AS.110.201 Linear Algebra Final Week Notes

Ben Marlin

December 2024

Problem 1: Let V and W be n dimensional real vector spaces. Let $\mathrm{Hom}(V, W)$ denote that set of all linear transformations $T : V \to W$.

1. Define natural operations of addition and scalar multiplication on $\mathrm{Hom}(V, W)$ so that it is a vector space with these operations.
2. What is the dimension of $\mathrm{Hom}(V, W)$?
3. Let $\mathrm{Mat}_{n\times n}(\mathbb{R})$ denote the vector space of $n \times n$ matrices with real coefficients. Define an explicit linear isomorphism $\varphi : \mathrm{Hom}(V, W) \to \mathrm{Mat}_{n\times n}(\mathbb{R})$.

Problem 2: Consider the vector space $C([a, b])$ consisting of all continuous functions $f : [a, b] \to \mathbb{R}$.

1. Show that the set of all polynomial functions $p : [a, b] \to \mathbb{R}$ forms a subspace of $C([a, b])$.
2. Is the above subspace finite dimensional? Find a basis.
3. Consider $C(\mathbb{R})$ (all continuous real-valued functions on $\mathbb{R}$). Is the function $T_a(f) = f(x + a)$ a linear transformation on $C(\mathbb{R})$?

Problem 3: Recall that matrices A and B are said to be similar if there exists an invertible matrix S such that $A = SBS^{-1}$. A matrix is said to be diagonalizable if it is similar to a diagonal matrix.

1. If a matrix is diagonalizable is it necessarily invertible? Prove or give a counterexample.
2. If a matrix is invertible is it necessarily diagonalizable? Prove or give a counterexample.
3. If a matrix has all nonzero eigenvalues is it true that it is invertible?
4. If a matrix has $\lambda = 0$ as one of its eigenvalues can it be invertible?

Problem 4: Recall that vectors v and w are orthogonal if their dot product is zero.

1. Suppose that v and w are orthogonal. Prove that the Pythagorean theorem holds: $||v||^2 + ||w||^2 = ||v + w||^2$.

2. Find the matrix that rotates a vector $v \in \mathbb{R}^2$ by an angle of $\frac{\pi}{2}$.

3. Challenge: Consider the 3-dimensional Euclidean space $\mathbb{R}^3$. Let $P_y : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear map that projects onto the plane $z = 0$. Is the matrix representation of P_y diagonalizable?

True or False:

1. An invertible matrix is the product of elementary matrices.

2. The kernel of RREF(A) is the same as the kernel of A for any matrix A.

3. The image of RREF(A) is the same as the image of A for any matrix A.

4. Symmetric matrices are diagonalizable.

5. The eigenvalues of a symmetric matrix are all imaginary.

6. A matrix is invertible if and only if its determinant is greater than zero.

7. If V is n-dimensional and $W \subseteq V$ is a subspace, then $\dim(W) \leq n$.

8. A matrix with real coefficients always has real eigenvalues.

9. If λ is an eigenvalue of a matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ with eigenvector $v \in \mathbb{R}^n$, then $\lambda \in \mathbb{R}$.

AS.110.201 Linear Algebra Week 9 Notes

Ben Marlin

October 2024

1 Orthogonality

In your last two lectures, the concept of orthogonality was introduced and you learned some basic properties. The goal of this week's section will be to practice with the concept of orthogonality and introduce some unifying terminology. We will rehash many of the results from the lecture notes (but it is good to see these multiple times). We will work with subspaces of $\mathbb{R}^n$, but everything presented here holds more generally for a finite-dimensional vector space over $\mathbb{C}$ equipped with a Hermitian inner product.

Recall that $v, w \in \mathbb{R}^n$ are said to be **orthogonal** if $v \cdot w = 0$. Given an arbitrary $x \in \mathbb{R}^n$ and a subspace $V \subseteq \mathbb{R}^n$, we want to show that x can be written as the sum of an element of V and an element orthogonal to every element in V. We can think of this as writing x as the sum of its orthogonal projection onto V and the vector along which we project. Let us introduce some terminology to make this idea precise.

Suppose V and W are subspaces of $\mathbb{R}^n$. We define the sum of the subspaces as $V + W = \{v + w \mid v \in V, w \in W\}$. We say that $\mathbb{R}^n$ is the **direct sum** of V and W, and we write $\mathbb{R}^n = V \oplus W$, provided $\mathbb{R}^n = V + W$ and $V \cap W = \{0\}$.

> (1) Quickly verify for yourself that $V + W$ is a subspace of $\mathbb{R}^n$.
>
> (2) Prove that $\mathbb{R}^n = V \oplus W$ if and only if each element of $\mathbb{R}^n$ can be written *uniquely* in the form $v + w$, where $v \in V$ and $w \in W$.

Let V be a subspace of $\mathbb{R}^n$. We define the **orthogonal complement** of V as $V^\perp = \{x \in \mathbb{R}^n \mid x \cdot v = 0, \forall v \in V\}$. Our goal will be show that for any subspace $V \subseteq \mathbb{R}^n$, $V \oplus V^\perp = \mathbb{R}^n$ (compare Theorem 5.13 in the lecture notes). We will also establish a basic fact about dimensions of direct sums. This takes a little bit of work, so I'll break it down into smaller steps.

(1) Prove that $V^\perp$ is a subspace of $\mathbb{R}^n$ for any subspace $V \subseteq \mathbb{R}^n$.

(2) Prove that $V \cap V^\perp = \{0\}$. Therefore, it just remains to show that every element of $\mathbb{R}^n$ has a decomposition as the sum of an element of V and one of $V^\perp$ to establish $\mathbb{R}^n = V \oplus V^\perp$.

(3) By the Gram-Schmidt orthonormalization process (which you'll learn soon), every subspace of $\mathbb{R}^n$ has an orthonormal basis. Therefore, $V \subseteq \mathbb{R}^n$ has some orthonormal basis $v_1, \ldots, v_k$.

Define $P_V : \mathbb{R}^n \to \mathbb{R}^n$ by $P_V(x) = \sum_{i=1}^{k} (x \cdot v_i) v_i$. Given $x \in \mathbb{R}^n$, we can use the trick of adding and subtracting the same thing to obtain the decomposition $x = (x - P_V(x)) + P_V(x)$. Show that $P_V(x) \in V$ and $x - P_V(x) \in V^\perp$, which establishes $\mathbb{R}^n = V \oplus V^\perp$.

(4) Show that P_V is linear and satisfies that $P_V \circ P_V = P_V$.

(5) In general, if $\mathbb{R}^n = V \oplus W$, then $n = \dim(V) + \dim(W)$. Start with bases for V and W and use them to construct a basis for $\mathbb{R}^n$ to establish this result. It follows that $n = \dim(V) + \dim(V^\perp)$ for any subspace $V \subseteq \mathbb{R}^n$.

2 Projections

Let V be a vector space and $T : V \to V$ a linear map. We say that T is a **projection** if $T^2 = T$. Intuitively, we can think of T as projecting onto its image. After we apply T once, the resulting vector is already in the image, so a second application does not change the result. If you like fancy words, projections $V \to V$ are the *idempotent* elements of the ring $\mathrm{End}(V)$ of linear endomorphisms on a vector space.

Our previous discuss established that the orthogonal projection operator P_V is a projection in this abstract sense.

(1) Show that if $T : V \to V$ is a projection, then $V = \text{im}(T) \oplus \ker(T)$. Hint: use the trick of adding and subtracting the same thing that we used in (3) above.

(2) Show that if $A \sim B$, then $\text{Tr}(A) = \text{Tr}(B)$. It is easiest to first show $\text{Tr}(AB) = \text{Tr}(BA)$, and then apply this to the case of similar matrices.

Given any linear map T on a finite-dimensional vector space V, we can take its matrix representation $[T]_\beta$ with respect to any basis β. The change of basis formula tells us that $[T]_\beta \sim [T]_\alpha$ for any two bases α and β of V. This means that $[T]_\beta = P[T]_\alpha P^{-1}$ for some invertible matrix P.

Therefore, the trace of the matrix representation of any linear transformation is the same, so we can define this common value as the trace of the linear transformation.

(3) Show that if T is a projection then $\text{Tr}(T) = \dim(\text{im}(T))$.

Hint: use the decomposition $V = \text{im}(T) \oplus \ker(T)$ and compute the matrix representation with respect to the basis for $\text{im}(T) \oplus \ker(T)$ that we found in (5) of the section on orthogonality.

AS.110.201 Linear Algebra Week 7 Notes

Ben Marlin

October 2024

Abstract Vector Spaces

Recall that an abstract vector space consists of a set of "vectors" along with operations of vector addition and scalar multiplication that satisfy certain axioms. The axioms that these operations satisfy force vector spaces to behave similarly to $\mathbb{R}^n$. If you take a course on algebra, you'll see that a vector space can be thought of as an abelian group along with scalar multiplication (this is how I remember the definition, for example). Let us go over some common examples of vector spaces.

Example 1: The set of $m \times n$ matrices over $\mathbb{R}$, denoted $\mathrm{Mat}_{m\times n}(\mathbb{R})$ forms a vector space. NB: When I say that this set forms a vector space, what I mean is that this set along with certain canonical addition and scalar multiplication operations forms a vector space. There may be multiple ways to equip a set with a vector space structure, but most of the time it is obvious what choice is intended.

Your job is to describe the canonical addition and scalar multiplication operations on $\mathrm{Mat}_{m\times n}(\mathbb{R})$ and prove that they fulfill the axioms of a vector space. What is the dimension of $\mathrm{Mat}_{m\times n}(\mathbb{R})$? Exhibit an explicit basis for $\mathrm{Mat}_{m\times n}(\mathbb{R})$. Hint: there is a reason that Professor Riehl uses the alternative notation $\mathbb{R}^{m\times n}$ for $\mathrm{Mat}_{m\times n}(\mathbb{R})$.

Example 2: Let V be a vector space over $\mathbb{R}$. Recall that a subspace of V is a subset $W \subseteq V$ which is closed under linear combinations. This means that $aw_1 + bw_2 \in W$ for $w_1, w_2 \in W$ and $a, b \in \mathbb{R}$.

Given a finite collection of vectors $\{v_1, \dots, v_n\} \subseteq V$, we define the *span* of the vectors as the set of linear combinations $\{\sum_{i=1}^n a_i v_i \mid a_i \in \mathbb{R}, 1 \le i \le n\}$. Show that the span of such a collection of vectors forms a subspace of V.

Next, suppose that W_1 and W_2 are subspaces of V. Show that their intersection $W_1 \cap W_2$ is a subspace of V.

Example 3: The complex numbers $\mathbb{C}$ form a vector space over $\mathbb{R}$. Recall that $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$. Describe the addition and scalar multiplication operations on $\mathbb{C}$. What is the dimension of $\mathbb{C}$ as a vector space over $\mathbb{R}$?

Example 4: The set $\mathbb{Q}(\sqrt{2}) = \{a + \sqrt{2}b \mid a, b \in \mathbb{Q}\}$ forms a vector space over $\mathbb{Q}$. Find a basis and give the dimension of $\mathbb{Q}(\sqrt{2})$. This is an example of a *number field*, which are the central objects of study in algebraic number theory.

Example 5: The set $C([a, b])$ of continuous functions $[a, b] \to \mathbb{R}$ form a vector space over $\mathbb{R}$. Show that $C([a, b])$ is infinite dimensional. Hint: Find linearly independent polynomials on $[a, b]$.

Example 6: This example covers a topic more advanced than all of the others, but I think it is a fun example of how varied the structure of vector spaces can be. It's definitely beyond the scope of the course, but it might be interesting to those of you interested in higher level math.

The set of all polynomials over $\mathbb{R}$ with variable x, denoted $\mathbb{R}[x]$, is an infinite-dimensional vector space over $\mathbb{R}$. However, if we restrict our polynomials to have degree at most n, then we obtain an $n + 1$-dimensional vector space. It turns out that vector spaces arising from polynomials are extremely important in mathematics. An additional important construction is the quotient space $\mathbb{R}[x]/(p(x))$, where $p(x) \in \mathbb{R}[x]$. We can think of this space as being the set of polynomials over $\mathbb{R}$ where two polynomials are considered to be equivalent if their difference is a multiple of $p(x)$ by another polynomial $q(x)$.

For example, we can consider $\mathbb{R}[x]/(x^2 + 1)$, which turns out to be the same (in some technical sense) as the complex numbers $\mathbb{C}$. Let me briefly explain this connection.

Let $f(x) \in \mathbb{R}[x]$. We can use polynomial long division to divide $f(x)$ by $x^2 + 1$. This gives $f(x) = (x^2 + 1)q(x) + r(x)$, where $q(x)$ is just some polynomial and $r(x)$ (the remainder) is a polynomial of degree less than two.

It follows that $f(x) - r(x) = (x^2 + 1)q(x)$, so $f(x)$ is equivalent to the linear polynomial $r(x)$. Further, it is clear to x^2+1 is a multiple of itself, so $(x^2+1)-0$ is a multiple of $x^2 + 1$ as well. This means that $x^2 + 1$ is equivalent to 0, or in other words, $x^2 = -1$.

Therefore, $\mathbb{R}[x]/(x^2+1)$ consists of all linear polynomials $a+bx$, where $a, b \in \mathbb{R}$ and $x^2 = -1$. The complex numbers consist of all $a + bi$, where $a, b \in \mathbb{R}$ and $i^2 = -1$. Therefore, $\mathbb{R}[x]/(x^2 + 1)$ is the same as $\mathbb{C}$ by simply changing our variable from x to i.

The technical terminology would be to say that $\mathbb{R}[x]/(x^2 + 1)$ is isomorphic to $\mathbb{C}$ as fields.

AS.110.201 Linear Algebra Week 6 Notes

Ben Marlin

October 2024

Midterm Solutions

These solutions are a supplement to my recitation section. I focus on the concepts involved in solving the problems more than the actual midterm problems in these notes. These notes are **not** meant to be a template for what your solutions on exams should look like. I leave out many explicit details and calculations that should be written in detail on an exam (because these will be written on the blackboard in the actual recitation section!)

Problem 1: There are several ways to solve a system of linear equations of the form $Ax = b$. Because our coefficient matrix A is already close to being in row reduced echelon form, the easiest way to find the set of solutions is to compute the RREF of the **augmented matrix of A**, and write the set of solutions parametrically, with the parameters varying over the free variables. Remember, when solving $Ax = b$ we must compute the RREF of the augmented matrix. If we simply compute RREF(A) and append our right-hand-side vector, we end up with a different system of equations (see if you can come up with an example to verify that this is true).

Let me describe another way to solve systems of the form $Ax = b$, $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, $b \in \mathbb{R}^m$, which uses the concepts of bases and subspaces. First, observe that the set of solutions to $Ax = b$ is an *affine subspace* of $\mathbb{R}^n$, that is, some translate of a genuine subspace of $\mathbb{R}^n$. Indeed, suppose that $Ay = b$ and let z be an arbitrary solution to our system. Then, $y - z \in \ker(A)$, hence $z \in y + \ker(A)$. Therefore, to find all of the solutions to $Ax = b$, we need to find all of the solutions to $Ax = 0$ and a single particular solution to $Ax = b$. Because $\ker(A)$ is a subspace of $\mathbb{R}^n$, it has a basis, so we can find the kernel of A by finding a basis and taking linear combinations. As we progress in this course and become more comfortable with bases and subspaces, you will see that computing bases for the image and kernel of a linear transformation are intimately related to row reduction.

Problem 2: Recall that a function $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear if and only if $T(ax + by) = aT(x) + bT(y)$ **for all** $a, b \in \mathbb{R}$ and $x, y \in \mathbb{R}^n$. Many students only verified that linearity holds for a particular choice of scalars and vectors. To establish linearity, we must show the linearity identity holds for arbitrary vectors and scalars. Another way to show T is linear is to show that there exists a matrix $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ such that $T(x) = Ax$, for all $x \in \mathbb{R}^n$.

This gives us two ways to approach this problem. First, we need to make sure that we understand what the function in question actually looks like concretely before attempting to verify its linearity. The function $\mathbb{R}^4 \to \mathbb{R}$ sends a vector $v \in \mathbb{R}^4$ to its **dot product** with the column vector $(3, 0, 1, -2)^T$. Recall that the dot product of two vectors is obtained by multiplying entries coordinate-wise and then summing everything up.

Therefore, given an arbitrary $v = (v_1, v_2, v_3, v_4)^T \in \mathbb{R}^4$, its dot product with $(3, 0, 1, -2)^T$ is $3v_1 + 0v_2 + v_3 - 2v_4 = 3v_1 + v_3 - 2v_4$.

The easiest way to see this function is linear is by observing that taking the dot product of two column vectors $x, y \in \mathbb{R}^n$ is the same thing as computing the product $x^T y$. Therefore, the function f in question is represented by the matrix $(3, 0, 1, -2)$ in the sense that $f(x) = (3, 0, 1, -2)x$ for all $x \in \mathbb{R}^4$.

Another way to see the function is linear is by verifying the linearity identity. This follows immediately by the properties of the dot product. Indeed, $x \cdot (ay_1 + by_2) = a(x \cdot y_1) + b(x \cdot y_2)$.

Problem 3: Recall that by definition the image of a matrix $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ is the set $\mathrm{im}(A) = \{Ax \mid x \in \mathbb{R}^n\}$. One can verify that this is the same as the set $\mathrm{span}\{a_1, \ldots, a_n\}$, where a_i is the vector defining the ith column of A.

Because the second row of the matrix A in question consists entirely of zeros, any element of the image of A must have a zero as its second coordinate. This means that $(1, 1, 1, 1)^T$ cannot be in the image because $0 \neq 1$.

Most people did quite well on this question. Some of the common mistakes were confusing the image and kernel of a matrix and incorrectly using row reduction to solve the system $Ax = b$. Remember, you must row reduce the **augmented matrix**, not the coefficient matrix, when solving a system by row reduction.

Problem 4: We are asked to find the kernel of a 2×2 matrix. Recall that for a square matrix A, $\ker(A) = \{0\}$ (we say A has trivial kernel) if and only if A is an invertible matrix if and only if the determinant of A is non-zero. Therefore, if our matrix in question is invertible, we know that it must have trivial kernel.

Last section, I went over many different ways to determine whether a square matrix A is invertible. Let me list them here for your convenience:

1. Explicitly exhibit some matrix B such that $AB = BA = I$.
2. Show that the determinant of A is non-zero.
3. Show that $\text{RREF}(A) = I$.
4. Show that $\ker(A) = \{0\}$.
5. Show that A defines an injective map.
6. Show that A defines a surjective map.
7. Show that A defines a bijective map.
8. Show that A defines an invertible map (in the sense of functions).
9. Show that the columns of A are linearly independent.
10. Show that the rows of A are linearly independent.
11. Show that A has full rank.
12. Show that the image of A is the entire codomain (A does not collapse our domain into a lower dimensional space).

I think that it is a **really good exercise** to try to think about why all of these conditions are equivalent. Some implications such as $9 \implies 10$ or those saying that injectivity or surjectivity on their own imply bijectivity are non-trivial. Several implications are just directly rewriting things in different language.

We can easily compute the determinant and verify it is non-zero. Therefore, A is invertible and its kernel must be just the zero vector.

Problem 5: This was the hardest question on the midterm in my opinion. It involves a solid understanding of the concepts of function composition and invertibility.

A common mistake was composing the functions in the wrong order. Recall that the function f in $g \circ f$ "acts first," however, when presented with composite function described diagrammatically, the leftmost function "acts first."

For example, if we have $f : A \to B$ and $g : B \to C$, we obtain the composite $g \circ f : A \to C$.

We can describe this information diagramatically as:

$g \circ f = A \xrightarrow{f} B \xrightarrow{g} C$.

Notice that in the diagram f appears on the left while in the standard notation it appears on the right. If $f : A \to B$ and $g : B \to A$, both of the compositions

$f \circ g$ and $g \circ f$ are well-defined, so we need to be careful that we do them in the right order.

For our specific problem, we are given $T_A : \mathbb{R}^3 \to \mathbb{R}$ defined by the row vector $(2, 3, 4)$. This is the function that sends $x = (x_1, x_2, x_3) \mapsto 2x_1 + 3x_2 + 4x_3$. Our function $T_B : \mathbb{R} \to \mathbb{R}^3$ is defined by the column vector $(7, -6, 5)^T$. This means that $T_B(\alpha) = (7\alpha, -6\alpha, 5\alpha)^T$ for any $\alpha \in \mathbb{R}$.

The composite function is a map $\mathbb{R}^3 \to \mathbb{R}^3$, so when we write down its matrix, we better end up with a 3×3 matrix.

Let $x = (x_1, x_2, x_3)^T \in \mathbb{R}^3$. As before, $T_A(x) = 2x_1 + 3x_2 + 4x_3$. This is the scalar that we will now feed into T_B.

Observe that:

$$T_B(T_A(x)) = (7(2x_1 + 3x_2 + 4x_3), -6(2x_1 + 3x_2 + 4x_3), 5(2x_1 + 3x_2 + 4x_3))^T.$$

Simplifying:

$$T_B(T_A(x)) = (14x_1 + 21x_2 + 28x_3, -12x_1 - 18x_2 - 24x_3, 10x_1 + 15x_2 + 20x_3)^T.$$

Recall that the matrix of $T = T_B \circ T_A$ has ith column $T(e_i)$, where e_i denotes the ith standard basis vector.

This means that the first column of A is $(14, -12, 10)^T$, its second is $(21, -18, 15)^T$, and its third is $(28, -24, 20)^T$. By inspection, we can see that the columns are **not** linearly independent, hence the matrix of T cannot be invertible.

That was a lot of work! Here is a far easier solution. By the rank-nullity theorem, any linear map $\mathbb{R}^n \to \mathbb{R}^m$ has non-trivial kernel if $m < n$. Therefore, $T_A : \mathbb{R}^3 \to \mathbb{R}$ has non-trivial kernel. Any element of the kernel of T_A is also in the kernel of $T_B \circ T_A$ (verify this if it is not clear). Therefore, $T = T_B \circ T_A$ has non-trivial kernel, so it cannot be invertible.

AS.110.201 Linear Algebra Week 5 Notes

Ben Marlin

September 2024

Practice Midterm Solutions

Problem 1: There is no solution to the system of linear equations defined by the augmented matrix because the standard basis e_4 is not a linear combination of elements of the set $\{e_i\}_{i=1}^3$. Phrased differently, e_4 is not in the image of the coefficient matrix. We can also immediately detect the inconsistency of the system by observing that there is a row of zeros in the coefficient matrix that corresponds to a nonzero entry in the augmented vector.

Problem 2: Recall the general form of a matrix that reflects vectors across the line L_θ that has counterclockwise angle θ. In our case, we are first reflecting across $L_{\frac{\pi}{4}}$ and then across $L_{\frac{\pi}{2}}$. This corresponds to the composition of linear transformations, which we know can be represented by matrix multiplication.

Basic trigonometry tells us that $Ref(\frac{\pi}{4})$ has its first column as e_2 and its second as e_1. Similarly, $Ref(\frac{\pi}{2})$ has first column $-e_1$ and second column e_2.

Before calculating the product of these matrices, let us check that these matrices actually do what we want them to do. Inspection of $Ref(\frac{\pi}{4})$ tells us that when we multiply it by a vector it just swaps the x and y coordinates, which makes sense for a reflection over $y = x$. Inspection of $Ref(\frac{\pi}{2})$ tells us that when we multiply it by a vector, it negates the x coordinate and leaves the y coordinate unchanged, which makes sense for a reflection over $x = 0$.

Matrix multiplication yields that the matrix of the linear transformation defining reflection over $y = x$ followed by reflection over $x = 0$ has first column e_2 and second column $-e_1$.

Note that we could have avoided touching matrices at all by recognizing that our transformation should swap coordinates and then negate the first one. Applying this to the standard basis vectors gives the same matrix.

Problem 3: The since matrix multiplication defines a linear map, we can obtain the image of $e_4 - \frac{1}{2}e_1$ under the matrix A as the difference between the fourth column and $\frac{1}{2}$ times the second column of A.

This is $(3,0,2,5) - \frac{1}{2}(6,-2,4,0) = (0,1,0,5)$ by definition of vector addition and scalar multiplication.

Another (albeit less efficient) way to do this problem is to write down $e_4 - \frac{1}{2}e_2$ explicitly and then multiply it by A. We obtain $e_4 - \frac{1}{2}e_1 = (-\frac{1}{2},0,0,1)$ and carrying out the matrix multiplication yields the same answer as above.

Problem 4: We want to determine if $e_3 - e_2$ is in the span of the vectors $e_1 + e_2$ and $e_1 + e_3$. It is easiest to first write our vectors explicitly in coordinates:

$e_3 - e_2 = (0,-1,1)$, $e_1 + e_2 = (1,1,0)$, and $e_1 + e_3 = (1,0,1)$.

We then see that $(0,-1,1) = (1,0,1) - (1,1,0)$, so recalling the definition of span, we see that it is true that $e_2 - e_3$ lies in the span of $e_1 + e_2$ and $e_1 + e_3$.

Another way to solve this problem is to do algebra with the standard basis vectors. Indeed, $e_3 - e_2 = (e_1 + e_3) - (e_1 + e_2)$, so we arrive at the same conclusion.

Problem 5: In general, a diagonal (or even upper triangular) matrix is invertible if and only if all of its diagonal entries are nonzero because the determinant of such a matrix is the product of its diagonal entries. This is a useful fact and tells us our matrix is invertible, but completely overkill for this problem.

Instead, we can explicitly exhibit an inverse matrix by considering the matrix whose diagonal elements are the reciprocals of those of our original matrix A. So, define $B = \text{diag}(\frac{1}{2}, \frac{1}{3}, -1, \frac{1}{4})$. It is easy to verify that $AB = BA = I_4$.

We could also row reduce our matrix and observe that $RREF(A) = I_4$, so A is invertible.

We could also remember that a square matrix is invertible if and only if its kernel is trivial if and only if it defines a surjective linear map. It is obvious that the only solution to $Ax = 0$ is the zero vector, so A is invertible.

We can show A defines a surjective map by noting that given any vector $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$, $Ay = x$, where $y = (\frac{1}{2}x_1, \frac{1}{3}x_2, -1x_3, \frac{1}{4}x_4)$.

AS.110.201 Linear Algebra Week 4 Notes

Ben Marlin

September 2024

1 Invertible functions

We say that a function $f : A \to B$ is **invertible** if there exists an inverse function $g : B \to A$ such that $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$.

<u>Exercise</u>: Show that the inverse of an invertible function is unique. That is, show that if there exists $g_1, g_2 : B \to A$ such that $g_i \circ f = \mathrm{id}_A$ and $f \circ g_i = \mathrm{id}_B$ for $1 \leq i \leq 2$, then $g_1 = g_2$.

The previous exercise tells us that we can speak of *the* inverse of a function $f : A \to B$, which we denote $f^{-1} : B \to A$.

Let us practice with the notion of invertibility by doing some concrete examples:

1. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$. Is f invertible? If so, find f^{-1}. What if we modify the domain and codomain and consider $\bar{f} : \mathbb{R}^+ \to \mathbb{R}^+$ with $x \mapsto x^2$?

2. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x + 1$. Graph the function f and determine if it is invertible. If f is invertible, find its inverse f^{-1}.

3. Consider the linear function $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(e_1) = (1, 0)$ and $T(e_2) = (4, 3)$ (by last week's notes, this uniquely defines a linear function!). Is T invertible? If so, find T^{-1}. Hint: find the matrix $[T]$ representing T and determine whether it is invertible.

It is not always obvious whether a function f is invertible or not. Indeed, from the given definition of invertibility, we need to construct an inverse function or prove that such a function cannot exist. This definition is somewhat external to our starting function f. The following section will provide a different characterization of invertible functions.

We say that a function $f : A \to B$ is **injective** provided that, for all $a, b \in A$, if $a \neq b$ then $f(a) \neq f(b)$. In other words, f sends distinct elements to distinct values. Such functions are also called one-to-one functions sometimes. In high school, you likely learned about the horizontal line test, which says that a function $f : \mathbb{R} \to \mathbb{R}$ is injective when any horizontal line intersects the graph of f in at most one place.

Exercise: Prove that the composition of two injective functions is injective. In other words, show that if $f : A \to B$ and $g : B \to C$ are injective, then $g \circ f : A \to C$ is injective.

We say that a function $f : A \to B$ is **surjective** if for every $b \in B$ there exists some $a \in A$ such that $f(a) = b$. We can think of such a function as one that "hits" everything in the codomain in the sense that the range of f is all of B. We also call surjective functions onto (this comes from thinking of surjections as mapping *onto* the entire codomain). Note that the codomain we be restricted to force a function to be surjective (look at the first example we did above).

Exercise: Come up with an example of a surjective function and a non-surjective function. You should write the function in the form $f : A \to B$, $x \mapsto f(x)$ to emphasize the domain and codomain in the definition.

We now have the language to give an equivalent definition of an invertible function. We say that a function $f : A \to B$ is **bijective** if f is both injective and surjective. By considering the definitions of injective and surjective, we see that f is bijective provided that for each $b \in B$ there exists a *unique* $a_b \in A$ such that $f(a_b) = b$.

Exercise: Prove that a function $f : A \to B$ is invertible if and only if f is bijective.

Because linear transformations (and matrices) can be viewed as just special types of functions, this tells us that we can determine whether a matrix is invertible by looking at its associated linear transformation and checking if it is bijective.

AS.110.201 Linear Algebra Week 3 Notes

Ben Marlin

September 2024

1 Understanding Linearity

Recall that a function $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear if $T(ax+by) = aT(x)+bT(y)$ for all $a, b \in \mathbb{R}$ and $x, y \in \mathbb{R}^n$. Furthermore, any linear function can be represented by a matrix. More precisely, if $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then there exists a unique matrix $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ such that $T(x) = Ax$ for all $x \in \mathbb{R}^n$. Recall from your lectures that the matrix representing T is precisely the matrix whose ith column is $T(e_i)$, where e_i denotes the ith standard basis vector in $\mathbb{R}^n$.

Because any matrix $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ defines a linear map T_A by $T_A(x) = Ax$, linear transformations and matrices are "the same thing." Indeed, given a linear transformation, I can construct its associated matrix, and given a matrix, I can construct its associated linear transformation. These constructions are inverses of one another.

<u>Exercise</u>: Verify that T_A as defined above is a linear transformation.

Let $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ be a matrix with columns $x_1, \dots, x_n$. Observe that when we multiply A by a vector y we obtain $Ay = \sum_{i=1}^n y_i x_i$, where $y = (y_1, \dots, y_n)$. Therefore, multiplying a matrix by a vector is determined by multiplying its columns by the corresponding coordinates of the vector and adding everything together. Similarly, a linear transformation T is determined by its values on the standard basis $\{e_1, \dots, e_n\}$.

<u>Exercise</u>: Explain to a peer why linear transformations are determined by their values on the standard basis.

<u>Exercise</u>: Let us make the previous exercise more formal. Suppose that the functions $T_1, T_2 : \mathbb{R}^n \to \mathbb{R}^m$ are linear transformations. Suppose that $T_1(e_i) = T_2(e_i)$ for all $1 \le i \le n$. Show that $T_1(x) = T_2(x)$ for all $x \in \mathbb{R}^n$.

<u>Exercise (Challenge!)</u>: Let $x_1, \dots, x_n \in \mathbb{R}^m$. Show that there exists a unique linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that $T(e_i) = x_i$ for $1 \le i \le n$. More generally, if $\beta_1, \dots, \beta_n$ is any basis of $\mathbb{R}^n$, then there exists a unique linear

transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that $T(\beta_i) = T(x_i)$. If you take an abstract algebra course in the future, you'll recognize this exercise is stating the universal property of free modules. If this exercise makes no sense right now, try revisiting it in a couple weeks as preparation for the first midterm.

NB: The above exercise show that it makes sense to talk about *the* linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that $T(e_i) = x_i$.

2 Finding the matrix of a linear function

Suppose that I tell you T is a linear function $\mathbb{R}^2 \to \mathbb{R}^2$ such that $T(e_1) = (2, 0)$ and $T(e_2) = (0, 1)$. What should its corresponding matrix look like? Well, in lecture Professor Riehl proved that the corresponding matrix should have its first column as $(2, 0)$ and its second as $(0, 1)$.

Note: when I write $(x_1, \ldots, x_n)$, I really mean the column vector obtained by writing the tuple vertically. It is easier to type the ordered tuples.

This is pretty straightforward. If I give you the values of $T(e_i)$, these just become the columns of the corresponding matrix. Let us make things a little bit harder.

<u>Exercise</u>: Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear transformation that satisfies $T(2, 0) = (1, 0)$ and $T(2, 1) = (3, 4)$. Find the matrix corresponding to T.

Hint: express the standard basis vectors in terms of $(2, 0)$ and $(2, 1)$ and then use the linearity of T to find $T(e_1)$ and $T(e_2)$.

<u>Exercise</u>: Suppose that $T : \mathbb{R}^3 \to \mathbb{R}^3$ is the linear transformation that satisfies $T(1, 0, 0) = (0, 1, 4)$, $T(1, 1, 0) = (2, 3, 1)$, and $T(0, 0, 2) = (1, 0, 4)$. Find the matrix corresponding to T.

<u>Exercise</u>: Does it make sense to talk about *the unique* linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T(1, 0) = (4, 1)$ and $T(2, 0) = (0, 1)$? Why or why not? How about *the unique* linear transformation such that $T(1, 0) = (2, 1)$ and $T(2, 3) = (5, 60012)$?

AS.110.201 Linear Algebra Week 2 Notes

Ben Marlin

September 2024

1 Dimension

The notion of the dimension of a vector space is fundamental to linear algebra. We have not yet formally introduced this concept, but we can still think about what the "right" definition should be. For example, we all agree that a sheet of paper is a 2-dimensional object and the world around us is 3-dimensional space. In mathematical language, we are essentially saying that $\mathbb{R}^2$ is 2-dimensional and $\mathbb{R}^3$ is 3-dimensional. This makes sense because it takes 2 coordinates to uniquely specify a point in the plane and 3 coordinates to uniquely specify a point in $\mathbb{R}^3$.

Although it might be tempting to think that any collection of points in, say, $\mathbb{R}^3$ should be regarded as a 3-dimensional space because to specify any point in $\mathbb{R}^3$ requires a value for x, y, and z, this does not capture our intuition that lines are 1-dimensional and planes are 2-dimensional.

Consider the line $L = \{t(2, 3, 4) \in \mathbb{R}^3 \mid t \in \mathbb{R}\}$. While each point on this line lies in $\mathbb{R}^3$, once we understand that the line is just multiples of the vector $(2, 3, 4)$, we can actually uniquely specify any point on the line by giving the scale factor t. Therefore, if we are smart about how we view the line, we only need to specify 1 value to give points on it, so the line should be a 1-dimensional object.

Let us try a more complicated example. Consider the equation $x + 2y + 3z = 0$. What sort of geometric object does the set of solutions to this equation define? Compute the dimension of the set of solutions. Find a collection of vectors such that every solution to the above equation can be written uniquely as a linear combination of the chosen collection. What does this collection have to do with the dimension?

(Recall that if $\{v_1, \dots, v_n\}$ is a collection of vectors in $\mathbb{R}^m$, then a linear combination of the vectors $v_1, \dots, v_n$ is a vector of the form $\sum_{i=1}^n a_i v_i$, where $a_i \in \mathbb{R}$. For example, if $v_1 = (1, 2)$ and $v_2 = (3, 4)$, then $2(1, 2) + 3(3, 4) = (11, 16)$ is a linear combination).

We will see later on that a collection of vectors β such that every vector in a vector space V (think $\mathbb{R}^m$ for now) can be written as a unique linear combination of elements of β is called a *basis* for V, and every vector space has a basis.

Next, let us understand how the dimension of the solutions to an equation of the form $Ax = b$ is related to the dimension of the solutions to an equation of the form $Ax = 0$.

Show that every solution to $Ax = b$ can be written as the sum of a fixed vector v and a solution to $Ax = 0$. We call the set of solutions to $Ax = 0$ the *kernel* of A, so we can restate this as saying that any solution to $Ax = b$ can be written as the sum of a fixed vector v and an element of the kernel of A.

Explain to your group how to compute the kernel of a matrix using Gauss-Jordan elimination (row reduction).

---

2 Linear transformations

We say that a function $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear provided that $T(ax + by) = aT(x) + bT(y)$ for all vectors $x, y \in \mathbb{R}^n$ and scalars $a, b \in \mathbb{R}$. You will see later on that any linear transformation can be represented by a matrix, which makes them particularly nice objects to work with.

The key insight to the connection between linear transformations and matrices is that a linear transformation is entirely determined by its values on a basis. An example of a basis for $\mathbb{R}^n$ is the standard basis, consisting of the vectors $e_1, \ldots, e_n$, where e_i is the vector whose ith component is 1 and every other component is 0.

Explain how every vector $v \in \mathbb{R}^n$ can be written in terms of the standard basis.

Explain how knowing the value of a linear map T on the standard basis allows us to know the value of T for every vector in $\mathbb{R}^n$. Put differently, if I give you the vectors $T(e_1), \ldots, T(e_n)$, how can you figure out the value of $T(v)$, where v is an arbitrary vector in $\mathbb{R}^n$.

If all of the abstraction and notation is confusing, try writing down simple examples for a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}$. For example, consider the <u>linear</u> transformation $T : \mathbb{R}^2 \to \mathbb{R}$ that satisfies $T(1, 0) = 2$ and $T(0, 1) = 1$. Find the value of $T(3, 4)$ using linearity. Can you find a 1×2 matrix A such that $Ax = T(x)$ for all $x \in \mathbb{R}^2$?

AS.110.201 Linear Algebra Week 1 Notes

Ben Marlin

August 2024

1 Introduction

Hello everyone, I'm Ben Marlin and I will be your TA for linear algebra this semester. I'm a fourth year undergraduate pursing a BA/MA in mathematics and a BA in philosophy. You can reach me by email at bmarlin1@jh.edu. Please feel free to reach out with any questions concerning course content or logistics and I will get back to you as soon as possible. I will be holding office hours through the Learning Den from 7-9pm in Gilman 313. Bring any questions about linear algebra (or math in general), and I'll be more than happy to help!

2 Systems of linear equations

Linear equations and the maps that define them are of central importance to linear algebra. In this course, you will learn about the properties of such systems and how to systematically solve them. Before giving a formal definition, let me give a few examples.

1. Consider the system of equations:

$$\begin{aligned} 3x + 6y &= 3 \\ 3x + 7y &= 4 \end{aligned}$$

This is a system of linear equations with two equations and two "unknowns" (variables). We can solve it by "eliminating variables" (which you probably learned in high school). Subtracting the first equation from the second gives the system

$$\begin{aligned} 3x + 6y &= 3 \\ y &= 1 \end{aligned}$$

from which it is easy to see that $(x, y) = (-1, 1)$ is the unique solution to the system.

2. Consider the system of equations:

$$x + 2y + 3z = 7$$
$$3x + 4y + z = 5$$

This is a system of linear equations with two equations and three "unknowns." We can proceed similarly. Subtracting three times the first equation from the second equation yields

$$x + 2y + 3z = 7$$
$$-2y - 8z = -16$$

We can then divide the second equation by -2 to obtain

$$x + 2y + 3z = 7$$
$$y + 4z = 8$$

We will now choose a value for z and use this to find the values of x and y such that (x, y, z) is a solution to the system. For simplicity, choose $z = 0$. We then see that

$$x + 2y = 7$$
$$y = 8$$

so by substituting in $y = 8$, we find that $(-9, 8, 0)$ is a solution. Note that this solution is not unique! By choosing $z = 1$ and arguing as before we see $(-4, 4, 1)$ is another solution. In fact, since both $(-9, 8, 0)$ and $(-4, 4, 1)$ satisfy the system, if we subtract the latter from the former we obtain $(-5, 4, -1)$, which is a solution to the system

$$x + 2y + 3z = 0$$
$$3x + 4y + z = 0$$

(think about why the above statement is true). Therefore, if $\alpha \in \mathbb{R}$ then $(-9, 8, 0) + \alpha(-5, 4, -1)$ is a solution to our system (think about why this is true). In particular, the system of equations has infinitely many solutions.

Exercise 1: Explain to a peer conceptually why the first system of equations has a unique solution and the second has infinitely many solutions.

Exercise 2: Explain to a peer why our method of "eliminating variables" does not change the set of solutions to the linear systems. This fact is crucial to the method of Gauss-Jordan elimination, and you will later see that it is a consequence of the invertibility of so-called elementary matrices.

Exercise 3: Does every system of linear equations with 2 equations and 2 unknowns have a solution? How about n equations and n unknowns, where n is a positive integer?

Exercise 4: Write down a system of equations with 3 unknowns and 4 equations and find a solution to it.

2.1 Reframing the problem using matrices

I will freely use standard mathematical notation in this section so that you can practice reading and writing mathematical prose. Try to figure out what the notation means by context, and please ask me if it is unclear.

Let us now treat systems of linear equations in a slightly more formal way. A linear system of equations is a set of equations of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_n \end{aligned}$$

where for $1 \leq i \leq m$ and $1 \leq j \leq n$, a_{ij} and b_i are real numbers. More generally, we can take our coefficients to be elements of a field (you'll learn what this means later on).

Using matrices and vectors, a linear system with m equations and n unknowns can be written as $Ax = b$, where $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ is an $m \times n$ matrix, and $x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ are vectors. We can view $(x_1, \ldots, x_n)^T = x \in \mathbb{R}^n$ as a vector of unknowns and $b = (b_1, \ldots, b_n)^T$ as a vector of the right-hand-side coefficients. One can check that the equation $Ax = b$ with $A = (a_{ij})_{m\times n}$ defines the same system of linear equations as our above system.

By using the matrix form $Ax = b$ of a system of linear equations, we can use the theory of matrices and linear transformations (what linear algebra is all about!) to determine how many solutions a given system has and then find them.

Exercise 5: Suppose there exists a solution to the system of linear equations given by $Ax = b$. Prove that the solution is unique if and only if there is a unique solution to the system given by $Ax = 0$.

2.2 Preview of what is to come

Disclaimer: The material in this section is extra and entirely optional. It is written at a higher level than the previous sections and presupposes some familiarity with linear algebra. It may be useful to look at later.

Let me explain how we can use linear algebra to understand systems of linear equations given as $Ax = b$ for $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$. Suppose there exists some $y \in \mathbb{R}^n$ such that $Ay = b$. If x is another solution, then $Ax = Ay$, implying $A(x - y) = 0$. Therefore, given a particular solution y, all of the solutions $Ax = b$ arises as $y + w$, where w is a solution to $Ax = 0$. If we denote the set of all solutions to $Ax = 0$ by $\ker(A)$, then the set of all solutions to $Ax = b$ is given by the set $y + \ker(A) := \{y + w \in \mathbb{R}^n \mid w \in \ker(A)\}$.

Therefore, we have reduced the problem of finding all solutions $Ax = b$ to finding a single solution of $Ax = b$ and all of the solutions to $Ax = 0$. Both of these are possible by bringing A to its row reduced echelon form (also known as Gaussian elimination). Moreover, because $\ker(A)$ is a subspace of $\mathbb{R}^n$, it has a basis and we can write all of its elements as linear combinations of finitely many basis vectors. Since we can find a basis for $\ker(A)$ using row reduction, solving $Ax = b$ can be completely dealt with using row reduction.

Some other observations are in order. If $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ and $\ker(A) = 0$, there exists a unique solution to $Ax = b$, for any $b \in \mathbb{R}^n$. Indeed, by the rank-nullity theorem, a trivial kernel implies that A is invertible, so $x = A^{-1}b$ is the unique solution to the linear system. Furthermore, we can compute A^{-1} using Cramer's rule to give a precise solution using determinants.

We can also say some interesting things when A is not square. If $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ and $n > m$, then $\ker(A) \neq 0$. Indeed, if $\ker(A) = 0$ then A can be viewed as an injective linear map $\mathbb{R}^n \hookrightarrow \mathbb{R}^m$, which implies there is an n-dimensional subspace of $\mathbb{R}^m$. But this contradicts $n > m$ because subspaces of $\mathbb{R}^m$ have dimension $\leq m$. Therefore, presented with a system of equations with n unknowns and m equations, $n > m$, if a solution exists then there are infinitely many solutions.