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

MATH 196, SECTION 57 (VIPUL NAIK)

1. Performance review

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

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

The question-wise answers and performance review are below:

(1) Option (E): 21 people
(2) Option (D): 15 people
(3) Option (D): 14 people
(4) Option (C): 18 people
(5) Option (D): 23 people
(6) Option (B): 21 people
(7) Option (E): 14 people
(8) Option (C): 5 people
(9) Option (A): 16 people
(10) Option (A): 15 people
(11) Option (B): 17 people
(12) Option (E): 4 people
(13) Option (E): 16 people

2. Solutions

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.

Answer: Option (E)

Explanation: We have the following formula for the i^{th} diagonal entry of the product:

$$c_{ii} = \sum_{j=1}^{n} a_{ij} b_{ji}$$

The sum of the diagonal entries of C is thus:

$$\sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ji}$$

The j^{th} diagonal entry of D is:

$$d_{jj} = \sum_{i=1}^{n} b_{ji} a_{ij}$$

The sum of the diagonal entries of D is thus:

$$\sum_{j=1}^{n} d_{jj} = \sum_{j=1}^{n} \sum_{i=1}^{n} b_{ji} a_{ij}$$

Thus, the sum of the diagonal entries of C is the same as the sum of the diagonal entries of D.

Performance review: 21 out of 25 got this. 2 chose (C), 1 each chose (B) and (D).

Historical note (last time): 23 out of 26 got this. 1 each chose (B), (C), and (D).

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

Answer: Option (D)

Explanation: We can write $X = A^{-1}(AX)$, whereas $AXA^{-1} = (AX)A^{-1}$. Thus, both X and AXA^{-1} are products of two matrices A^{-1} and AX but in opposite orders. Hence, by the preceding question, they have the same trace.

Performance review: 15 out of 25 got this. 3 each chose (A) and (C), 2 each chose (B) and (E).

Historical note (last time): 20 out of 26 got this. 5 chose (A), 1 chose (C).

(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.

Answer: Option (D)

Explanation: Essentially, the conjugation operation preserves all aspects of the multiplicative structure, hence it preserves the properties of being invertible, nilpotent, and idempotent.

Let us illustrate this with idempotent. We have that:

$$AX^2A^{-1} = (AXA^{-1})^2$$

If $X^2 = X$, we get:

$$AXA^{-1} = (AXA^{-1})^2$$

showing that AXA^{-1} is also idempotent. We can work backwards to show that the reverse implication also holds.

Performance review: 14 out of 25 got this. 4 chose (B), 3 chose (A), 2 chose (D), 1 chose (C), and 1 left the question blank.

Historical note (last time): 18 out of 26 got this. 3 each chose (A) and (B), 2 chose (C).

(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

Answer: Option (C)

Explanation: Start with:

$$AXA^{-1} = X$$

Multiply both sides of the equation on the *right* with the matrix A. We get:

$$AX = XA$$

Note that the algebraic manipulation is reversible: starting from $AX = XA$ and multiplying both sides by A^{-1} on the right gives the original relation.

Performance review: 18 out of 25 got this. 4 chose (E), 1 each chose (A), (B), and (D).

Historical note (last time): 18 out of 26 got this. 4 chose (D) 2 chose (A), 1 each chose (B) and (E).

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

Answer: Option (D)

Explanation: We can just carry out the matrix multiplication. Note that:

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

A *sanity check* is that the new matrix C should have the same trace as the original matrix B, because the trace is invariant under conjugation. Among the options, the only matrix with the correct trace (3) is Option (D).

Performance review: 23 out of 25 got this. 2 chose (B).

Historical note (last time): 23 out of 26 got this. 2 chose (B), 1 chose (C).

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

Answer: Option (B)

Explanation: We calculate:

$$C^8 = \begin{bmatrix} 1 & 0 \\ 0 & 256 \end{bmatrix}$$

We now recover B^8 as $A^{-1}C^8A$.

Performance review: 21 out of 25 got this. 2 chose (A), 1 each chose (C) and (E).

Historical note (last time): 22 out of 26 got this. 3 chose (C), 1 chose (E).

(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

Answer: Option (E)

Explanation: The linear transformation corresponding to A preserves lengths of vectors. Thus, the images of the standard basis vectors are all unit vectors. Recall that the images of the standard basis vectors under the linear transformation corresponding to a matrix are the columns of that matrix. Therefore, each column of the matrix is a unit vector, i.e., the sum of the squares of the coordinates there is 1.

As an illustration, consider the case of a rotation matrix in two dimensions:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Performance review: 14 out of 25 got this. 6 chose (B), 3 chose (A), 2 chose (D).

Historical note (last time) :12 out of 26 got this. 4 each chose (A), (C), and (D). 1 chose (B). 1 left the question blank.

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 $C^1(\mathbb{R})$
Answer: Option (C)
Explanation: $C^1(\mathbb{R})$ is the space of continuously differentiable functions, and differentiating any continuously differentiable function gives rise to a continuous function. We cannot go as broad as $C^0(\mathbb{R})$ because not all functions in $C^0(\mathbb{R})$ are differentiable. For instance, the absolute value function is not differentiable at 0.
Note that everything in $C^0(\mathbb{R})$ does get hit, because every continuous function is the derivative of its antiderivative.
Performance review: 5 out of 25 got this. 8 chose (A), 6 chose (B), 5 chose (E), 1 chose (D).
Historical note (last time): 18 out of 26 got this. 4 chose (A), 2 each chose (B) and (E).

(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) $C^\infty(\mathbb{R})$
Answer: Option (A)
Explanation: $C^k(\mathbb{R})$ is the space of functions that are at least k times continuously differentiable. Thus, the derivative of any function here is at least $(k-1)$ times continuously differentiable. Thus, the iamge is in $C^{k-1}(\mathbb{R})$. The image is the whole of $C^{k-1}(\mathbb{R})$ because any function in $C^{k-1}(\mathbb{R})$ can be integrated to get a function in $C^k(\mathbb{R})$.
Performance review: 16 out of 25 got this. 9 chose (C).
Historical note (last time): 20 out of 26 got this. 2 each chose (C) and (D). 1 each chose (B) and (E).

(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) $C^1(\mathbb{R})$
Answer: Option (A)
Explanation: The derivative of a function on $\mathbb{R}$ is zero if and only if the function is a constant function.
Performance review: 15 out of 25 got this. 7 chose (E), 2 chose (B), 1 chose (D).
Historical note (last time): 11 out of 26 got this. 8 chose (B), 4 chose (C), 2 chose (E), and 1 left the question blank.

(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

Answer: Option (B)

Explanation: Each differentiation reduces the degree of the polynomial by 1, unless we are already at a constant, in which case we differentiate to 0. So, if the degree of the polynomial is at most $k-1$, differentiating k times gives 0. Conversely, if differentating k times gives zero, then repeated integration gives a generic polynomial of degree at most $k-1$.

Performance review: 17 out of 25 got this. 4 chose (D), 3 chose (C), 1 chose (E).

Historical note (last time): 17 out of 26 got this. 3 each chose (A) and (D), 2 chose (C), and 1 left the question blank.

(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) $C^\infty(\mathbb{R})$

Answer: Option (E)

Explanation: *All* polynomials are infinitely differentiable, so they are all in $C^\infty(\mathbb{R})$, the smallest of the spaces listed.

Performance review: 4 out of 25 got this. 11 chose (D), 7 chose (C), 2 chose (B), 1 left the question blank.

Historical note (last time): 7 out of 26 got this. 11 chose (D), 6 chose (C), 1 chose (B), and 1 left the question blank.

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) $f \mapsto \int_0^1 (f(x))^2\,dx$

Answer: Option (E)

Explanation: This is easy to see from the description. The key point here is that if we square *after* evaluation, then that is not linear. Squaring prior to evaluation is fine. Explicitly, the point is that:

$$(f+g)(x^2) = f(x^2) + g(x^2)$$

But the following is *not* true generally:

$$((f+g)(x))^2 = (f(x))^2 + (g(x))^2$$

In fact, the left side simplifies to $(f(x))^2 + (g(x))^2 + 2f(x)g(x)$.

Performance review: 16 out of 25 got this. 6 chose (D), 1 each chose (A), (B), and (C).

Historical note (last time): 15 out of 26 got this. 8 chose (D), 1 each chose (B) and (C), and 1 left the question blank.