to 1

Your answer: ______________________________

A *real vector space* (just called *vector space* for short) is a set V equipped with the following structures:

- A binary operation $+$ on V called addition that is commutative and associative.
- A special element $0 \in V$ that is an identity for addition.
- A scalar multiplication operation $\mathbb{R} \times V \to V$ denoted by concatenation such that:
  - $0\vec{v} = 0$ (the 0 on the right side being the vector 0) for all $\vec{v} \in V$.
  - $1\vec{v} = \vec{v}$ for all $\vec{v} \in V$.
  - $a(b\vec{v}) = (ab)\vec{v}$ for all $a, b \in \mathbb{R}$ and $\vec{v} \in V$.
  - $a(\vec{v} + \vec{w}) = a\vec{v} + a\vec{w}$ for all $a \in \mathbb{R}$ and $\vec{v}, \vec{w} \in V$.
  - $(a + b)\vec{v} = a\vec{v} + b\vec{v}$ for all $a, b \in \mathbb{R}$, $\vec{v} \in V$.

A *subspace* of a vector space is defined as a nonempty subset that is closed under addition and scalar multiplication. In particular, any subspace must contain the zero vector. A subspace of a vector space can be viewed as being a vector space in its own right.

Suppose V and W are vector spaces. A function $T : V \to W$ is termed a *linear transformation* if T preserves addition and scalar multiplication, i.e., we have the following two conditions:

- $T(\vec{v_1} + \vec{v_2}) = T(\vec{v_1}) + T(\vec{v_2})$ for all vectors $\vec{v_1}, \vec{v_2} \in V$.
- $T(a\vec{v}) = aT(\vec{v})$ for all $a \in \mathbb{R}$, $\vec{v} \in V$.

The *kernel* of a linear transformation T is defined as the set of all vectors $\vec{v}$ such that $T(\vec{v})$ is the zero vector. The *image* of a linear transformation T is defined as its range as a set map.

Denote by $C(\mathbb{R})$ (or alternatively by $C^0(\mathbb{R})$) the vector space of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$, with pointwise addition and scalar multiplication. Note that the elements of this vector space, which we would ordinarily call "vectors", are now *functions*.

For k a positive integer, denote by $C^k(\mathbb{R})$ the subspace of $C(\mathbb{R})$ comprising those continuous functions that are at least k times *continuously* differentiable. Note that $C^{k+1}(\mathbb{R})$ is a subspace of $C^k(\mathbb{R})$, so we have a descending chain of subspaces:

$$C(\mathbb{R}) = C^0(\mathbb{R}) \supseteq C^1(\mathbb{R}) \supseteq C^2(\mathbb{R}) \supseteq \dots$$

The intersection of these spaces is the vector space $C^\infty(\mathbb{R})$, defined as the subspace of $C(\mathbb{R})$ comprising those functions that are *infinitely* differentiable.

(8) We can think of differentiation as a linear transformation. Of the following options, which is the broadest way of viewing differentiation as a linear transformation? By "broadest" we mean "with the largest domain that makes sense among the given options."
(A) From $C^\infty(\mathbb{R})$ to $C^\infty(\mathbb{R})$
(B) From $C^0(\mathbb{R})$ to $C^1(\mathbb{R})$
(C) From $C^1(\mathbb{R})$ to $C^0(\mathbb{R})$
(D) From $C^1(\mathbb{R})$ to $C^2(\mathbb{R})$
(E) From $C^2(\mathbb{R})$ to $C^1(\mathbb{R})$

Your answer: ____________________________

(9) Under the differentiation linear transformation, what is the image of $C^k(\mathbb{R})$ for a positive integer k?
(A) $C^{k-1}(\mathbb{R})$
(B) $C^k(\mathbb{R})$
(C) $C^{k+1}(\mathbb{R})$
(D) $C^1(\mathbb{R})$
(E) $C^\infty(\mathbb{R})$

Your answer: ____________________________

(10) What is the kernel of differentiation?
(A) The vector space of all constant functions
(B) The vector space of all linear functions (i.e., functions of the form $x \mapsto mx + c$ with $m, c \in \mathbb{R}$)
(C) The vector space of all polynomial functions
(D) $C^\infty(\mathbb{R})$
(E) $C^1(\mathbb{R})$

Your answer: ____________________________

(11) Suppose k is a positive integer greater than 2. Consider the operation of "differentiating k times." This is a linear transformation that can be defined as the k-fold composite of differentiation with itself. Viewed most generally, this is a linear transformation from $C^k(\mathbb{R})$ to $C(\mathbb{R})$. What is the kernel of this linear transformation?
(A) The set of all constant functions
(B) The set of all polynomial functions of degree at most $k - 1$
(C) The set of all polynomial functions of degree at most k
(D) The set of all polynomial functions of degree at most $k + 1$
(E) The set of all polynomial functions

Your answer: ____________________________

(12) Suppose k is a positive integer greater than 2. Consider the set P_k of all polynomial functions of degree at most k. This set is a vector subspace of $C(\mathbb{R})$. Of the following subspaces of $C(\mathbb{R})$, which is the *smallest* subspace of which P_k is a subspace?
(A) $C^1(\mathbb{R})$
(B) $C^{k-1}(\mathbb{R})$
(C) $C^k(\mathbb{R})$
(D) $C^{k+1}(\mathbb{R})$
(E) $C^\infty(\mathbb{R})$

Your answer: ____________________________

Two more definitions of use. A *linear functional* on a vector space V is a linear transformation from V to $\mathbb{R}$, where $\mathbb{R}$ is viewed as a one-dimensional vector space over itself in the obvious way.

We define $C([0, 1])$ as the set of all continuous functions from $[0, 1]$ to $\mathbb{R}$ with pointwise addition and scalar multiplication.

(13) Which of the following is *not* a linear functional on $C([0, 1])$?

(A) $f \mapsto f(0)$
(B) $f \mapsto f(1)$
(C) $f \mapsto \int_0^1 f(x)\,dx$
(D) $f \mapsto \int_0^1 f(x^2)\,dx$
(E) $f \mapsto \int_0^1 (f(x))^2\,dx$

Your answer: ______________________________

TAKE-HOME CLASS QUIZ: DUE MONDAY NOVEMBER 11: USING LINEAR SYSTEMS FOR MEASUREMENT

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ______________________________________

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

The purpose of this quiz is two-fold. First, many of the ideas you saw early on in the course (in the second and third week) may be on the verge of fading out. Drawing from the best research on *spaced repetition* (see for instance `http://en.wikipedia.org/wiki/Spaced_repetition`) it's high time we tried recalling some of that stuff. But with a twist, because we can now use some concepts from later topics we've seen to refine our past understanding.

The second purpose is to prepare you for what we hope to eventually get to: a deep and rich understanding of linear algebra as it's *used*: in linear regressions, computing correlations, and more fancy applications like factor analysis and principal component analysis. The third question, in particular, relates to the central idea behind linear regression (specifically, ordinary least squares regression). The questions also relate, albeit not very directly, to the broad ideas behind factor analysis and principal component analysis.

For the questions here, assume two dimensions of a person's general cognitive ability: verbal and mathematical. Denote by g_v the person's general verbal ability and by g_m the person's general mathematical ability.

Various ability tests can be devised that aim to test for the person's abilities. However, it is not possible to construct a test that *solely* measures g_v or *solely* measure g_m. Different tests measure g_v and g_m to different extents. For instance, an ordinary numerical computation test might measure mostly g_m. On the other hand, a test similar to the quizzes in this course might measure both g_v and g_m a fair amount, given how much you have to read to answer the quiz questions.

Of course, the score on a given test could depend on a lot of factors other than general abilities. Some of them could be systematic: a student with poor mathematical abilities in general may have "trained for the test." As an example, using a calculus test to test for general mathematical ability might mean that people who have happened to take calculus do a lot better than people who haven't, but have similar general mathematical ability. Some are more ephemeral, such as students guessing answers, mood fluctuations, and other context-specific factors that affect scores.

For simplicity, we will assume that there are no systemic factors other than general verbal and general mathematical ability that the test is measuring. For even greater simplicity, assume that the (expected) test score is linear in g_v and g_m with zero intercept, i.e, the score is of the form $w_v g_v + w_m g_m$ where w_v and w_m are real numbers that serve as *weights*. Note that this assumes that there is no *interaction* between the verbal and mathematical skills in determining the score.

The assumption may or may not be realistic. For instance, a question (such as those on this quiz!) that requires a lot of reading *and* strong math skills would probably have an expected score formula that is *multiplicative* in g_v and g_m: having zero or near-zero verbal ability means you will be unable to do the question, even if your mathematical ability is awesome. Similarly, having zero or near-zero mathematical ability means you will be unable to do the question, even if your verbal ablity is awesome. Multiplicatively separable functions are better suited to capture this sort of dependence. However, even if the test has questions of this sort, we can take logs on test scores and make them additively separable, so the additive model may still work well.

The assumption of *linearity* goes further, but this too might be realistic.

My goal is to use one or more tests in order to determine the true values of a student's g_v and g_m. Another formulation is that my goal is to determine the vector:

$$\vec{g} = \begin{bmatrix} g_v \\ g_m \end{bmatrix}$$

(1) I administer two tests to a student. The student's score s_1 on the first test is $2g_v + 3g_m$ while the score s_2 on the second test is $3g_v + 5g_m$. How do I recover g_v and g_m from s_1 and s_2?
(A) $g_v = 2s_1 + 3s_2$, $g_m = 3s_1 + 5s_2$
(B) $g_v = 2s_1 - 3s_2$, $g_m = 3s_1 - 5s_2$
(C) $g_v = 5s_1 + 3s_2$, $g_m = 3s_1 + 2s_2$
(D) $g_v = 5s_1 - 3s_2$, $g_m = 3s_1 - 2s_2$
(E) $g_v = 5s_1 - 3s_2$, $g_m = -3s_1 + 2s_2$

Your answer: ____________________________

(2) In order to combat the problem of uncertainty about my model, I decide to administer more than two tests. I administer a total of n tests. The score on the i^{th} test is $s_i = w_{i,v}g_v + w_{i,m}g_m$. The score vector $\vec{s}$ has coordinates $s_i, 1 \le i \le n$.

If there is no measurement error and the student's actual score in each test equals the student's expected score, then we have a system of n simultaneous linear equations in 2 variables.

Let W be the matrix:

$$\begin{bmatrix} w_{1,v} & w_{1,m} \\ w_{2,v} & w_{2,m} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ w_{n,v} & w_{n,m} \end{bmatrix}$$

Assume that all entries of W are positive, i.e., each test tests to a nonzero extent for both verbal and mathematical ability.

Our goal is to "solve for" $\vec{g}$ the following vector equation:

$$W\vec{g} = \vec{s}$$

What is the necessary and sufficient condition on W so that the equation has at most one solution for $\vec{g}$ for each $\vec{s}$? If $\vec{s}$ arises from an actual $\vec{g}$, i.e., it is a true score vector, then note that there will be a solution.
(A) All the ratios $w_{i,v} : w_{i,m}$ are the same.
(B) All the ratios $w_{i,v} : w_{i,m}$ are different.
(C) At least two of the ratios $w_{i,v} : w_{i,m}$ are the same.
(D) At least two of the ratios $w_{i,v} : w_{i,m}$ are different.

Your answer: ____________________________

(3) Use notation as in the previous question. Suppose that there is some measurement error, so that instead of getting the true score vector $\vec{s}$, I have a somewhat distorted score vector $\vec{t}$. How do I go about recovering my "best guess" for $\vec{s}$ from $\vec{t}$?
(A) Find the closest vector to $\vec{t}$ in the kernel of the linear transformation corresponding to W.
(B) Find the closest vector to $\vec{t}$ in the image of the linear transformation corresponding to W.
(C) Find the farthest vector from $\vec{t}$ in the kernel of the linear transformation corresponding to W.
(D) Find the farthest vector from $\vec{t}$ in the image of the linear transformation corresponding to W.

Your answer: ____________________________

(4) Suppose I want to introduce a *new* test that tests for both verbal and mathematical ability with expected score of the form $w_v g_v + w_m g_m$, but the values w_v and w_m are currently unknown. My

strategy is as follows. I find two students. I administer a bunch of tests with *known* w_v and w_m values to those students. I use those tests to find the g_v and g_m values for both students. Then, I administer the new test to both students and try to determine the values of w_v and w_m. Assume no measurement error.

Of course, I want the matrix W of the *known* tests to satisfy the condition of Question 2. What additional criteria would I wish of the two students I use for this in order to correctly determine g_v and g_m? Note that it will not be possible to be sure of this in advance, but one can still pick student pairs who are more likely to satisfy the criterion and thus avoid waste of effort.

(A) The students should have the same $g_v + g_m$ value.
(B) The students should have different $g_v + g_m$ values.
(C) The students should have the same $g_v : g_m$ ratio.
(D) The students should have different $g_v : g_m$ ratios.

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE WEDNESDAY NOVEMBER 13: MATRIX MULTIPLICATION: ROWS, COLUMNS, ORTHOGONALITY, AND OTHER MISCELLANEA

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____________________

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

The purpose of this quiz is two-fold. First, many of the ideas related to matrix multiplication are at the stage where a bit of review will help prevent their fading out. Drawing from the best research on *spaced repetition* (see for instance `http://en.wikipedia.org/wiki/Spaced_repetition`) we will try to recall some of the stuff. But with a twist, because we consider it from a somewhat different angle.

Second, the new angle will also turn out to be useful for later material.

For Questions 1-5: Given a n-dimensional vector $\langle a_1, a_2, \dots, a_n \rangle \in \mathbb{R}^n$, the vector can be interpreted as a $n \times 1$ matrix (a column vector). This is the default interpretation. But there are also two other interpretations: as a $1 \times n$ matrix (a row vector) and as a diagonal $n \times n$ matrix.

Also note that for Questions 1-5, all the three ways of representing vectors coincide with each other for $n = 1$, so the questions are uninteresting for $n = 1$ because all answer options are equivalent. You may therefore assume that $n > 1$ for these questions, though obviously the correct answers are correct for $n = 1$ as well.

(1) Suppose I want to add two vectors $\vec{a} = \langle a_1, a_2, \dots, a_n \rangle$ and $\vec{b} = \langle b_1, b_2, \dots, b_n \rangle$ to obtain the output vector $\langle a_1 + b_1, a_2 + b_2, \dots, a_n + b_n \rangle$ using matrix addition. What format (row vector, column vector, or diagonal matrix) should I use? Please see Option (D) before answering and select the option that best describes your view.
  (A) Represent both $\vec{a}$ and $\vec{b}$ as row vectors and interpret the sum as a row vector.
  (B) Represent both $\vec{a}$ and $\vec{b}$ as column vectors and interpret the sum as a column vector.
  (C) Represent both $\vec{a}$ and $\vec{b}$ as diagonal matrices and interpret the sum as a diagonal matrix.
  (D) We can use any of the above.

  Your answer: ____________________

(2) Suppose I want to perform coordinate-wise multiplication on two vectors. Explicitly, I have two vectors $\vec{a} = \langle a_1, a_2, \dots, a_n \rangle$ and $\vec{b} = \langle b_1, b_2, \dots, b_n \rangle$ and I want to obtain the output vector $\langle a_1 b_1, a_2 b_2, \dots, a_n b_n \rangle$ using matrix multiplication (with the matrix for $\vec{a}$ written on the left and the matrix for $\vec{b}$ written on the right). What format (row vector, column vector, or diagonal matrix) should I use? Please see Option (D) before answering and select the option that best describes your view.
  (A) Represent both $\vec{a}$ and $\vec{b}$ as row vectors and interpret the matrix product as a row vector.
  (B) Represent both $\vec{a}$ and $\vec{b}$ as column vectors and interpret the matrix product as a column vector.
  (C) Represent both $\vec{a}$ and $\vec{b}$ as diagonal matrices and interpret the matrix product as a diagonal matrix.
  (D) We can use any of the above.

  Your answer: ____________________

(3) Suppose I am given two vectors $\vec{a} = \langle a_1, a_2, \dots, a_n \rangle$ and $\vec{b} = \langle b_1, b_2, \dots, b_n \rangle$ and I want to obtain a 1×1 matrix with entry $\sum_{i=1}^n a_i b_i$ using matrix multiplication (with the matrix for $\vec{a}$ written on the left and the matrix for $\vec{b}$ written on the right). What format (row vector, column vector, or diagonal matrix) should I use?
  (A) Represent both $\vec{a}$ and $\vec{b}$ as row vectors.

(B) Represent both $\vec{a}$ and $\vec{b}$ as column vectors.
(C) Represent both $\vec{a}$ and $\vec{b}$ as diagonal matrices.
(D) Represent $\vec{a}$ as a row vector and $\vec{b}$ as a column vector.
(E) Represent $\vec{a}$ as a column vector and $\vec{b}$ as a row vector.

Your answer: ____________________________

(4) Suppose I am given three vectors $\vec{a} = \langle a_1, a_2, \dots, a_n \rangle$, $\vec{b} = \langle b_1, b_2, \dots, b_n \rangle$, and $\vec{c} = \langle c_1, c_2, \dots, c_n \rangle$. I want to obtain a 1×1 matrix with entry $\sum_{i=1}^{n}(a_i b_i c_i)$ using matrix multiplication (with the matrix for $\vec{a}$ written on the left, the matrix for $\vec{b}$ written in the middle, and the matrix for $\vec{c}$ written on the right). What format should I use?
(A) $\vec{a}$ as a row vector, $\vec{b}$ as a column vector, $\vec{c}$ as a diagonal matrix.
(B) $\vec{a}$ as a column vector, $\vec{b}$ as a row vector, $\vec{c}$ as a diagonal matrix.
(C) $\vec{a}$ as a diagonal matrix, $\vec{b}$ as a row vector, $\vec{c}$ as a column vector.
(D) $\vec{a}$ as a column vector, $\vec{b}$ as a diagonal matrix, $\vec{c}$ as a row vector.
(E) $\vec{a}$ as a row vector, $\vec{b}$ as a diagonal matrix, $\vec{c}$ as a column vector.

Your answer: ____________________________

The next few questions rely on the concept of orthogonality (*orthogonal* is a synonym for *perpendicular* or *at right angles*). We say that two vectors (of the same dimension) are orthogonal if their dot product is zero. By this definition, the zero vector of a given dimension is orthogonal to every vector of that dimension. Note that it does not make sense to talk of orthogonality for vectors with different dimensions, i.e., with different numbers of coordinates.

(5) Suppose A is a $n \times m$ matrix. We can think of solving the system $A\vec{x} = \vec{0}$ (where $\vec{x}$ is a $m \times 1$ column vector of unknowns) as trying to find all the vectors orthogonal to all the vectors in a given set of vectors. What set of vectors is that?
(A) The set of row vectors of A, i.e., the rows of A, viewed as m-dimensional vectors.
(B) The set of column vectors of A, i.e., the columns of A, viewed as n-dimensional vectors.

Your answer: ____________________________

(6) Suppose A is a $p \times q$ matrix and B is a $q \times r$ matrix where p, q, and r are positive integers. The matrix product AB is a $p \times r$ matrix. What orthogonality condition corresponds to the condition that the matrix product AB is a zero matrix (i.e., all its entries are zero)?
(A) Every row of A is orthogonal to every row of B.
(B) Every row of A is orthogonal to every column of B.
(C) Every column of A is orthogonal to every row of B.
(D) Every column of A is orthogonal to every column of B.

Your answer: ____________________________

(7) Suppose A is an invertible $n \times n$ square matrix. Which of the following correctly characterizes the $n \times n$ matrix A^{-1} using orthogonality? Recall that AA^{-1} and $A^{-1}A$ are both equal to the $n \times n$ identity matrix.
(A) For every i in $\{1, 2, \dots, n\}$, the i^{th} row of A is orthogonal to the i^{th} row of A^{-1}. The dot product of the i^{th} row of A and the j^{th} row of A^{-1} for distinct i, j in $\{1, 2, \dots, n\}$ equals 1.
(B) For every i in $\{1, 2, \dots, n\}$, the i^{th} column of A is orthogonal to the i^{th} column of A^{-1}. The dot product of the i^{th} row of A and the j^{th} column of A^{-1} for distinct i, j in $\{1, 2, \dots, n\}$ equals 1.
(C) For every distinct i, j in $\{1, 2, \dots, n\}$, the i^{th} row of A is orthogonal to the j^{th} row of A^{-1}. The dot product of the i^{th} row of A with the i^{th} row of A^{-1} equals 1.
(D) For every i in $\{1, 2, \dots, n\}$, the i^{th} row of A is orthogonal to the i^{th} column of A^{-1}. The dot product of the i^{th} row of A and the j^{th} column of A^{-1} for distinct i, j in $\{1, 2, \dots, n\}$ equals 1.

(E) For every distinct i, j in $\{1, 2, \dots, n\}$, the i^{th} row of A is orthogonal to the j^{th} column of A^{-1}. The dot product of the i^{th} row of A and the i^{th} column of A^{-1} equals 1.

Your answer: ____________________________

The remaining questions review your skills at abstract behavior prediction.

(8) Suppose n is a positive integer greater than 1. Which of the following is always true for two invertible $n \times n$ matrices A and B?
(A) $A + B$ is invertible, and $(A + B)^{-1} = A^{-1} + B^{-1}$
(B) $A + B$ is invertible, and $(A + B)^{-1} = B^{-1} + A^{-1}$
(C) $A + B$ is invertible, though neither of the formulas of the preceding two options is correct
(D) AB is invertible, and $(AB)^{-1} = A^{-1}B^{-1}$
(E) AB is invertible, and $(AB)^{-1} = B^{-1}A^{-1}$

Your answer: ____________________________

(9) Suppose n is a positive integer greater than 1. For a nilpotent $n \times n$ matrix C, define the *nilpotency* of C as the smallest positive integer r such that $C^r = 0$. Note that the nilpotency is not defined for a non-nilpotent matrix. Given two $n \times n$ matrices A and B, what is the relation between the nilpotencies of AB and BA?
(A) AB is nilpotent if and only if BA is nilpotent, and if so, their nilpotencies must be equal.
(B) AB is nilpotent if and only if BA is nilpotent, and if so, their nilpotencies must differ by 1.
(C) AB is nilpotent if and only if BA is nilpotent, and if so, their nilpotencies must either be equal or differ by 1.
(D) It is possible for AB to be nilpotent and BA to be non-nilpotent; however, *if* both are nilpotent, then their nilpotencies must be equal.
(E) It is possible for AB to be nilpotent and BA to be non-nilpotent, however *if* both are nilpotent, then their nilpotencies must differ by 1.

Your answer: ____________________________

(10) What is the smallest n for which there exist examples of invertible $n \times n$ matrices A and B such that $A \neq B$ but $A^2 = B^2$?
(A) 1
(B) 2
(C) 3
(D) 4
(E) This is not possible for any n.

Your answer: ____________________________

(11) What is the smallest n for which there exist examples of invertible $n \times n$ matrices A and B such that $A \neq B$ but $A^3 = B^3$?
(A) 1
(B) 2
(C) 3
(D) 4
(E) This is not possible for any n.

Your answer: ____________________________

(12) What is the smallest n for which there exist examples of invertible $n \times n$ matrices A and B such that $A \neq B$ but $A^2 = B^2$ and $A^3 = B^3$?
(A) 1

(B) 2
(C) 3
(D) 4
(E) This is not possible for any n.

Your answer: ______________________________

(13) What is the smallest n for which there exist examples of (not necessarily invertible) $n \times n$ matrices A and B such that $A \neq B$ but $A^2 = B^2$ and $A^3 = B^3$?
(A) 1
(B) 2
(C) 3
(D) 4
(E) This is not possible for any n.

Your answer: ______________________________

TAKE-HOME CLASS QUIZ: DUE FRIDAY NOVEMBER 15: IMAGE AND KERNEL

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____________________________________

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

The purpose of this quiz is to review in greater depth the ideas behind image and kernel. The goal of the first seven questions is to review the ideas of injectivity, surjectivity, and bijectivity in the context of arbitrary functions between sets. The purpose is two-fold: (i) to give a functions-based approach to justifying, intuitively and formally, facts about the effect of matrix multiplication on rank, and (ii) to hint at ways in which linear transformations behave better than other types of functions.

The corresponding lecture notes are titled `Image and kernel of a linear transformation` and the corresponding section of the text is Section 3.1.

Just as a reminder, a function $f : A \to B$ between sets A and B is said to be:

- *injective* if for every $b \in B$, there is *at most* one value of a such that $f(a) = b$. In other words, if we denote by $f^{-1}(b)$ the set $\{a \in A \mid f(a) = b\}$, then $|f^{-1}(b)| \le 1$ for all $b \in B$ (here $|f^{-1}(b)|$ denotes the size of the set $f^{-1}(b)$).
- *surjective* if for every $b \in B$, there is *at least* one value of a such that $f(a) = b$. In other words, if we denote by $f^{-1}(b)$ the set $\{a \in A \mid f(a) = b\}$, then $|f^{-1}(b)| \ge 1$ for all $b \in B$.
- *bijective* if for every $b \in B$, there is *exactly* one value of a such that $f(a) = b$. In other words, if we denote by $f^{-1}(b)$ the set $\{a \in A \mid f(a) = b\}$, then $|f^{-1}(b)| = 1$ for all $b \in B$.

(1) Suppose $g : A \to B$ and $f : B \to C$ are functions. The composite $f \circ g$ is a function from A to C. What can we say the relationship between the injectivity of $f \circ g$, the injectivity of f, and the injectivity of g?
   (A) $f \circ g$ is injective if and only if f and g are both injective.
   (B) If f and g are both injective, then $f \circ g$ is injective. However, $f \circ g$ being injective does not imply anything about the injectivity of either f or g.
   (C) If f and g are both injective, then $f \circ g$ is injective. If $f \circ g$ is injective, then at least one of f and g is injective, but we cannot conclusively say for any specific one of the two that it must be injective.
   (D) If f and g are both injective, then $f \circ g$ is injective. If $f \circ g$ is injective, then f is injective, but we do not have enough information to deduce whether g is injective.
   (E) If f and g are both injective, then $f \circ g$ is injective. If $f \circ g$ is injective, then g is injective, but we do not have enough information to deduce whether f is injective.

   Your answer: ____________________________

(2) Suppose $g : A \to B$ and $f : B \to C$ are functions. The composite $f \circ g$ is a function from A to C. What can we say the relationship between the surjectivity of $f \circ g$, the surjectivity of f, and the surjectivity of g?
   (A) $f \circ g$ is surjective if and only if f and g are both surjective.
   (B) If f and g are both surjective, then $f \circ g$ is surjective. However, $f \circ g$ being surjective does not imply anything about the surjectivity of either f or g.
   (C) If f and g are both surjective, then $f \circ g$ is surjective. If $f \circ g$ is surjective, then at least one of f and g is surjective, but we cannot conclusively say for any specific one of the two that it must be surjective.
   (D) If f and g are both surjective, then $f \circ g$ is surjective. If $f \circ g$ is surjective, then f is surjective, but we do not have enough information to deduce whether g is surjective.

(E) If f and g are both surjective, then $f \circ g$ is surjective. If $f \circ g$ is surjective, then g is surjective, but we do not have enough information to deduce whether f is surjective.

Your answer: ____________________________

(3) Suppose $g : A \to B$ and $f : B \to C$ are functions. The composite $f \circ g$ is a function from A to C. Suppose $f \circ g$ is bijective. What can we say about f and g individually?
(A) Both f and g must be bijective.
(B) Both f and g must be injective, but neither of them need be surjective.
(C) Both f and g must be surjective, but neither of them need be injective.
(D) f must be injective but need not be surjective. g must be surjective but need not be injective.
(E) f must be surjective but need not be injective. g must be injective but need not be surjective.

Your answer: ____________________________

(4) $g : A \to B$ and $f : B \to C$ are functions. The composite $f \circ g$ is a function from A to C. Suppose both f and g are surjective. Further, suppose that for every $b \in B$, $g^{-1}(b)$ has size m (for a fixed positive integer m) and for every $c \in C$, $f^{-1}(c)$ has size n (for a fixed positive integer n). Then, what can we say about the sizes of the fibers (i.e., the inverse images of points in C) under the composite $f \circ g$?
(A) The size is $\min\{m, n\}$
(B) The size is $\max\{m, n\}$
(C) The size is $m + n$
(D) The size is mn
(E) The size is m^n

Your answer: ____________________________

(5) **PLEASE READ THIS VERY CAREFULLY AND CONSIDER A WIDE VARIETY OF POLYNOMIAL EXAMPLES**: Suppose f is a polynomial function of degree $n > 2$ from $\mathbb{R}$ to $\mathbb{R}$. What can we say about the fibers of f, i.e., the sets of the form $f^{-1}(x)$, $x \in \mathbb{R}$?

Hint: At the one extreme, consider a polynomial of the form x^n. Consider the sizes of the fibers $f^{-1}(0)$ and $f^{-1}(x)$ for a positive value of x (the fiber size for the latter will depend on whether n is even or odd). Alternatively, consider a polynomial of the form $(x-1)(x-2)\dots(x-n)$. Consider the size of the fiber $f^{-1}(0)$.
(A) Every fiber has size n.
(B) The minimum of the sizes of fibers is exactly n, but every fiber need not have size n.
(C) The maximum of the sizes of fibers is exactly n, but every fiber need not have size n.
(D) The minimum of the sizes of fibers is at least n, but need not be exactly n.
(E) The maximum of the sizes of fibers is at most n, but need not be exactly n.

Your answer: ____________________________

(6) Suppose f is a continuous injective function from $\mathbb{R}$ to $\mathbb{R}$. What can we say about the nature of f?
(A) f must be an increasing function on all of $\mathbb{R}$.
(B) f must be a decreasing function on all of $\mathbb{R}$.
(C) f must be a constant function on all of $\mathbb{R}$.
(D) f must be either an increasing function on all of $\mathbb{R}$ or a decreasing function on all of $\mathbb{R}$, but the information presented is insufficient to decide which case occurs.
(E) f must be either an increasing function or a decreasing function or a constant function on all of $\mathbb{R}$, but the information presented is insufficient for deciding anything stronger.

Your answer: ____________________________

(7) **PLEASE READ THIS CAREFULLY, MAKE CASES, AND CHECK YOUR REASONING**: Suppose f, g, and h are continuous bijective functions from $\mathbb{R}$ to $\mathbb{R}$. What can we say about the functions $f + g$, $f + h$, and $g + h$?

Hint: Based on the preceding question, you know something about the nature of f, g, and h individually as functions, but there is some degree of ambiguity in your knowledge. Make cases based on the possibilities and see what you can deduce in the best and worst case.

(A) They are all continuous bijective functions from $\mathbb{R}$ to $\mathbb{R}$.

(B) At least two of them are continuous bijective functions from $\mathbb{R}$ to $\mathbb{R}$. However, we cannot say more.

(C) At least one of them is a continuous bijective function from $\mathbb{R}$ to $\mathbb{R}$. However, we cannot say more.

(D) Either all three sums are continuous bijective functions from $\mathbb{R}$ to $\mathbb{R}$, or none is.

(E) It is possible that none of the sums is a continuous bijective functions from $\mathbb{R}$ to $\mathbb{R}$; it is also possible that one, two, or all the sums are continuous bijective functions from $\mathbb{R}$ to $\mathbb{R}$.

Your answer: ____________________________

The questions that follow tripped up students quite a bit last time, so I urge you to proceed with caution. You can do each of these questions in either of two ways:

- Using abstract, general reasoning.
- Constructing concrete examples.

While the former approach is one you should eventually be able to embrace without trepidation, feel free to rely on the latter approach for now. For this, consider matrices describing the linear transformations and use matrix multiplication to compute the composite where needed. Compute the kernel, image, and rank using the methods known to you. Take matrices such as those arising from finite state automata (as described in the "linear transformations and finite state automata" quiz) or their generalizations to rectangular matrices.

For instance, you might try taking a matrix such as $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. This describes a linear transformation $\mathbb{R}^5 \to \mathbb{R}^4$ and has rank three. The dimension of the kernel (inside $\mathbb{R}^5$) is 2 (explicitly, the kernel is precisely the set of vectors in $\mathbb{R}^5$ whose first three coordinates are zero) and the dimension of the image (inside $\mathbb{R}^4$) is 3 (explicitly, the image is precisely the set of vectors in $\mathbb{R}^4$ whose fourth coordinate is 0).

(8) *This is the analogue for linear transformations of Question 1*: Suppose m, n, p are positive integers. Suppose A is a $m \times n$ matrix and B is a $n \times p$ matrix. The product AB is a $m \times p$ matrix. Denote by T_A, T_B, and T_{AB} respectively the linear transformations corresponding to A, B, and AB. We have $T_A : \mathbb{R}^n \to \mathbb{R}^m$, $T_B : \mathbb{R}^p \to \mathbb{R}^n$, and $T_{AB} : \mathbb{R}^p \to \mathbb{R}^m$. Note that $T_{AB} = T_A \circ T_B$.

Recall that a matrix has full column rank if and only if the corresponding linear transformation is injective.

Which of the following describes correctly the relationship between A having full column rank (i.e., rank n), B having full column rank (i.e., rank p), and AB having full column rank (i.e., rank p)?

(A) AB has full column rank (i.e., rank p) if and only if A and B both have full column rank (ranks n and p respectively).

(B) If A and B both have full column rank, then AB has full column rank. However, AB having full column rank does not imply anything (separately or jointly) regarding whether A or B has full column rank.

(C) If A and B both have full column rank, then AB has full column rank. If AB has full column rank, then at least one of A and B has full column rank, but we cannot definitively say for any particular one of A and B that it must have full column rank.

(D) If A and B both have full column rank, then AB has full column rank. AB having full column rank implies that A has full column rank, but it does not tell us for sure that B has full column rank.

(E) If A and B both have full column rank, then AB has full column rank. AB having full column rank implies that B has full column rank, but it does not tell us for sure that A has full column rank.

Your answer: ____________________________

(9) *This is the analogue for linear transformations of Question 2*: Suppose m, n, p are positive integers. Suppose A is a $m \times n$ matrix and B is a $n \times p$ matrix. The product AB is a $m \times p$ matrix. Denote by T_A, T_B, and T_{AB} respectively the linear transformations corresponding to A, B, and AB. We have $T_A : \mathbb{R}^n \to \mathbb{R}^m$, $T_B : \mathbb{R}^p \to \mathbb{R}^n$, and $T_{AB} : \mathbb{R}^p \to \mathbb{R}^m$. Note that $T_{AB} = T_A \circ T_B$.

Recall that a matrix has full row rank if and only if the corresponding linear transformation is surjective.

Which of the following describes correctly the relationship between A having full row rank (i.e., rank m), B having full row rank (i.e., rank n), and AB having full row rank (i.e., rank m)?

(A) AB has full row rank if and only if A and B both have full row rank.

(B) If A and B both have full row rank, then AB has full row rank. However, AB having full row rank does not imply anything (separately or jointly) regarding whether A or B has full row rank.

(C) If A and B both have full row rank, then AB has full row rank. If AB has full row rank, then at least one of A and B has full row rank, but we cannot definitively say for any particular one of A and B that it must have full row rank.

(D) If A and B both have full row rank, then AB has full row rank. AB having full row rank implies that A has full row rank, but it does not tell us for sure that B has full row rank.

(E) If A and B both have full row rank, then AB has full row rank. AB having full row rank implies that B has full row rank, but it does not tell us for sure that A has full row rank.

Your answer: ____________________________

(10) *This is the analogue for linear transformations of Question 3*: Suppose m and n are positive integers. Suppose A is a $m \times n$ matrix and B is a $n \times m$ matrix. The product AB is a $m \times m$ matrix. The corresponding linear transformations are $T_A : \mathbb{R}^n \to \mathbb{R}^m$, $T_B : \mathbb{R}^m \to \mathbb{R}^n$, and $T_{AB} : \mathbb{R}^m \to \mathbb{R}^m$.

Suppose the square matrix AB has full rank m. What can we deduce about the ranks of A and B?

(A) Both A and B have full row rank, *and* both A and B have full column rank.

(B) Both A and B have full column rank, but neither of them need have full row rank.

(C) Both A and B have full row rank, but neither of them need have full column rank.

(D) A must have full column rank but need not have full row rank. B must have full row rank but need not have full column rank.

(E) A must have full row rank but need not have full column rank. B must have full column rank but need not have full row rank.

Your answer: ____________________________

For the coming questions, we will denote vector spaces by letters such as U, V, and W. You can, however, consider them to be finite-dimensional vector spaces of the form $\mathbb{R}^n$. However, you should take care not to use a letter for the dimension of a vector space if the letter is already in use elsewhere in the question. Also, you should take care to use different letters for the dimensions of different vector spaces, unless it is given to you that the vector spaces have the same dimension. The results also hold for infinite-dimensional vector spaces, but you can work on all the problems assuming you are working in the finite-dimensional setting.

(11) *This is an analogue for linear transformations of Question 4*: Suppose $T_1 : U \to V$ and $T_2 : V \to W$ are linear transformations. The composite $T_2 \circ T_1$ is also a linear transformation, this time from U to W. Suppose the kernel of T_1 has dimension m and the kernel of T_2 has dimension n. Suppose both T_1 and T_2 are surjective. What can you say about the dimension of the kernel of $T_2 \circ T_1$?

Please note this carefully: Although this question is analogous to Question 4, the correct answer options differ for the two questions. Here is an intuitive explanation for the relationship between the questions. Question 4 asked abot the *sizes* of the fibers. This question asks about the dimensions of the kernels. The fibers do correspond to the kernels. But the relationship between dimension and size is of a *logarithmic nature*. What we mean is that the dimension can be thought of as the logarithm of the size. This isn't literally true, because the size is infinite. But metaphorically, it makes sense, because, for instance, the dimension of $\mathbb{R}^p$ is the exponent p, and that comports with the laws of logarithms (similar to how the $\log_2(2^p) = p$).

(A) The dimension is $\min\{m, n\}$.
(B) The dimension is $\max\{m, n\}$.
(C) The dimension is $m + n$.
(D) The dimension is mn.
(E) The dimension is m^n.

Your answer: ____________________________

(12) Suppose $T_1 : U \to V$ and $T_2 : V \to W$ are linear transformations. The composite $T_2 \circ T_1$ is also a linear transformation, this time from U to W. Suppose the kernel of T_1 has dimension m and the kernel of T_2 has dimension n. However, unlike the preceding question, we are not given any information about the surjectivity of either T_1 or T_2. The answer to the preceding question gives an (inclusive) *upper* bound on the dimension of the kernel of $T_2 \circ T_1$. Which of the following is the best *lower* bound we can manage in general?

(A) $|m - n|$
(B) m
(C) n
(D) $m + n$

Your answer: ____________________________

(13) Suppose $T_1, T_2 : U \to V$ are linear transformations. Which of the following is true? Please see Options (D) and (E) before answering and select the single option that best reflects your view.

(A) If both T_1 and T_2 are injective, then $T_1 + T_2$ is injective.
(B) If both T_1 and T_2 are surjective, then $T_1 + T_2$ is surjective.
(C) If both T_1 and T_2 are bijective, then $T_1 + T_2$ is bijective.
(D) All of the above
(E) None of the above

Your answer: ____________________________

(14) Suppose $T_1, T_2 : U \to V$ are linear transformations. Which of the following best describes the relation between the kernels of T_1, T_2, and $T_1 + T_2$?

(A) The kernel of $T_1 + T_2$ equals the intersection of the kernel of T_1 and the kernel of T_2.
(B) The kernel of $T_1 + T_2$ is contained inside the intersection of the kernel of T_1 and the kernel of T_2, but need not be equal to the intersection.
(C) The kernel of $T_1 + T_2$ contains the intersection of the kernel of T_1 and the kernel of T_2, but need not be equal to the intersection.
(D) The kernel of $T_1 + T_2$ is contained inside the sum of the kernel of T_1 and the kernel of T_2, but need not be equal to the sum.
(E) The kernel of $T_1 + T_2$ contains the sum of the kernel of T_1 and the kernel of T_2, but need not be equal to the sum.

Your answer: ____________________________

(15) Suppose $T_1, T_2 : U \to V$ are linear transformations. Which of the following best describes the relation between the images of T_1, T_2, and $T_1 + T_2$?
(A) The image of $T_1 + T_2$ equals the intersection of the image of T_1 and the image of T_2.
(B) The image of $T_1 + T_2$ is contained inside the intersection of the image of T_1 and the image of T_2, but need not be equal to the intersection.
(C) The image of $T_1 + T_2$ contains the intersection of the image of T_1 and the image of T_2, but need not be equal to the intersection.
(D) The image of $T_1 + T_2$ is contained inside the sum of the image of T_1 and the image of T_2, but need not be equal to the sum.
(E) The image of $T_1 + T_2$ contains the sum of the image of T_1 and the image of T_2, but need not be equal to the sum.

Your answer: ____________________________

(16) Suppose T is a linear transformation from a vector space V to itself. Note that V may be an infinite-dimensional space, such as $C^\infty(\mathbb{R})$ (with T being differentiation), but for convenience, you can imagine V to be finite-dimensional (we will not reference the dimension of V in this question, however). Suppose the kernel of T has dimension n. What can you say from this information about the dimension of the kernel of T^r for a positive integer r?
(A) It is at least n and at most $n + r$.
(B) It is at least n and at most nr.
(C) It is at least $n + r$ and at most nr.
(D) It is at least $n + r$ and at most n^r.

Your answer: ____________________________

The next few questions deal with the relationship between the rows and columns of the matrix on the one hand, and the image and kernel of the linear transformation on the other hand.

(17) Suppose A is a $n \times m$ matrix and $T_A : \mathbb{R}^m \to \mathbb{R}^n$ is the corresponding linear transformation. Which of the following correctly describes the relationship between the rows and columns of A and the image and kernel of T_A?
(A) The kernel of T_A is precisely the subspace of $\mathbb{R}^m$ spanned by the rows of A. The image of T_A is precisely the subspace of $\mathbb{R}^n$ spanned by the columns of A.
(B) The kernel of T_A is precisely the subspace of $\mathbb{R}^m$ spanned by the columns of A. The image of T_A is precisely the subspace of $\mathbb{R}^n$ spanned by the rows of A.
(C) The kernel of T_A is precisely the subspace of $\mathbb{R}^m$ comprising the vectors that are *orthogonal* to the rows of A. The image of T_A is precisely the subspace of $\mathbb{R}^n$ comprising the vectors that are *orthogonal* to the columns of A.
(D) The kernel of T_A is precisely the subspace of $\mathbb{R}^m$ comprising the vectors that are *orthogonal* to the rows of A. The image of T_A is the subspace of $\mathbb{R}^n$ spanned by the columns of A.
(E) The kernel of T_A is precisely the subspace of $\mathbb{R}^m$ spanned by the rows of A. The image of T_A is precisely the subspace of $\mathbb{R}^n$ comprising the vectors that are *orthogonal* to the columns of A.

Your answer: ____________________________

(18) Suppose A and B are $n \times m$ matrices, $T_A : \mathbb{R}^m \to \mathbb{R}^n$ is the linear transformation corresponding to A, and $T_B : \mathbb{R}^m \to \mathbb{R}^n$ is the linear transformation corresponding to B. Which of the following correctly describes the relation between the rows, columns, image and kernel? Please see Option (E) before answering.
(A) If B can be obtained from A by a sequence of row interchange operations, then T_A and T_B have the same kernel as each other and also the same image as each other.
(B) If B can be obtained from A by a sequence of column interchange operations, then T_A and T_B have the same kernel as each other and also the same image as each other.

(C) If B can be obtained from A by a sequence of row interchange operations, then T_A and T_B have the same kernel as each other. If B can be obtained from A by a sequence of column interchange operations, then T_A and T_B have the same image as each other.

(D) If B can be obtained from A by a sequence of row interchange operations, then T_A and T_B have the same image as each other. If B can be obtained from A by a sequence of column interchange operations, then T_A and T_B have the same kernel as each other.

(E) None of the above.

Your answer: ____________________________

DIAGNOSTIC IN-CLASS QUIZ: ORIGINALLY DUE FRIDAY NOVEMBER 15, DELAYED TO WEDNESDAY NOVEMBER 20: LINEAR DEPENDENCE, BASES, AND SUBSPACES

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____

PLEASE DO NOT DISCUSS *ANY* QUESTIONS

The purpose of this quiz is to review some basic ideas from part of the lecture notes titled `Linear dependence, bases, and subspaces`. The corresponding sections of the book are Sections 3.2 and 3.3.

(1) *Do not discuss this!*: Suppose S is a finite nonempty set of vectors in $\mathbb{R}^n$, and T is a nonempty subset of S. What can we say about S and T?
  (A) S is linearly dependent if and only if T is linearly dependent. S is linearly independent if and only if T is linearly independent.
  (B) If S is linearly dependent, then T is linearly dependent. If S is linearly independent, then T is linearly independent. However, we cannot deduce anything about the linear dependence or independence of S from the linear dependence or independence of T.
  (C) If T is linearly dependent, then S is linearly dependent. If T is linearly independent, then S is linearly independent. However, we cannot deduce anything about the linear dependence or independence of T from the linear dependence or independence of S.
  (D) If S is linearly dependent, then T is linearly dependent. If T is linearly independent, then S is linearly independent. We cannot make either of the two other deductions.
  (E) If T is linearly dependent, then S is linearly dependent. If S is linearly independent, then T is linearly independent. We cannot make either of the other two deductions.

  Your answer: ____________________________

(2) *Do not discuss this!*: Suppose S is a finite set of vectors in $\mathbb{R}^n$. Consider the three statements: (i) S is linearly independent, (ii) S does not contain the zero vector, (iii) S does not contain any two vectors that are scalar multiples of one another. Which of the following options best describes the relationship between these statements?
  (A) (i) is equivalent to (ii), and both imply (iii), but the reverse implication does not hold.
  (B) (i) is equivalent to (iii), and both imply (ii), but the reverse implication does not hold.
  (C) (i) is equivalent to (ii) and (iii) combined.
  (D) (i) implies both (ii) and (iii), but (ii) and (iii), even if combined, do not imply (i).

  Your answer: ____________________________

(3) *Do not discuss this!*: Suppose V is a linear subspace of $\mathbb{R}^n$ for some n, and W is a linear subspace of V. Assume also that $W \neq V$, i.e., W is a *proper* subspace of V. Which of the following correctly describes the relationship between bases of V and bases of W?
  (A) Given a basis of V, we can find a subset of that basis that is a basis of W. Also, given a basis of W, we can find a set containing that basis that is a basis of V.
  (B) Given a basis of V, we can find a subset of that basis that is a basis of W. However, given a basis of W, we may not necessarily be able to find a set containing that basis that is a basis of V.
  (C) Given a basis of V, we may not necessarily be able to find a subset of that basis that is a basis of W. However, given a basis of W, we can find a set containing that basis that is a basis of V.

  Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE WEDNESDAY NOVEMBER 20: IMAGE AND KERNEL: APPLICATIONS TO CALCULUS

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____________________________________

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

The goal of this quiz is to use the setting of calculus to practice our skill of understanding linear transformations, specifically their injectivity, surjectivity, bijectivity, kernel and image. It builds on the November 8 quiz, but goes further. Please refer back to the November 8 quiz for the definitions of vector space, subspace, and linear transformation.

Please read these questions *very* carefully. For the first few questions, the interpretation of the question in the language of calculus is provided. Please refer to that if the linear algebra-based description is unclear.

(1) Let $\mathbb{R}[x]$ denote the vector space of all polynomials in one variable with real coefficients, with the usual addition and scalar multiplication of polynomials. There is an obvious linear transformation from $\mathbb{R}[x]$ to $C^\infty(\mathbb{R})$ that sends any polynomial to the function it describes, e.g., the polynomial $x^2 + 1$ gets sent to the function $x \mapsto x^2 + 1$. What can you say about this map $\mathbb{R}[x] \to C^\infty(\mathbb{R})$?

Please note: We are *not* talking here about whether the polynomial functions themselves are injective or surjective as functions from $\mathbb{R}$ to $\mathbb{R}$. Rather, we are talking about whether the mapping from *the set of polynomials* (which itself is a vector space over the reals) to *the set of infinitely differentiable functions* (which itself is another vector space).

(A) The map is neither injective nor surjective, i.e., different polynomials may define the same function, and not every infinitely differentiable function can be expressed using a polynomial.
(B) The map is injective but not surjective, i.e., different polynomials always define different functions, and not every infinitely differentiable function can be expressed using a polynomial.
(C) The map is surjective but not injective, i.e., different polynomials may define the same function, and every infinitely differentiable function can be expressed using a polynomial.
(D) The map is bijective, i.e., different polynomials always define different functions, and every infinitely differentiable function can be expressed using a polynomial.

Your answer: ________________________

(2) Denote by $\mathbb{R}[[x]]$ the vector space of all *formal power series* in one variable with real coefficients, with coefficient-wise addition and scalar multiplication. Explicitly, an element $a \in \mathbb{R}[[x]]$ is of the form:

$$a = \sum_{i=0}^{\infty} a_i x^i = a_0 + a_1 x + a_2 x^2 + \dots$$

where $a_i \in \mathbb{R}$ for $i \in \mathbb{N}_0$. Addition is coefficient-wise, i.e., if:

$$a = \sum_{i=0}^{\infty} a_i x^i, b = \sum_{i=0}^{\infty} b_i x^i$$

Then we have:

$$a + b = \sum_{i=0}^{\infty} (a_i + b_i) x^i$$

and for any real number λ, we have:

$$\lambda a = \sum_{i=0}^{\infty} (\lambda a_i) x^i$$

Note that a formal power series may have any radius of convergence. The radius of convergence could range from being 0 (which means that the formal power series converges only at the point $\{0\}$) to being ∞ (which means that the formal power series converges on all of $\mathbb{R}$). In other words, a formal power series need not define an actual function on $\mathbb{R}$.

Aside: If you remember sequences and series from single-variable calculus, you will recall that the radius of convergence is the reciprocal of the exponential growth rate of coefficients. In particular, if the coefficients *grow superexponentially*, the radius of convergence is zero. On the other hand, if the coefficients *decay superexponentially*, the radius of convergence is ∞. If the coefficients have exponential growth, the radius of convergence is less than 1. If the coefficients have exponential decay, the radius of convergence is greater than 1. Finally, if the coefficients grow or decay subexponentially, the radius of convergence is 1.

Note that $\mathbb{R}[x]$ can be viewed as a subspace of $\mathbb{R}[[x]]$ by thinking of each polynomial as a formal power series where there are only finitely many nonzero coefficients.

Let Ω be the subset of $\mathbb{R}[[x]]$ comprising those formal power series that converge globally, i.e., the radius of convergence is ∞. Note that Ω is a sub*space* of $\mathbb{R}[[x]]$.

What is the relation between $\mathbb{R}[x]$ and Ω?

Note that by *proper* subspace we mean a subspace that is not equal to the whole space.

(A) $\mathbb{R}[x] = \Omega$, i.e., a power series is globally convergent if and only if it is a polynomial (i.e., it has only finitely many nonzero coefficients).
(B) $\mathbb{R}[x]$ is a proper subspace of Ω, i.e., every polynomial is a globally convergent power series, but there exist globally convergent power series that are not polynomials.
(C) Ω is a proper subspace of $\mathbb{R}[x]$, i.e., every globally convergent power series is a polynomial, but there are polynomials that are not globally convergent power series.
(D) $\mathbb{R}[x]$ and Ω are incomparable, i.e., there exist polynomials that are not globally convergent power series and there exist globally convergent power series that are not polynomials.

Your answer: ____________________________

(3) The *Taylor series operator* can be viewed as a linear transformation from $C^\infty(\mathbb{R})$ to $\mathbb{R}[[x]]$. This operator sends any infinitely differentiable function to its Taylor series centered at 0. Explicitly, the operator is:

$$f \mapsto \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$$

What can we say about the kernel of this linear transformation?

(A) The kernel is the set of functions f satisfying $f(0) = 0$
(B) The kernel is the set of functions f satisfying $f'(0) = 0$
(C) The kernel is the set of functions f such that f and all its derivatives take the value 0 at 0.
(D) The kernel is the set of polynomial functions.
(E) The kernel is the set of functions that have globally convergent power series.

Your answer: ____________________________

(4) Which of the following is the best explanation for why we put the $+C$ when performing indefinite integration?

(A) The kernel of differentiation is a zero-dimensional space (namely, the zero function only), hence the fibers (inverse images or pre-images) for differentiation are all zero-dimensional spaces, i.e., single functions.
(B) The kernel of differentiation is a one-dimensional space (namely, the vector space of constant functions), hence the fibers (inverse images or pre-images) for differentiation are all one-dimensional spaces, i.e., lines that are translates of the space of constant functions.

(C) The image of differentiation is a zero-dimensional space (namely, the zero function only), hence the fibers (inverse images or pre-images) for differentiation are all zero-dimensional spaces, i.e., single functions.
(D) The image of differentiation is a one-dimensional space (namely, the vector space of constant functions), hence the fibers (inverse images or pre-images) for differentiation are all one-dimensional spaces, i.e., lines that are translates of the space of constant functions.

Your answer: ____________________________

(5) When finding all functions f on $\mathbb{R}$ such that $f''(x) = g(x)$ for some known continuous function g on $\mathbb{R}$, we get a general description of the form $G(x) + C_1x + C_2$ where C_1, C_2, are arbitrary real numbers. Which of the following is the best explanation for this?
(A) The kernel of the operation of differentiating twice is precisely the set of constant functions.
(B) The kernel of the operation of differentiating twice is precisely the set of nonconstant linear functions.
(C) The kernel of the operation of differentiating twice is the union of the set of constant functions and the set of nonconstant linear functions.
(D) The image of the operation of differentiating twice is precisely the set of constant functions.
(E) The image of the operation of differentiating twice is precisely the set of nonconstant linear functions.

Your answer: ____________________________

(6) Consider a second-order homogeneous linear differential equation of the form:

$$y'' + p_1(x)y' + p_2(x)y = 0$$

where x is the independent variable, y is the dependent variable, and p_1 and p_2 are known functions. We are trying to find global solutions, i.e., functions defined on all of $\mathbb{R}$. One way of thinking of this is to consider the linear transformation L that sends a function y of x to $L(y) = y'' + p_1(x)y' + p_2(x)y$, a new function of x. Which of the following best describes what we are trying to do?
(A) L is a linear transformation $C^2(\mathbb{R}) \to C(\mathbb{R})$, and the solution space we are interested in is the kernel of L.
(B) L is a linear transformation $C^2(\mathbb{R}) \to C(\mathbb{R})$, and the solution space we are interested in is the image of L.
(C) L is a linear transformation $C(\mathbb{R}) \to C^2(\mathbb{R})$, and the solution space we are interested in is the kernel of L.
(D) L is a linear transformation $C(\mathbb{R}) \to C^2(\mathbb{R})$, and the solution space we are interested in is the image of L.

Your answer: ____________________________

(7) Consider a second-order non-homogeneous linear differential equation of the form:

$$y'' + p_1(x)y' + p_2(x)y = q(x)$$

where x is the independent variable, y is the dependent variable, and p_1, p_2, and q are known functions. We are trying to find global solutions, i.e., functions defined on all of $\mathbb{R}$. One way of thinking of this is to consider the linear transformation L that sends a function y of x to $L(y) = y'' + p_1(x)y' + p_2(x)y$, a new function of x. Which of the following best describes what we are trying to do?
(A) We are trying to find the inverse image under L of $q(x)$, and we know this is a translate of the solution space of the corresponding homogeneous linear differential equation (the one from the preceding question).

(B) We are trying to find the image under L of $p_1(x)$, and we know this is a translate of the solution space of the corresponding homogeneous linear differential equation (the one from the preceding question).

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE FRIDAY NOVEMBER 22: LINEAR DYNAMICAL SYSTEMS

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ______________________

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

This quiz covers a topic that we will not be able to get to formally in the course due to time constraints. The corresponding section of the book is Section 7.1, and there is more relevant material discussed in the later sections of Chapter 7. However, you do not need to read those sections in order to attempt this quiz. Also, simply mastering the computational techniques in those sections of the book will not help you much with the quiz questions.

The questions here consider a linear dynamical system. Consider a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$. Let A be the matrix of T, so that A is a $n \times n$ matrix. For any positive integer r, the matrix A^r is the matrix for the linear transformation T^r (note here that T^r refers to the r-fold *composite* of T). The goal is to determine, starting off with an arbitrary vector $\vec{x} \in \mathbb{R}^n$, how the following sequence behaves:

$$\vec{x}, T(\vec{x}), T^2(\vec{x}), T^3(\vec{x}), \ldots$$

More explicitly, each term of the sequence is obtained by applying T to the preceding term. In other words, the sequence is:

$$\vec{x}, T(\vec{x}), T(T(\vec{x})), T(T(T(\vec{x}))), \ldots$$

(1) What is the necessary and sufficient condition on A such that for *every* choice of $\vec{x} \in \mathbb{R}^n$, the sequence described above eventually reaches, and stays at, the zero vector? Note that if it reaches the zero vector, it must do so in at most n steps. Please see Option (E) before answering.
(A) A is a nilpotent matrix.
(B) A is an idempotent matrix.
(C) A is an invertible matrix.
(D) A is a non-invertible matrix.
(E) None of the above.

Your answer: ______________________

(2) What is the necessary and sufficient condition on A such that there *exists* a nonzero vector $\vec{x} \in \mathbb{R}^n$ for which the sequence described above eventually reaches, and stays at, the zero vector? Note that if it reaches the zero vector, it must do so in at most n steps. Please see Option (E) before answering.
(A) A is a nilpotent matrix.
(B) A is an idempotent matrix.
(C) A is an invertible matrix.
(D) A is a non-invertible matrix.
(E) None of the above.

Your answer: ______________________

(3) What is the necessary and sufficient condition on A such that for *every* choice of $\vec{x} \in \mathbb{R}^n$, the sequence described above returns to $\vec{x}$ after a finite and positive number of steps? Please see Option (E) before answering.
(A) A is a nilpotent matrix.
(B) A is an idempotent matrix.
(C) A is an invertible matrix.

(D) A is a non-invertible matrix.
(E) None of the above.

Your answer: ____________________________

(4) What is the necessary and sufficient condition on A such that there *exists* a nonzero vector $\vec{x} \in \mathbb{R}^n$ for which the sequence described above returns to $\vec{x}$ after a finite and positive number of steps? Please see Option (E) before answering.
(A) A is a nilpotent matrix.
(B) A is an idempotent matrix.
(C) A is an invertible matrix.
(D) A is a non-invertible matrix.
(E) None of the above.

Your answer: ____________________________

(5) Suppose $n = 2$ and T is a rotation by an angle that is a rational multiple of π. What can we say about the range of the sequence

$$\vec{x}, T(\vec{x}), T^2(\vec{x}), T^3(\vec{x}), \dots$$

starting from a nonzero vector $\vec{x}$?
(A) The range is finite, i.e., there are only finitely many distinct vectors in the sequence.
(B) The range is infinite and forms a dense subset of the circle centered at the origin and with radius equal to the length of the vector $\vec{x}$. However, it is not the entire circle.
(C) The range is infinite and is the entire circle centered at the origin and with radius equal to the length of the vector $\vec{x}$.
(D) The range is infinite and forms a dense subset of the line of the vector $\vec{x}$ (excluding the origin), but is not the entire line (excluding the origin).
(E) The range is infinite and is the entire line of the vector $\vec{x}$, excluding the origin.

Your answer: ____________________________

(6) Suppose $n = 2$ and T is a rotation by an angle that is a irrational multiple of π. What can we say about the range of the sequence

$$\vec{x}, T(\vec{x}), T^2(\vec{x}), T^3(\vec{x}), \dots$$

starting from a nonzero vector $\vec{x}$?
(A) The range is finite, i.e., there are only finitely many distinct vectors in the sequence.
(B) The range is infinite and forms a dense subset of the circle centered at the origin and with radius equal to the length of the vector $\vec{x}$. However, it is not the entire circle.
(C) The range is infinite and is the entire circle centered at the origin and with radius equal to the length of the vector $\vec{x}$.
(D) The range is infinite and forms a dense subset of the line of the vector $\vec{x}$ (excluding the origin), but is not the entire line (excluding the origin).
(E) The range is infinite and is the entire line of the vector $\vec{x}$, excluding the origin.

Your answer: ____________________________

We return to generic n now.

(7) A nonzero vector $\vec{x}$ is termed an *eigenvector* for a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ with *eigenvalue* a real number $\lambda \in \mathbb{R}$ if $T(\vec{x}) = \lambda\vec{x}$. Note that λ is allowed to be 0. We sometimes conflate the roles of T and its matrix A, so that we call $\vec{x}$ an eigenvector for A and λ an eigenvalue for A.

If $\vec{x}$ is an eigenvector of T (or equivalently, of A) with eigenvalue λ, which of the following is true? We denote by I_n the identity transformation from $\mathbb{R}^n$ to itself.
(A) $\vec{x}$ must be in the kernel of the linear transformation $T + \lambda I_n$
(B) $\vec{x}$ must be in the image of the linear transformation $T + \lambda I_n$
(C) $\vec{x}$ must be in the kernel of the linear transformation $T - \lambda I_n$
(D) $\vec{x}$ must be in the image of the linear transformation $T - \lambda I_n$
(E) $\vec{x}$ must be in the kernel of the linear transformation λT

Your answer: ____________________________

(8) As above, let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation with matrix A. Use the terminology of eigenvector and eigenvalue from the preceding question. Which of the following is a characterization of the situation that A is a diagonal matrix?
(A) Every nonzero vector in $\mathbb{R}^n$ is an eigenvector for T.
(B) Every standard basis vector in $\mathbb{R}^n$ is an eigenvector for T.
(C) Every vector with at least one zero coordinate in $\mathbb{R}^n$ is an eigenvector for T.
(D) T has a unique eigenvector (up to scalar multiples, i.e., all eigenvectors of T are scalar multiples of each other).
(E) T has no eigenvector.

Your answer: ____________________________

(9) As above, let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation with matrix A. Use the terminology of eigenvector and eigenvalue from the preceding question. Which of the following is a characterization of the situation that A is a scalar matrix (i.e., a diagonal matrix with all diagonal entries equal)?
(A) Every nonzero vector in $\mathbb{R}^n$ is an eigenvector for T.
(B) Every standard basis vector in $\mathbb{R}^n$ is an eigenvector for T.
(C) Every vector with at least one zero coordinate in $\mathbb{R}^n$ is an eigenvector for T.
(D) T has a unique eigenvector (up to scalar multiples, i.e., all eigenvectors of T are scalar multiples of each other).
(E) T has no eigenvector.

Your answer: ____________________________

(10) Suppose A is a strictly upper-triangular $n \times n$ matrix, i.e., all entries of A that are on or below the main diagonal are zero. T is the linear transformation corresponding to A. It will turn out that the only eigenvalue for T is 0. What can we say about the eigenvectors for T for this eigenvalue?
(A) All nonzero vectors in $\mathbb{R}^n$ are eigenvectors for T with eigenvalue 0.
(B) All standard basis vectors in $\mathbb{R}^n$ are eigenvectors for T with eigenvalue 0.
(C) The vector $\vec{e}_1$ is an eigenvector for T with eigenvalue 0. The information presented is not sufficient to determine whether any of the other standard basis vectors is an eigenvector.
(D) The vector $\vec{e}_n$ is an eigenvector for T with eigenvalue 0. The information presented is not sufficient to determine whether any of the other standard basis vectors is an eigenvector.
(E) At least one of the standard basis vectors is an eigenvector for T with eigenvalue 0. However, the information presented is not sufficient to say definitively for any particular standard basis vector that it is an eigenvector.

Your answer: ____________________________

(11) Suppose A is a strictly upper-triangular $n \times n$ matrix, i.e., all entries of A that are on or below the main diagonal are zero. T is the linear transformation corresponding to A. Which of the following is A guaranteed to be? Please see Options (D) and (E) before answering.
(A) Nilpotent
(B) Idempotent
(C) Invertible
(D) All of the above

(E) None of the above

Your answer: ____________________________

(12) Consider the case $n = 2$ and let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a rotation by an angle that is *not* an *integer* multiple of π. What can we say about the set of eigenvectors and eigenvalues for T?

(A) T has no eigenvectors

(B) T has one eigenvector (up to scalar multiples) with eigenvalue 1

(C) T has one eigenvector (up to scalar multiples) and the eigenvalue depends on the angle of rotation

(D) T has two linearly independent eigenvectors (so that the set of all eigenvectors is obtained as the set of scalar multiples of either one of these vectors) with the same eigenvalue

(E) T has two linearly independent eigenvectors (so that the set of all eigenvectors is obtained as the set of scalar multiples of either one of these vectors) with distinct eigenvalues

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE MONDAY NOVEMBER 25: STOCHASTIC MATRICES

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

This quiz can be viewed as a continuation of the quiz on linear dynamical systems. The book defines column-stochastic matrices using the jargon "transition matrix" on Page 53 (Definition 2.1.4) and uses them throughout the text when describing (a simplified version of) Google's PageRank algorithm. The quiz questions are self-contained and do not require you to read the book, but you may benefit from skimming through the book's discussion of PageRank to complement these questions. Note that the 4th Edition does not include the discussion of transition matrices and PageRank.

In this quiz, we discuss the dynamics of a very special type of linear transformation. A $n \times n$ matrix A is termed a *row-stochastic matrix* if all its entries are in the interval $[0, 1]$ and all the row sums are equal to 1. A $n \times n$ matrix is termed a *column-stochastic matrix* if all its entries are in the interval $[0, 1]$ and all the column sums are equal to 1. A $n \times n$ matrix A is termed a *doubly stochastic matrix* if it is both row-stochastic and column-stochastic, i.e., all the entries are in the interval $[0, 1]$, all the row sums are equal to 1, and all the column sums are equal to 1.

(1) Suppose A and B are two $n \times n$ row-stochastic matrices. Which of the following is *guaranteed* to be row-stochastic? Please see Options (D) and (E) before answering.
   (A) $A + B$
   (B) $A - B$
   (C) AB
   (D) All of the above
   (E) None of the above

   Your answer: ____________________________

(2) Suppose A and B are two $n \times n$ column-stochastic matrices. Which of the following is *guaranteed* to be column-stochastic? Please see Options (D) and (E) before answering.
   (A) $A + B$
   (B) $A - B$
   (C) AB
   (D) All of the above
   (E) None of the above

   Your answer: ____________________________

(3) Suppose A and B are two $n \times n$ doubly stochastic matrices. Which of the following is *guaranteed* to be doubly stochastic? Please see Options (D) and (E) before answering.
   (A) $A + B$
   (B) $A - B$
   (C) AB
   (D) All of the above
   (E) None of the above

   Your answer: ____________________________

We now consider the case $n = 2$. In this case, the doubly stochastic matrices have the form:

$$\begin{bmatrix} a & 1-a \\ 1-a & a \end{bmatrix}$$

where $a \in [0,1]$. Denote this matrix by $D(a)$ for short.

(4) Suppose $a, b \in [0,1]$ (they are allowed to be equal). The product $D(a)D(b)$ equals $D(c)$ for some $c \in [0,1]$. What is that value of c?
(A) $a + b$
(B) ab
(C) $2ab + a + b$
(D) $(1-a)(1-b)$
(E) $1 - a - b + 2ab$

Your answer: ______________________________

(5) For what value(s) of a is the matrix $D(a)$ non-invertible? Note that when judging invertibility, we do not insist that the inverse matrix also be doubly stochastic.
(A) $a = 0$ only
(B) $a = 1/2$ only
(C) $a = 1$ only
(D) $0 < a < 1$ (i.e., $D(a)$ is invertible only at $a = 0$ and $a = 1$)
(E) $a \neq 1/2$

Your answer: ______________________________

(6) For what value(s) of a is it true that the matrix $D(a)$ does not have an inverse that is a doubly stochastic matrix? In other words, either $D(a)$ should be non-invertible or it should be invertible but the inverse is not a doubly stochastic matrix.
(A) $a = 0$ only
(B) $a = 1/2$ only
(C) $a = 1$ only
(D) $0 < a < 1$ (i.e., $D(a)$ has an inverse that is also doubly stochastic only if $a = 0$ or $a = 1$)
(E) $a \neq 1/2$

Your answer: ______________________________

For the next few questions, denote by T_a the linear transformation whose matrix is $D(a)$. For any vector $\vec{x} \in \mathbb{R}^2$, we can consider the sequence:

$$\vec{x}, T_a(\vec{x}), T_a^2(\vec{x}), \dots$$

Note that if we were to start with a vector $\vec{x} \in \mathbb{R}^2$ with both coordinates equal, it would be invariant under T_a.

Thus, for the questions below, assume that we start with a nonzero vector $\vec{x} \in \mathbb{R}^2$ for which the two coordinates are not equal to each other.

(7) For what value of a is it the case that $\lim_{r \to \infty} T_a^r(\vec{x})$ does *not* exist?
(A) $a = 0$ only
(B) $a = 1/2$ only
(C) $a = 1$ only
(D) $0 < a < 1$
(E) $a \neq 1/2$

Your answer: ______________________________

(8) For what value of a is it the case that the sequence

$$\vec{x}, T_a(\vec{x}), T_a^2(\vec{x}), \dots$$

is a constant sequence?

(A) $a = 0$ only
(B) $a = 1/2$ only
(C) $a = 1$ only
(D) $0 < a < 1$
(E) $a \neq 1/2$

Your answer: ____________________________

(9) For what value of a is it the case that the sequence

$$\vec{x}, T_a(\vec{x}), T_a^2(\vec{x}), \dots$$

is not a constant sequence but becomes constant from $T_a(\vec{x})$ onward?

(A) $a = 0$ only
(B) $a = 1/2$ only
(C) $a = 1$ only
(D) $0 < a < 1$
(E) $a \neq 1/2$

Your answer: ____________________________

(10) For a other than 0, 1/2, or 1, what is the limit $\lim_{r \to \infty}(D(a))^r$? Here, when we talk of taking the limit of a sequence of matrices, we are taking the limit entry-wise.

(A) The matrix $D(0)$
(B) The matrix $D(1/2)$
(C) The matrix $D(1)$
(D) The matrix $D(a)$
(E) The matrix $D(1 - a)$

Your answer: ____________________________

DIAGNOSTIC IN-CLASS QUIZ: DUE MONDAY NOVEMBER 25: SUBSPACE, BASIS, AND DIMENSION

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ________________________________

PLEASE DO NOT DISCUSS *ANY* QUESTIONS.

This quiz covers material related to the `Linear dependence, bases and subspaces` notes corresponding to Sections 3.2 and 3.3 of the text.

Keep in mind the following facts. Suppose $T : \mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation. Suppose A is the matrix for T, so that $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^m$. Then, A is a $n \times m$ matrix. Further, the following are true:

- The dimension of the image of T equals the rank of A.
- The dimension of the kernel of T, called the *nullity* of A, is m minus the rank of A.

(1) *Do not discuss this!*: Suppose $T : \mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation. What is the best we can say about the dimension of the image of T?
(A) It is at least 0 and at most $\min\{m, n\}$. However, we cannot be more specific based on the given information.
(B) It is at least 0 and at most $\max\{m, n\}$. However, we cannot be more specific based on the given information.
(C) It is at least $\min\{m, n\}$ and at most $\max\{m, n\}$. However, we cannot be more specific based on the given information.
(D) It is at least $\min\{m, n\}$ and at most $m + n$. However, we cannot be more specific based on the given information.
(E) It is at least $\max\{m, n\}$ and at most $m + n$. However, we cannot be more specific based on the given information.

Your answer: ____________________

(2) *Do not discuss this!*: Suppose $T_1, T_2 : \mathbb{R}^m \to \mathbb{R}^n$ are linear transformations. Suppose the images of T_1 and T_2 have dimensions d_1 and d_2 respectively. What can we say about the dimension of the image of $T_1 + T_2$? Assume that both m and n are larger than $d_1 + d_2$.
(A) It is precisely $|d_2 - d_1|$.
(B) It is precisely $\min\{d_1, d_2\}$.
(C) It is precisely $\max\{d_1, d_2\}$.
(D) It is precisely $d_1 + d_2$.
(E) Based on the information, it could be any integer r with $|d_2 - d_1| \leq r \leq d_1 + d_2$.

Your answer: ____________________

(3) *Do not discuss this!*: Suppose V_1 and V_2 are subspaces of $\mathbb{R}^n$. We define the sum $V_1 + V_2$ as the subset of $\mathbb{R}^n$ comprising all vectors that can be expressed as a sum of a vector in V_1 and a vector in V_2. Define $V_1 \cup V_2$ as the set-theoretic union of V_1 and V_2, i.e., the set of all vectors that are either in V_1 or in V_2. What can we say about these?
(A) $V_1 \cup V_2 = V_1 + V_2$ and it is a subspace of $\mathbb{R}^n$.
(B) $V_1 \cup V_2$ is contained in $V_1 + V_2$ and both are subspaces of $\mathbb{R}^n$.
(C) $V_1 \cup V_2$ is contained in $V_1 + V_2$, and $V_1 + V_2$ is a subspace of $\mathbb{R}^n$. $V_1 \cup V_2$ is generally not a subspace of $\mathbb{R}^n$ (though it might be in special cases).
(D) $V_1 \cup V_2$ contains $V_1 + V_2$, and both are subspaces of $\mathbb{R}^n$.
(E) $V_1 \cup V_2$ contains $V_1 + V_2$, and $V_1 \cup V_2$ is a subspace of $\mathbb{R}^n$. $V_1 + V_2$ is generally not a subspace of $\mathbb{R}^n$ (though it might be in special cases).

Your answer: ____________________

TAKE-HOME CLASS QUIZ SOLUTIONS: DUE MONDAY NOVEMBER 25: SUBSPACE, BASIS, DIMENSION, AND ABSTRACT SPACES: APPLICATIONS TO CALCULUS

MATH 196, SECTION 57 (VIPUL NAIK)

1. Performance review

22 people took this 13-question quiz. The score distribution was as follows:

- Score of 2: 1 person
- Score of 4: 3 people
- Score of 5: 2 people
- Score of 6: 2 people
- Score of 7: 1 person
- Score of 8: 5 people
- Score of 9: 3 people
- Score of 10: 2 people
- Score of 11: 3 people

The mean score was about 7.41.

The question-wise answers and performance review are as follows:

(1) Option (A): 18 people
(2) Option (A): 15 people
(3) Option (E): 17 people
(4) Option (D): 13 people
(5) Option (C): 15 people
(6) Option (E): 10 people
(7) Option (E): 12 people
(8) Option (E): 13 people
(9) Option (E): 9 people
(10) Option (D): 1 person
(11) Option (C): 15 people
(12) Option (D): 13 people
(13) Option (B): 12 people

REVIEW NOTE: Please make sure to read the corresponding lecture notes on abstract vector spaces rather than simply going over the quiz solutions.

2. Solutions

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

This quiz builds on the November 8 and November 20 quizzes that apply ideas we are learning about linear transformations to the calculus setting. The November 8 quiz went over some basic ideas related to differentiation as a linear transformation. The November 20 quiz explored the ideas in greater depth. We now look at questions that apply the ideas of basis, dimension, and subspace to the calculus setting.

We begin by recalling some notation and facts we already saw in earlier quizzes. Denote by $C(\mathbb{R})$ (or alternatively by $C^0(\mathbb{R})$) the vector space of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$, with pointwise addition and scalar multiplication. Note that the elements of this vector space, which we would ordinarily call "vectors", are now *functions*.

For k a positive integer, denote by $C^k(\mathbb{R})$ the subspace of $C(\mathbb{R})$ comprising those continuous functions that are at least k times *continuously* differentiable. Note that $C^{k+1}(\mathbb{R})$ is a subspace of $C^k(\mathbb{R})$, so we have a descending chain of subspaces:

$$C(\mathbb{R}) = C^0(\mathbb{R}) \supseteq C^1(\mathbb{R}) \supseteq C^2(\mathbb{R}) \supseteq \dots$$

The intersection of these spaces is the vector space $C^\infty(\mathbb{R})$, defined as the subspace of $C(\mathbb{R})$ comprising those functions that are *infinitely* differentiable.

We had also noted that:

- The kernel of differentiation is the vector space of constant functions.
- The kernel of k times differentiating is the vector space of polynomials of degree at most $k-1$.
- The fiber of any function for differentiation is a translate of the space of constant functions. That's what explains the $+C$ when you perform indefinite integration.

Note: For finite-dimensional spaces, a linear transformation T from a vector space to itself is injective if and only if it is surjective. This follows from dimension and rank considerations: T is injective if and only its kernel is zero, which happens if and only if the matrix has full column rank, which happens if and only if the matrix has full row rank (because the matrix is a square matrix), which happens if and only if T is surjective. The rank-nullity theorem provides an equivalent explanation. We had also seen that if $T : \mathbb{R}^m \to \mathbb{R}^n$ is injective, then $m \leq n$, and if $T : \mathbb{R}^m \to \mathbb{R}^n$ is surjective, then $m \geq n$. In particular, we cannot have a surjective map from a proper subspace to the whole space.

With infinite-dimensional spaces, however, we can have funny phenomena. Examples of these phenomena are strewn across the quizzes.

- We can have a map from an infinite-dimensional vector space to itself that is injective but not surjective.
- We can have a map from an infinite-dimensional vector space to itself that is surjective but not injective.
- We can have a surjective map from a proper subspace to the whole space (for instance, differentiation $C^1(\mathbb{R}) \to C(\mathbb{R})$ is surjective, even though $C^1(\mathbb{R})$ is a proper subspace of $C(\mathbb{R})$).
- We can have an injective map from a space to a proper subspace.

Note that we will use the terms *subspace* and *vector subspace* synonymously with *linear subspace* in this quiz.

(1) Suppose V is a vector subspace of the vector space $C^\infty(\mathbb{R})$. We know that differentiation is linear. How is that information computationally useful?

(A) It tells us that knowing how to differentiate all functions in any spanning set for V tells us how to differentiate any function in V (assuming we know how to express any function in V as a linear combination of the functions in the spanning set).

(B) It tells us that knowing how to differentiate all functions in any linearly independent set in V tells us how to differentiate any function in V.

Answer: Option (A)

Explanation: **MAKE SURE TO READ THE LECTURE NOTES ON ABSTRACT VECTOR SPACES, SECTION 5**.

Given the knowledge of the derivatives of all functions in a spanning set for V, we can differentiate any function in V as follows: first, express it as a linear combination of the functions in the spanning set. Now, use the linearity of differentiation to express its derivative as the corresponding linear combination of the derivatives.

For instance, suppose we know that the derivative of sin is cos and the derivative of exp is exp. Then the derivative of the function:

$$f(x) = 2\sin x + 5\exp(x)$$

is:

$$f'(x) = 2\sin' x + 5\exp'(x) = 2\cos x + 5\exp(x)$$

Note that it is the fact of the functions *spanning* V that is crucial in allowing us to be able to write *any* function in V as a linear combination of the functions.

Performance review: 18 out of 22 people got this. 4 chose (B).

Historical note (last time): 20 out of 26 got this. 6 chose (B).

(2) Suppose V is a vector subspace of the vector space $C^\infty(\mathbb{R})$. We know that differentiation is linear. How is that information computationally useful?

(A) It tells us that knowing the antiderivatives of all functions in any spanning set for V tells us the antiderivative of every function in V (assuming we know how to express any function in V as a linear combination of the functions in the spanning set).

(B) It tells us that knowing the antiderivatives of all functions in any linearly independent set in V tells us the antiderivative of every function in V.

Answer: Option (A)

Explanation: **MAKE SURE TO READ THE LECTURE NOTES ON ABSTRACT VECTOR SPACES, SECTION 5**.

The reasoning is similar to that for differentiation, except that we put in the obligatory $+C$ of indefinite integration to account for the fact that the kernel of differentiation is the one-dimensional space of constant functions.

For instance, suppose we know that an antiderivative of sin is $-\cos$ and an antiderivative of exp is exp. Then, the indefinite integral of the function:

$$f(x) = 2\sin x + 5\exp(x)$$

is:

$$\int f(x)\,dx = 2(-\cos x) + 5\exp x + C$$

Performance review: 15 out of 22 people got this. 7 chose (B).

Historical note (last time): 23 out of 26 got this. 3 chose (B).

We now consider two related vector spaces. $\mathbb{R}[x]$ is defined as the vector space of polynomials with real coefficients in the single variable x, with the usual addition and scalar multiplication. There is a natural injective homomorphism from $\mathbb{R}[x]$ to $C^\infty(\mathbb{R})$ that sends any polynomial to the same polynomial viewed as a function.

$\mathbb{R}(x)$ is defined as the vector space of all rational functions where the numerator and denominator are both polynomials with the denominator nonzero, up to equivalence (i.e., two rational functions $p_1(x)/q_1(x)$ and $p_2((x)/q_2(x)$ are equivalent if $p_1(x)q_2(x) = q_1(x)p_2(x)$). Addition and scalar multiplication are defined the usual way. Note that there is a natural injective homomorphism from $\mathbb{R}[x]$ to $\mathbb{R}(x)$ that sends any polynomial $p(x)$ to the rational function $p(x)/1$.

Also note that $\mathbb{R}(x)$ does not map to $C^\infty(\mathbb{R})$, for the reason that a rational function, viewed *qua* function, is not necessarily defined everywhere. Specifically, if written in simplified form, it is not defined at the set of roots of its denominator.

Note that both $\mathbb{R}[x]$ and $\mathbb{R}(x)$ are infinite-dimensional vector spaces, i.e., they do not have finite spanning sets.

(3) Which of the following is *not* a basis for $\mathbb{R}[x]$? Please see Option (E) before answering.

(A) $1, x, x^2, x^3, \dots$

(B) $1, x, x(x-1), x(x-1)(x-2), x(x-1)(x-2)(x-3), \dots$

(C) $1, x+1, x^2+x+1, x^3+x^2+x+1, \dots$

(D) $1, x, x^2-x, x^3-x^2, x^4-x^3, \dots$

(E) None of the above, i.e., each of them is a basis.

Answer: Option (E)

Explanation: **MAKE SURE TO READ THE LECTURE NOTES ON ABSTRACT VECTOR SPACES, SECTION 5**.

Option (A) clearly is a basis: polynomials are by definition linear combinations of $1, x, x^2, \dots$ and the method of expressing a polynomial as a linear combination is unique. All the other options are equivalent to Option (A) in the following sense: if we think of how the span grows as we go from left to right, it's the same in all options. In each option, the span of the first n vectors is the same as the span of $1, x, x^2, \dots, x^{n-1}$. In particular, in all options, there are no redundant vectors, and the span of all vectors together is all of $\mathbb{R}[x]$. In other words, each option gives a basis.

Performance review: 17 out of 22 people got this. 2 each chose (B) and (C), 1 chose (D).

Historical note (last time): 16 out of 26 got this. 6 chose (C), 2 chose (D), 1 each chose (A) and (B).

Let's now revisit the topic of *partial fractions* as a tool for integrating rational functions. The idea behind partial fractions is to consider an integration problem with respect to a variable x with integrand of the following form:

$$\frac{a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1}}{p(x)}$$

where p is a polynomial of degree n. For convenience, we may take p to be a monic polynomial, i.e., a polynomial with leading coefficient 1. For p fixed, the set of all rational functions of the form above forms a vector subspace of dimension n inside $\mathbb{R}(x)$. A natural choice of basis for this subspace is:

$$\frac{1}{p(x)}, \frac{x}{p(x)}, \dots, \frac{x^{n-1}}{p(x)}$$

The goal of partial fraction theory is to provide an *alternate basis* for this space of functions with the property that those basis elements are particularly easy to integrate (recurring to one of our earlier questions). Let's illustrate one special case: the case that p has n distinct real roots $\alpha_1, \alpha_2, \dots, \alpha_n$. The alternate basis in this case is:

$$\frac{1}{x - \alpha_1}, \frac{1}{x - \alpha_2}, \dots, \frac{1}{x - \alpha_n}$$

The explicit goal is to rewrite a partial fraction:

$$\frac{a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1}}{p(x)}$$

in terms of the basis above. If we denote the numerator as $r(x)$, we want to write:

$$\frac{r(x)}{p(x)} = \frac{c_1}{x - \alpha_1} + \frac{c_2}{x - \alpha_2} + \cdots + \frac{c_n}{x - \alpha_n}$$

The explicit formula is:

$$c_i = \frac{r(\alpha_i)}{\prod_{j \neq i}(\alpha_i - \alpha_j)}$$

Once we rewrite the original rational function as a linear combination of the new basis vectors, we can integrate it easily because we know the antiderivatives of each of the basis vectors. The antiderivative is thus:

$$\left(\sum_{i=1}^{n} \frac{r(\alpha_i)}{\prod_{j \neq i}(\alpha_i - \alpha_j)} \ln|x - \alpha_i|\right) + C$$

where the obligatory $+C$ is put for the usual reasons.

Note that this process only handles rational functions that are proper fractions, i.e., the degree of the numerator must be less than that of the denominator.

We now consider cases where p is a polynomial of a different type.

(4) Suppose p is a monic polynomial of degree n that is a product of pairwise distinct irreducible factors that are all either monic linear or monic quadratic. Call the roots for the linear polynomials $\alpha_1, \alpha_2, \dots, \alpha_s$ and call the monic quadratic factors $q_1, q_2, \dots, q_t$. Which of the following sets forms a basis for the vector space that we are interested in, namely all rational functions of the form $r(x)/p(x)$ where the degree of r is less than n? Please see Option (E) before answering.

(A) All rational functions of the form $1/(x - \alpha_i), 1 \leq i \leq s$ together with all rational functions of the form $1/q_j(x), 1 \leq j \leq t$

(B) All rational functions of the form $1/(x - \alpha_i), 1 \leq i \leq s$ together with all rational functions of the form $q_j'(x)/q_j(x), 1 \leq j \leq t$

(C) All rational functions of the form $1/q_j(x), 1 \leq j \leq t$ together with all rational functions of the form $q_j'(x)/q_j(x), 1 \leq j \leq t$

(D) All rational functions of the form $1/(x - \alpha_i), 1 \leq i \leq s$ together with all rational functions of the form $1/q_j(x), 1 \leq j \leq t$ *and* all rational functions of the form $q_j'(x)/q_j(x), 1 \leq j \leq t$

(E) None of the above

Answer: Option (D)

Explanation: **MAKE SURE TO READ THE LECTURE NOTES ON ABSTRACT VECTOR SPACES, SECTION 5**.

This should be familiar to you from the halcyon days of doing partial fractions. For instance, consider the example where $p(x) = (x - 1)(x^2 + x + 1)$. In this case, the basis is:

$$\frac{1}{x-1}, \frac{1}{x^2 + x + 1}, \frac{2x + 1}{x^2 + x + 1}$$

Note that an easy sanity check is that the *size* of the basis should be n. This is clear in the above example with $n = 3$, but let's reason generically.

We have that:

$$p(x) = \left[\prod_{i=1}^{s}(x - \alpha_i)\right]\left[\prod_{j=1}^{t} q_j(x)\right]$$

By degree considerations, we get that:

$$s + 2t = n$$

Now, the vector space for which we are trying to obtain a basis has dimension n. This means that the basis we are looking for should have size n. Of the given options, Option (D) (which gives one basis element for each of the s linear factors and two basis elements for each of the t quadratic factors) is the most attractive.

Also recall that the reciprocals of the linear factors integrate to logarithms. The expressions of the form $1/q_j(x)$ integrate to an expression involving arctan. The expressions of the form $q_j'(x)/q_j(x)$ integrate to logarithms.

Performance review: 13 out of 22 people got this. 3 each chose (A) and (B), 2 chose (E), 1 chose (C).

Historical note (last time): 4 out of 26 got this. 13 chose (A), 5 chose (B), and 4 chose (C).

(5) Suppose $p(x) = (x - \alpha)^n$. Which of the following sets forms a basis for the vector space that we are interested in, namely all rational functions of the form $r(x)/p(x)$ where the degree of r is less than n? Please see Options (D) and (E) before answering.

(A) The single function $1/(x - \alpha)$

(B) The single function $1/(x-\alpha)^n$
(C) All the functions $1/(x-\alpha), 1/(x-\alpha)^2, \ldots, 1/(x-\alpha)^n$
(D) Any of the above works
(E) None of the above works

Answer: Option (C)

Explanation: **MAKE SURE TO READ THE LECTURE NOTES ON ABSTRACT VECTOR SPACES, SECTION 5**.

This is the obvious choice by size considerations. It also makes sense based on what you remember about partial fractions.

Performance review: 15 out of 22 people got this. 4 chose (D), 2 chose (B), 1 chose (A).

Historical note (last time): 14 out of 26 got this. 8 chose (D), 3 chose (B), and 1 chose (E).

We now recall our earlier discussion of the solution process for first-order linear differential equations. Consider a first-order linear differential equation with independent variable x and dependent variable y, with the equation having the form:

$$y' + p(x)y = q(x)$$

where $p, q \in C^\infty(\mathbb{R})$.

We solve this equation as follows. Let H be an antiderivative of p, so that $H'(x) = p(x)$.

$$\frac{d}{dx}\left(ye^{H(x)}\right) = q(x)e^{H(x)}$$

This gives:

$$ye^{H(x)} = \int q(x)e^{H(x)}\,dx$$

So:

$$y = e^{-H(x)}\int q(x)e^{H(x)}\,dx$$

The indefinite integration gives a $+C$, so overall, we get:

$$y = Ce^{-H(x)} + \text{particular solution}$$

It's now time to understand this in terms of linear algebra.

Define a linear transformation $L : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ as:

$$f(x) \mapsto f'(x) + p(x)f(x)$$

(6) The kernel of L is one-dimensional. Which of the following functions spans the kernel?
(A) $p(x)$
(B) $q(x)$
(C) $H(x)$
(D) $e^{H(x)}$
(E) $e^{-H(x)}$

Answer: Option (E)

Explanation: **MAKE SURE TO READ THE LECTURE NOTES ON ABSTRACT VECTOR SPACES, SECTION 7**.

This is pretty obvious from the description of the general solution above. In particular, set $q(x) = 0$ and get the generic element of the kernel as:

$$y = Ce^{-H(x)}, C \in \mathbb{R}$$

This is spanned by $e^{-H(x)}$.

Performance review: 10 out of 22 people got this. 4 each chose (A) and (B), 2 each chose (C) and (D).

Historical note (last time): 17 out of 26 got thus. 5 chose (C), 2 each chose (A) and (B).

(7) I would like to argue that L is *surjective* as a linear transformation from $C^\infty(\mathbb{R})$ to $C^\infty(\mathbb{R})$. Why is that true?

(A) The kernel of L is zero-dimensional.
(B) The image of L is zero-dimensional.
(C) The kernel of L is one-dimensional.
(D) The image of L is one-dimensional.
(E) For any q, we have a formula above that describes a solution function that maps to q.

Answer: Option (E)

Explanation: **MAKE SURE TO READ THE LECTURE NOTES ON ABSTRACT VECTOR SPACES, SECTION 7**.

Obvious! Note that Option (C), while correct, is not an *explanation* for the surjectivity of the linear transformation, because the dimension of the kernel speaks to how far the linear transformation is from being injective, and, particularly in the infinite-dimensional case, does not provide any direct information regarding whether or not the linear transformation is surjective.

Performance review: 12 out of 22 people got this. 8 chose (C), 2 chose (A).

Historical note (last time): 2 out of 26 got this. 12 chose (C), 7 chose (D), 3 chose (A), 2 chose (B).

Let n be a nonnegative integer. Denote by P_n the vector space of all polynomials in one variable x that have degree $\leq n$. P_n is a subspace of $\mathbb{R}[x]$, which in turn can be viewed as a subspace of $C^\infty(\mathbb{R})$ through the natural injective map. For convenience and completeness, define P_{-1} to be the zero subspace.

Differentiation defines a linear transformation from $C^\infty(\mathbb{R})$ to itself.

(8) What are the kernel and image of the restriction of differentiation to P_n? The result should be valid for all positive integers n.

(A) The kernel and image are both P_n
(B) The kernel is the zero subspace and the image is P_n
(C) The kernel is P_n and the image is the zero subspace
(D) The kernel is P_{n-1} and the image is P_0 (the subspace of constant functions)
(E) The kernel is P_0 and the image is P_{n-1}

Answer: Option (E)

Explanation: **MAKE SURE TO READ THE LECTURE NOTES ON ABSTRACT VECTOR SPACES, SECTION 5**.

The only functions that differentiate to 0 are the constant functions, which is P_0. The derivative of any polynomial of degree $\leq n$ is a polynomial of degree $\leq n-1$. Further, *every* polynomial of degree $\leq n-1$ arises as the derivative of a polynomial of degree $\leq n$. So the image is P_{n-1}.

Performance review: 13 out of 22 people got this. 5 chose (B), 3 chose (D), 1 chose (C).

Historical note (last time): 8 out of 26 got this. 8 chose (D), 6 chose (B), 2 each chose (A) and (C).

(9) What are the kernel and image of the restriction of differentiation to all of $\mathbb{R}[x]$?

(A) The kernel and image are both $\mathbb{R}[x]$
(B) The kernel is the zero subspace and the image is $\mathbb{R}[x]$
(C) The kernel is $\mathbb{R}[x]$ and the image is the zero subspace
(D) The kernel is $\mathbb{R}[x]$ and the image is P_0 (the subspace of constant functions)
(E) The kernel is P_0 and the image is $\mathbb{R}[x]$

Answer: Option (E)

Explanation: **MAKE SURE TO READ THE LECTURE NOTES ON ABSTRACT VECTOR SPACES, SECTION 5**.

The only functions that differentiate to the zero function are the constant functions, so the kernel is P_0. The image is all of $\mathbb{R}[x]$, because every polynomial arises as the derivative of some polynomial.

Performance review: 9 out of 22 people got this. 9 chose (B), 2 chose (D), 1 chose (A), 1 left the question blank.

Historical note (last time): 9 out of 26 got this. 8 chose (B), 7 chose (D), 2 chose (C).

(10) We can use differentiation to define a linear transformation from $\mathbb{R}(x)$ to $\mathbb{R}(x)$, where we differentiate a rational function using the quotient rule for differentiation and the known rules for differentiating polynomials. What can we say about this linear transformation?
(A) The differentiation linear transformation is bijective from $\mathbb{R}(x)$ to $\mathbb{R}(x)$, i.e., every rational function is the derivative of a unique rational function.
(B) The differentiation linear transformation is injective but not surjective from $\mathbb{R}(x)$ to $\mathbb{R}(x)$, i.e., every rational function is the derivative of *at most one* rational function, but there do exist rational functions that are not expressible as the derivative of any rational function.
(C) The differentiation linear transformation is surjective but not injective from $\mathbb{R}(x)$ to $\mathbb{R}(x)$, i.e., every rational function is the derivative of *at least one* rational function, but there do exist rational functions that occur as derivatives of more than one rational function.
(D) The differentiation linear transformation is neither injective nor surjective from $\mathbb{R}(x)$ to $\mathbb{R}(x)$.

Answer: Option (D)

Explanation: **MAKE SURE TO READ THE LECTURE NOTES ON ABSTRACT VECTOR SPACES, SECTION 5**.

Differentiation is not injective because it has a nonzero kernel comprising constants. It is not surjective because there exist rational functions, such as $1/(x^2 + 1)$, that have no rational function antiderivative: the indefinite integral of $1/(x^2 + 1)$ is of the form $(\arctan x) + C$, $C \in \mathbb{R}$, and no possible function here is a rational function.

Performance review: 1 (?????) out of 22 people got this. 14 chose (C), 6 chose (B), 1 chose (A).

Historical note (last time): 16 out of 26 got this. 5 chose (B), 2 chose(C), 1 chose (A), 1 chose (E) (????), 1 non-attempt.

(11) Denote by $\mathbb{R}[[x]]$ the vector space of all formal power series with real coefficients in one variable, i.e., series of the form:

$$\sum_{i=0}^{\infty} a_i x^i$$

Formal differentiation defines a linear transformation from $\mathbb{R}[[x]]$ to itself. What can we say about this linear transformation?
(A) The formal differentiation linear transformation is bijective from $\mathbb{R}[[x]]$ to $\mathbb{R}[[x]]$.
(B) The formal differentiation linear transformation is injective but not surjective from $\mathbb{R}[[x]]$ to $\mathbb{R}[[x]]$.
(C) The formal differentiation linear transformation is surjective but not injective from $\mathbb{R}[[x]]$ to $\mathbb{R}[[x]]$.
(D) The formal differentiation linear transformation is neither injective nor surjective from $\mathbb{R}[[x]]$ to $\mathbb{R}[[x]]$.

Answer: Option (C)

Explanation: **MAKE SURE TO READ THE LECTURE NOTES ON ABSTRACT VECTOR SPACES, SECTION 5**.

The linear transformation is surjective because every formal power series can be integrated termwise to obtain another formal power series (note that we have flexibility in choosing the constant term). It is not injective, because constant formal power series are in the kernel.

Performance review: 15 out of 22 people got this. 5 chose (A), 2 chose (D).

Historical note (last time): 17 out of 26 got this. 3 each chose (A), (B), and (D).

(12) Consider the following two linear transformations $T_1, T_2 : \mathbb{R}[x] \to \mathbb{R}[x]$: T_1 is differentiation, and T_2 is multiplication by x. Which of the following is true?

(A) Both T_1 and T_2 are injective, but neither is surjective.
(B) Both T_1 and T_2 are surjective, but neither is injective.
(C) T_1 is injective but not surjective. T_2 is surjective but not injective.
(D) T_1 is surjective but not injective. T_2 is injective but not surjective.
(E) Neither T_1 nor T_2 is injective. Neither T_1 nor T_2 is surjective.

Answer: Option (D)

Explanation: We already discussed that the image of T_1 (differentiation) is all of $\mathbb{R}[x]$, and that it is not injective because it has constants in its kernel. T_2 is injective because $xp(x) = xq(x) \implies p(x) = q(x)$. However, it is not surjective because its image comprises only the polynomials with zero constant term.

Performance review: 13 out of 22 people got this. 4 chose (C), 3 chose (B), 2 chose (A).

Historical note (last time): 14 out of 26 got this. 7 chose (C), 3 chose (B), and 2 chose (E).

(13) Consider the linear transformations T_1 and T_2 of the preceding question. What can we say regarding whether T_1 and T_2 commute?

(A) T_1 and T_2 commute.
(B) T_1 and T_2 do not commute.

Answer: Option (B)

Explanation: Consider the input x. $T_1(T_2(x)) = T_1(x^2) = 2x$. On the other hand, $T_2(T_1(x)) = T_2(1) = x$.

Performance review: 12 out of 22 people got this. 10 chose (A).

Historical note (last time): 16 out of 26 got this. 7 chose (A), 3 chose (E) (?????).

TAKE-HOME CLASS QUIZ: DUE WEDNESDAY NOVEMBER 27: SIMILARITY OF LINEAR TRANSFORMATIONS

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ______________________________

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

This quiz corresponds to material discussed in the lecture notes titled `Coordinates`. It also corresponds to Section 3.4 of the text.

Recall that $n \times n$ matrices A and B are termed *similar* if there exists an invertible $n \times n$ matrix S such that $A = SBS^{-1}$. The relation of matrices being similar is an *equivalence relation* (please refer to the notes for an explanation of the terminology).

For these questions, assume $n > 1$, because a lot of phenomena are much simpler in the case $n = 1$ and you may be misled if you look only at that case. In other words, just because an equality is true for 1×1 matrices, do not assume it is always true. On the other hand, if you can find *counterexamples* to a statement for 1×1 matrices, you can probably use that to construct counterexamples for all sizes of matrices by using scalar matrices.

(1) Which of the following can we say about two (possibly equal, possibly distinct) similar $n \times n$ matrices A and B? Please see Options (D) and (E) before answering.
 (A) A is invertible if and only if B is invertible.
 (B) A is nilpotent if and only if B is nilpotent.
 (C) A is idempotent if and only if B is idempotent.
 (D) All of the above.
 (E) None of the above.

 Your answer: ______________________

(2) Which of the following can we say about two (possibly equal, possibly distinct) similar $n \times n$ matrices A and B? Please see Options (D) and (E) before answering.
 (A) A is scalar if and only if B is scalar.
 (B) A is diagonal if and only if B is diagonal.
 (C) A is upper triangular if and only if B is upper triangular.
 (D) All of the above.
 (E) None of the above.

 Your answer: ______________________

(3) Suppose A_1, A_2, B_1, B_2 are $n \times n$ matrices such that A_1 is similar to B_1 and A_2 is similar to B_2. Which of the following is *definitely* true? Please see Options (D) and (E) before answering.
 (A) $A_1 + A_2$ is similar to $B_1 + B_2$.
 (B) $A_1 - A_2$ is similar to $B_1 - B_2$.
 (C) $A_1 A_2$ is similar to $B_1 B_2$.
 (D) All of the above.
 (E) None of the above.

 Your answer: ______________________

(4) Suppose A_1, A_2, B_1, B_2 are $n \times n$ matrices such that A_1 is similar to B_1 and A_2 is similar to B_2. Which of the following is *definitely* true? Please see Options (D) and (E) before answering.
 (A) $A_1 + B_1$ is similar to $A_2 + B_2$.
 (B) $A_1 - B_1$ is similar to $A_2 - B_2$.

(C) A_1B_1 is similar to A_2B_2.
(D) All of the above.
(E) None of the above.

Your answer: ____________________________

(5) Suppose A and B are both $n \times n$ matrices (but they are not given to be similar). Which of the following holds?
(A) A is similar to B if and only if $-A$ is similar to $-B$.
(B) If A is similar to B, then $-A$ is similar to $-B$. However, $-A$ being similar to $-B$ does not imply that A is similar to B.
(C) If $-A$ is similar to $-B$, then A is similar to B. However, A being similar to B does not imply that $-A$ is similar to $-B$.
(D) A being similar to B does not imply that $-A$ is similar to $-B$. Also, $-A$ being similar to $-B$ does not imply that A is similar to B.

Your answer: ____________________________

(6) Suppose A and B are both $n \times n$ matrices (but they are not given to be similar). Which of the following holds?
(A) A is similar to B if and only if $2A$ is similar to $2B$.
(B) If A is similar to B, then $2A$ is similar to $2B$. However, $2A$ being similar to $2B$ does not imply that A is similar to B.
(C) If $2A$ is similar to $2B$, then A is similar to B. However, A being similar to B does not imply that $2A$ is similar to $2B$.
(D) A being similar to B does not imply that $2A$ is similar to $2B$. Also, $2A$ being similar to $2B$ does not imply that A is similar to B.

Your answer: ____________________________

(7) Suppose A and B are both invertible $n \times n$ matrices (but they are not given to be similar). Which of the following holds?
(A) A is similar to B if and only if A^{-1} is similar to B^{-1}.
(B) If A is similar to B, then A^{-1} is similar to B^{-1}. However, A^{-1} being similar to B^{-1} does not imply that A is similar to B.
(C) If A^{-1} is similar to B^{-1}, then A is similar to B. However, A being similar to B does not imply that A^{-1} is similar to B^{-1}
(D) A being similar to B does not imply that A^{-1} is similar to B^{-1}. Also, A^{-1} being similar to B^{-1} does not imply that A is similar to B.

Your answer: ____________________________

(8) Suppose A and B are both $n \times n$ matrices (but they are not given to be similar). Which of the following holds?
(A) A is similar to B if and only if A^2 is similar to B^2.
(B) If A is similar to B, then A^2 is similar to B^2. However, A^2 being similar to B^2 does not imply that A is similar to B.
(C) If A^2 is similar to B^2, then A is similar to B. However, A being similar to B does not imply that A^2 is similar to B^2.
(D) A being similar to B does not imply that A^2 is similar to B^2. Also, A^2 being similar to B^2 does not imply that A is similar to B.

Your answer: ____________________________

(9) Suppose A and B are $n \times n$ matrices (but they are not given to be similar and they are not given to be invertible). We say that A and B are *quasi-similar* (not a standard term!) if there exist $n \times n$

matrices C and D such that $A = CD$ and $B = DC$. What can we say is the relation between being similar and being quasi-similar?

(A) A and B are similar if and only if they are quasi-similar.

(B) If A and B are similar, they are quasi-similar. However, the converse is not necessarily true: A and B may be quasi-similar but not similar.

(C) If A and B are quasi-similar, they are similar. However, the converse is not necessarily true: A and B may be similar but not quasi-similar.

(D) Neither implies the other. A and B may be similar but not quasi-similar. Also, A and B may be quasi-similar but not similar.

Your answer: ____________________________

(10) With the notion of quasi-similar as defined in the preceding question, what can we say about the relation between being similar and being quasi-similar for $n \times n$ matrices A and B that are both given to be *invertible*?

(A) A and B are similar if and only if they are quasi-similar.

(B) If A and B are similar, they are quasi-similar. However, the converse is not necessarily true: A and B may be quasi-similar but not similar.

(C) If A and B are quasi-similar, they are similar. However, the converse is not necessarily true: A and B may be similar but not quasi-similar.

(D) Neither implies the other. A and B may be similar but not quasi-similar. Also, A and B may be quasi-similar but not similar.

Your answer: ____________________________

(11) Suppose A and B are two $n \times n$ matrices. Which of the following best describes the relation between similarity and having the same rank?

(A) A and B are similar if and only if they have the same rank.

(B) If A and B are similar, then they have the same rank. However, it is possible for A and B to have the same rank but not be similar.

(C) If A and B have the same rank, then they are similar. However, it is possible for A and B to be similar but not have the same rank.

(D) A and B may be similar but have different ranks. Also, A and B may have the same rank but not be similar.

Your answer: ____________________________

(12) Suppose A and B are two $n \times n$ matrices. Which of the following best describes the relation between quasi-similarity and having the same rank?

(A) A and B are quasi-similar if and only if they have the same rank.

(B) If A and B are quasi-similar, then they have the same rank. However, it is possible for A and B to have the same rank but not be quasi-similar.

(C) If A and B have the same rank, then they are quasi-similar. However, it is possible for A and B to be quasi-similar but not have the same rank.

(D) A and B may be quasi-similar but have different ranks. Also, A and B may have the same rank but not be quasi-similar.

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE MONDAY DECEMBER 2: SIMILARITY OF LINEAR TRANSFORMATIONS (APPLIED)

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ____

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

This quiz corresponds to material discussed in the lecture notes titled `Coordinates`. It also corresponds to Section 3.4 of the text.

Recall that $n \times n$ matrices A and B are termed *similar* if there exists an invertible $n \times n$ matrix S such that $A = SBS^{-1}$. The relation of matrices being similar is an *equivalence relation*.

Recall that $n \times n$ matrices A and B are termed *quasi-similar* if there exist $n \times n$ matrices C and D such that $A = CD$ and $B = DC$. Recall that similar matrices are always quasi-similar, but quasi-similar matrices need not be similar. However, for *invertible* matrices, similarity and quasi-similarity are equivalent.

Also, note that if A and B are quasi-similar matrices, then A and B have the same trace. However, the converse is not true: it is possible to have two matrices A and B that have the same trace but are not quasi-similar.

For these questions, assume $n > 1$, because a lot of phenomena are much simpler in the case $n = 1$ and you may be misled if you look only at that case.

Note also that the trace of a square matrix is defined as the sum of its diagonal entries.

The *determinant* of a 2×2 matrix, denoted det, is defined as follows:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

The following are some important facts about the determinant:

- The determinant of a 2×2 diagonal matrix is the product of the diagonal entries.
- The determinant of a 2×2 matrix is zero if and only if the matrix is non-invertible.
- The determinant of the product of two 2×2 matrices is the product of the determinants.
- The determinant of the inverse of an invertible 2×2 matrix is the reciprocal of the determinant.
- If A and B are similar 2×2 matrices, they have the same determinant.
- If A and B are quasi-similar 2×2 matrices, they have the same determinant.
- If the determinant of A is positive, then the linear transformation given by A is an orientation-preserving linear automorphism of $\mathbb{R}^2$.
- If the determinant of A is negative, then the linear transformation given by A is an orientation-reversing linear automorphism of $\mathbb{R}^2$.

(1) Suppose A and B are both $n \times n$ matrices (but they are not given to be similar). Denote by I_n the $n \times n$ identity matrix. Which of the following holds?
  (A) A is similar to B if and only if $A - I_n$ is similar to $B - I_n$.
  (B) If A is similar to B, then $A - I_n$ is similar to $B - I_n$. However, $A - I_n$ being similar to $B - I_n$ does not imply that A is similar to B.
  (C) If $A - I_n$ is similar to $B - I_n$, then A is similar to B. However, A being similar to B does not imply that $A - I_n$ is similar to $B - I_n$.
  (D) A being similar to B does not imply that $A - I_n$ is similar to $B - I_n$. Also, $A - I_n$ being similar to $B - I_n$ does not imply that A is similar to B.

  Your answer: ____________________________

Suppose f is a polynomial of degree r in one variable with real coefficients. For a $n \times n$ matrix X, we denote by $f(X)$ we mean the matrix we get by applying the polynomial to f, where constant terms are interpreted as scalar matrices. For instance, if $f(x) = x^2 + 3x + 5$, then $f(X) = X^2 + 3X + 5I_n$.

(2) Suppose A and B are both $n \times n$ matrices (but they are not given to be similar). Suppose f is a polynomial of degree r in one variable, where $r \geq 2$. Which of the following holds?
(A) A is similar to B if and only if $f(A)$ is similar to $f(B)$.
(B) If A is similar to B, then $f(A)$ is similar to $f(B)$. However, $f(A)$ being similar to $f(B)$ does not imply that A is similar to B.
(C) If $f(A)$ is similar to $f(B)$, then A is similar to B. However, A being similar to B does not imply that $f(A)$ is similar to $f(B)$.
(D) A being similar to B does not imply that $f(A)$ is similar to $f(B)$. Also, $f(A)$ being similar to $f(B)$ does not imply that A is similar to B.

Your answer: ____________________________

(3) Suppose A and B are both $n \times n$ matrices (but they are not given to be similar). Suppose f is a polynomial of degree r in one variable, where $r = 1$. Which of the following holds?
(A) A is similar to B if and only if $f(A)$ is similar to $f(B)$.
(B) If A is similar to B, then $f(A)$ is similar to $f(B)$. However, $f(A)$ being similar to $f(B)$ does not imply that A is similar to B.
(C) If $f(A)$ is similar to $f(B)$, then A is similar to B. However, A being similar to B does not imply that $f(A)$ is similar to $f(B)$.
(D) A being similar to B does not imply that $f(A)$ is similar to $f(B)$. Also, $f(A)$ being similar to $f(B)$ does not imply that A is similar to B.

Your answer: ____________________________

Suppose p and q are real numbers (possibly equal, possibly distinct). The diagonal matrices:

$$A = \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix}$$

and

$$B = \begin{bmatrix} q & 0 \\ 0 & p \end{bmatrix}$$

are similar. Explicitly, the two matrices are similar under the change-of-basis transformation that interchanges the coordinates, i.e., if we set:

$$S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

then:

$$S = S^{-1}$$

and we have:

$$B = S^{-1}AS$$

Moreover, the only diagonal matrices similar to A are A and B (in the special case that $p = q$, we get $A = B$ is a scalar matrix, so A is the only diagonal matrix similar to A).

(4) What is the necessary and sufficient condition on p and q such that the matrix $A = \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix}$ is similar to $-A$?

(A) $p = q$
(B) $p = -q$
(C) $p = 1/q$
(D) $p = -1/q$
(E) $p + q = 1$

Your answer: ____________________________

(5) Which of the following is a necessary and sufficient condition on p and q so that the matrix $A = \begin{bmatrix} p & 0 \\ 0 & q \end{bmatrix}$ is invertible and similar to $-A^{-1}$?
(A) $p = q$
(B) $p = -q$
(C) $p = 1/q$
(D) $p = -1/q$
(E) $p + q = 1$

Your answer: ____________________________

Consider the matrix:

$$S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

used above. We have $S = S^{-1}$. For a general matrix:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

we have:

$$S^{-1}AS = \begin{bmatrix} d & c \\ b & a \end{bmatrix}$$

In other words, it swaps the rows *and* swaps the columns. This observation may be useful for some of the following questions.

(6) For an angle θ with $-\pi \leq \theta \leq \pi$, the rotation matrix for θ is given as:

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Note that $R(-\pi) = R(\pi)$, but other than that equality, all the $R(\theta)$s are distinct.

Which of these describes the relation between the rotation matrices for different values of θ?
(A) All the rotation matrices $R(\theta)$, $-\pi < \theta \leq \pi$, are similar to each other.
(B) The rotation matrix $R(\theta)$ is similar to itself and to the rotation matrix $R(-\theta)$. However, it is not in general similar to any other rotation matrix.
(C) No two different rotation matrices are similar.
(D) The rotation matrix $R(\theta)$ is similar to itself and to the rotation matrix $R(\pi - \theta)$ (or $R(-\pi - \theta)$, depending on which of the two angles lies within the specified range). However, it is not in general similar to any other rotation matrix.

Your answer: ____________________________

(7) Consider the linear automorphisms of $\mathbb{R}^2$ that are given as *reflections* about lines in $\mathbb{R}^2$ through the origin. (Note that we need the line of reflection to pass through the origin for the automorphism to be *linear* rather than merely being *affine linear*). Which of these describes the relation between reflection matrices for different possible lines of reflection through the origin?

(A) All the reflection matrices are similar to each other.
(B) No two reflection matrices for different lines of reflection are similar.
(C) The reflection matrices for two different lines of reflection are similar if and only if the lines of reflection are perpendicular.
(D) The reflection matrices for two different lines of reflection are similar if and only if the lines are reflection make an angle that is a rational multiple of π.

Your answer: ____________________________

(8) Suppose m and n are positive integers with $m < n$. Denote by P_m the "orthogonal projection onto the first m coordinates" linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$, defined as follows. This takes as input a n-dimensional vector, sends each of the first m coordinates to itself, and sends the remaining coordinates to zero. What is the trace of the matrix of P_m?
(A) 1
(B) m
(C) n
(D) $n - m$
(E) $m - n$

Your answer: ____________________________

(9) It is a fact that if A, B are $n \times n$ matrices that describe orthogonal projections onto (possibly different) m-dimensional subspaces of $\mathbb{R}^n$, then A and B are similar. What can we say must be the trace of an orthogonal projection onto any m-dimensional subspace of $\mathbb{R}^n$?
(A) 1
(B) m
(C) n
(D) $n - m$
(E) $m - n$

Your answer: ____________________________

(10) Suppose A, B and C are $n \times n$ matrices. Which of the following matrices is *not* guaranteed (based on the given information) to have the same trace as the product ABC? Please see (and read carefully) Options (D) and (E) before answering.
(A) BCA
(B) CBA
(C) CAB
(D) None the above, i.e., they are all guaranteed to have the same trace as ABC.
(E) All of the above, i.e., none of them is guaranteed to have the same trace as ABC.

Your answer: ____________________________

(11) Which of the following gives a pair of matrices A and B that have the same trace as each other *and* the same determinant as each other, but that are *not* similar to each other?
(A) $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$
(B) $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
(C) $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$
(D) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$
(E) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

Your answer: ___________________________

(12) Suppose A and B are 2×2 matrices. Which of the following correctly describes the relation between $\det A$, $\det B$, and $\det(A+B)$? Please see Option (E) before answering.
(A) $\det(A+B) = \det A + \det B$
(B) $\det(A+B) \leq \det A + \det B$, but equality need not necessarily hold.
(C) $\det(A+B) \geq \det A + \det B$, but equality need not necessarily hold.
(D) $|\det(A+B)| \leq |\det A| + |\det B|$, but equality need not necessarily hold.
(E) None of the above.

Your answer: ___________________________

Let n be a natural number greater than 1. Suppose $f : \{0, 1, 2, \dots, n\} \to \{0, 1, 2, \dots, n\}$ is a function satisfying $f(0) = 0$. Let T_f denote the linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ satisfying the following for all $i \in \{1, 2, \dots, n\}$:

$$T_f(\vec{e}_i) = \begin{cases} \vec{e}_{f(i)}, & f(i) \neq 0 \\ \vec{0}, & f(i) = 0 \end{cases}$$

Let M_f denote the matrix for the linear transformation T_f. M_f can be described explicitly as follows: the i^{th} column of M_f is $\vec{0}$ if $f(i) = 0$ and is $\vec{e}_{f(i)}$ if $f(i) \neq 0$.

Note that if $f, g : \{0, 1, 2, \dots, n\} \to \{0, 1, 2, \dots, n\}$ are functions with $f(0) = g(0) = 0$, then $M_{f \circ g} = M_f M_g$ and $T_{f \circ g} = T_f \circ T_g$.

For the following questions, the discussion prior to Question 3 might be helpful. Note, however, that while that discussion gives one possible candidate for the matrix S of the similarity transformation, it is not the only possible candidate. For some but not all of the following questions, in the case that two matrices are similar, the matrix S described there works. In the case that they are not similar, the lack of similarity can be inferred from the traces not being equal, or from the determinants not being equal.

(13) $n = 2$ for this question. For the following three functions f, g, and h, consider the corresponding matrices M_f, M_g, M_h. Either two of them are similar and the third is not similar to either (in which case, select the matrix that is not similar to the other two), or all three are similar (if so, select Option (D)), or no two are similar (if so, select Option (E)).
(A) $f(0) = 0, f(1) = 1, f(2) = 0$
(B) $g(0) = 0, g(1) = 0, g(2) = 2$
(C) $h(0) = 0, h(1) = 1, h(2) = 1$
(D) All the above give similar matrices.
(E) No two of the corresponding matrices are similar.

Your answer: ___________________________

(14) $n = 2$ for this question. For the following three functions f, g, and h, consider the corresponding matrices M_f, M_g, M_h. Either two of them are similar and the third is not similar to either (in which case, select the matrix that is not similar to the other two), or all three are similar (if so, select Option (D)), or no two are similar (if so, select Option (E)).
(A) $f(0) = 0, f(1) = 0, f(2) = 1$
(B) $g(0) = 0, g(1) = 2, g(2) = 0$
(C) $h(0) = 0, h(1) = 2, h(2) = 1$
(D) All the above give similar matrices.
(E) No two of the corresponding matrices are similar.

Your answer: ___________________________

(15) $n = 3$ for this question. For the following three functions f, g, and h, consider the corresponding matrices M_f, M_g, M_h. Either two of them are similar and the third is not similar to either (in which case, select the matrix that is not similar to the other two), or all three are similar (if so, select Option (D)), or no two are similar (if so, select Option (E)).

(A) $f(0) = 0, f(1) = 2, f(2) = 1, f(3) = 3$
(B) $g(0) = 0, g(1) = 1, g(2) = 3, g(3) = 2$
(C) $h(0) = 0, h(1) = 3, h(2) = 2, h(3) = 1$
(D) All the above give similar matrices.
(E) No two of the corresponding matrices are similar.

Your answer: ____________________________

(16) $n = 3$ for this question. For the following three functions f, g, and h, consider the corresponding matrices M_f, M_g, M_h. Either two of them are similar and the third is not similar to either (in which case, select the matrix that is not similar to the other two), or all three are similar (if so, select Option (D)), or no two are similar (if so, select Option (E)).
(A) $f(0) = 0, f(1) = 1, f(2) = 2, f(3) = 3$
(B) $g(0) = 0, g(1) = 2, g(2) = 3, g(3) = 1$
(C) $h(0) = 0, h(1) = 3, h(2) = 1, h(3) = 2$
(D) All the above give similar matrices.
(E) No two of the corresponding matrices are similar.

Your answer: ____________________________

TAKE-HOME CLASS QUIZ: DUE WEDNESDAY DECEMBER 4: ORDINARY LEAST SQUARES REGRESSION

MATH 196, SECTION 57 (VIPUL NAIK)

Your name (print clearly in capital letters): ______________________________

PLEASE FEEL FREE TO DISCUSS *ALL* QUESTIONS.

(1) Assume no measurement error. Consider the situation where we have a function f of the form $f(x) = a_0 + a_1x$ with unknown values of the parameters a_0 and a_1. We collect n distinct input-output pairs, i.e., we collect n distonct inputs and compute the outputs for them. The coefficient matrix for the system is a $n \times 2$ matrix (the rows correspond to the input values, and the columns correspond to the unknown parameters). What is the rank of this matrix?
(A) It is always 2
(B) It is always n
(C) It is always $\min\{2, n\}$
(D) It is always $\max\{2, n\}$

Your answer: ______________________

(2) Assume no measurement error. Consider the situation where we have a function f of the form $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m$ with unknown values of the parameters a_0, a_1, ..., a_m. We collect n distinct input-output pairs, i.e., we collect n distinct inputs and compute the outputs for them. The coefficient matrix for the system is a $n \times (m + 1)$ matrix (the rows correspond to the input values, and the columns correspond to the unknown parameters). What is the rank of this matrix?
(A) It is always $m + 1$
(B) It is always n
(C) It is always $\min\{m + 1, n\}$
(D) It is always $\max\{m + 1, n\}$

Your answer: ______________________

(3) Assume no measurement error. Consider the situation where we have a function f of the form $f(x, y) = a_0 + a_1x + a_2y$ with unknown values of the parameters a_0, a_1, and a_2. We collect n distinct input-output pairs, i.e., we collect n distinct inputs (here an input specification involves specifying both the x-value and the y-value) and compute the outputs for them. The coefficient matrix for the system is a $n \times 3$ matrix (the rows correspond to the input values, and the columns correspond to the unknown parameters). What is the rank of this matrix?
(A) It is always $\min\{3, n\}$
(B) It is always $\max\{3, n\}$
(C) For $n = 1$, it is 1. For $n \geq 2$, it is 2 if the input points are all collinear in the xy-plane. Otherwise, it is 3.
(D) For $n = 1$, it is 1. For $n \geq 2$, it is 3 if the input points are all collinear in the xy-plane. Otherwise, it is 2.

Your answer: ______________________

(4) Which of the following is closest to correct in the setting where we use a linear system to find the parameters using input-output pairs given a functional form that is linear in the parameters? Assume for simplicity that we are dealing with a functional form $y = f(x)$ with one input and one output, but possibly multiple parameters in the general description.

(A) The solutions to the linear system that we set up correspond to possibilities for the inputs to the function, and geometrically correspond to choices of points x for the graph $y = f(x)$.
(B) The solutions to the linear system that we set up correspond to possibilities for the inputs to the function, and geometrically correspond to different possible choices for the line or curve that is the graph $y = f(x)$.
(C) The solutions to the linear system that we set up correspond to possibilities for the parameters, and geometrically correspond to choices of points x for the graph $y = f(x)$.
(D) The solutions to the linear system that we set up correspond to possibilities for the parameters, and geometrically correspond to different possible choices for the line or curve that is the graph $y = f(x)$.

Your answer: ____________________________

(5) Continuing with the notation and setup of the preceding question, consider the coefficient matrix of the linear system. This matrix defines a linear transformation from the vector space of possible parameter values to the vector space of the outputs of the function. What is the image of this linear transformation?
(A) The image is the set of possible output values for which the linear system is consistent, i.e., we can find *at least one* function f of the required functional form that fits all the input-output pairs with *no measurement error*.
(B) The image is the set of possible output values for which the linear system has *at most one solution*, i.e., the set of output valus for which we can find *at most one* function f of the required functional form that fits all the input-output pairs with *no measurement error*.

Your answer: ____________________________

(6) Consider the case of polynomial regression for a polynomial function of one variable, allowing for measurement error. We believe that a function has the form of a polynomial. We can tentatively choose a degree m for the polynomial we are trying to fit, and a value n for the number of distinct inputs for which we compute the corresponding outputs to obtain input-output pairs (i.e., data points). We will get a $n \times (m + 1)$ coefficient matrix. Which of the following correctly describes what we should try for?
(A) We should choose n and $m + 1$ to be exactly equal, so that we get a unique polynomial.
(B) We should choose n to be greater than $m+1$, so that the system is guaranteed to be consistent and we can find the polynomial.
(C) We should choose n to be less than $m + 1$, so that the system is guaranteed to be consistent and we can find the polynomial.
(D) We should choose n to be greater than $m + 1$, so that the system is *not* guaranteed to be consistent, but we do have a unique solution after we project the output vector to a vector for which the system is consistent.
(E) We should choose n to be less than $m+1$, so that the system is *not* guaranteed to be consistent, but we do have a unique solution after we project the output vector to a vector for which the system is consistent.

Your answer: ____________________________

(7) Consider the general situation of linear regression. Denote by X the coefficient matrix for the linear system (also called the design matrix). Denote by $\vec{\beta}$ the parameter vector that we are trying to solve for. Denote by $\vec{y}$ an observed output vector. The idea in ordinary least squares regression is to choose a suitable vector $\vec{\varepsilon}$ such that the linear system $X\vec{\beta} = \vec{y} - \vec{\varepsilon}$ can be solved for $\vec{\beta}$. Among the many possibilities that we can choose for $\vec{\varepsilon}$, what criterion do we use to select the appropriate choice? Recall that the *length* of a vector is the square root of the sum of squares of its coordinates.
(A) We choose $\vec{\varepsilon}$ to have the minimum length possible subject to the constraint that $X\vec{\beta} = \vec{y} - \vec{\varepsilon}$ has a solution.

(B) We choose $\vec{\varepsilon}$ such that the system $X\vec{\beta} = \vec{y} - \vec{\varepsilon}$ can be solved and such that the solution vector $\vec{\beta}$ has the minimum possible length (among all such choices of $\vec{\varepsilon}$).

(C) We choose $\vec{\varepsilon}$ such that the system $X\vec{\beta} = \vec{y} - \vec{\varepsilon}$ can be solved and such that the difference vector $\vec{y} - \vec{\varepsilon}$ has the minimum possible length (among all such choices of $\vec{\varepsilon}$).

Your answer: ___________________________

(8) We have data for the logarithm of annual per capita GDP for a country for the last 100 years. We want to see if this fits a polynomial model. The idea is to try to first fit a polynomial of degree 0 (i.e., per capita GDP remains constant), then fit a polynomial of degree ≤ 1 (i.e., per capita GDP grows or decays exponentially), then fit a polynomial of degree ≤ 2 (i.e., per capita GDP grows or decays as the exponential of a quadratic function), and so on.

What happens to the length of the error vector $\vec{\varepsilon}$ as we increase the degree of the polynomial that we are trying to fit?

(A) The error vector $\vec{\varepsilon}$ keeps getting smaller and smaller in length, with a probability of almost 1 that it keeps *strictly* decreasing in length at each step, until the error vector becomes $\vec{0}$ (which we expect will happen when we get to the stage of trying to fit the function using a polynomial of degree 99).

(B) The error vector $\vec{\varepsilon}$ keeps getting larger and larger in length, with a probability of almost 1 that it keeps *strictly* increasing in length at each step, until the error vector becomes $\vec{y}$ (which we expect will happen when we get to the stage of trying to fit the function using a polynomial of degree 99).

Your answer: ___________________________