Daily Update

Math 290: Elementary Linear Algebra

Fall 2018

Lecture 27: Tuesday, December 4. After reviewing the definitions of a linear transformation, and the kernel and range of a linear transformation, we recalled what is means for a linear transformation to be **one-to-one** and **onto**, which ensure that it is an **isomorphism**. If V and W are vector spaces, a linear transformation

$$T : V \to W$$

is *one-to-one* if whenever $T(\mathbf{v}_1) = T(\mathbf{v}_2)$, then $\mathbf{v}_1 = \mathbf{v}_2$. On the other hand, T is *onto* if given any $\mathbf{w} \in W$, there is some $\mathbf{v} \in V$ for which $T(\mathbf{v}) = \mathbf{w}$. Equivalently, T is one-to-one if the kernel of T is $\{\mathbf{0}\}$, and is onto if the image of T is W.

If there is an isomorphism from V to W (or W to V), we say that V and W are **isomorphic**. In fact, two vector spaces are isomorphic if and only if they have the same dimension!

We gave an example of a linear transformation that is an isomorphism.

We described a linear transformation that is a *projection*, and one given by *differentiation*, and calculated the image and range of each.

Next, we described the *eigenvalue problem*: If A is a $n \times n$ matrix and λ is a scalar, if

$$A\mathbf{x} = \lambda\mathbf{x}$$

for some nonzero $n \times 1$ column vector $\mathbf{x}$, then we say that $\mathbf{x}$ is an **eigenvalue** of A with corresponding **eigenvalue** λ.

We verified that given scalar-vector pairs are in fact eigenvectors and corresponding eigenvalues.

Lecture 26: Thursday, November 29. We started class today by reviewing the definition of a linear transformation

$$T : V \to W$$

between vector spaces V and W, and then defined the **kernel** of T as the set of element in the domain sent to zero

$$kernel(T) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\} \subseteq V$$

and the **range** of T as the set of all outputs

$$\text{range}(T) = \{T(\mathbf{v}) \mid \mathbf{v} \in V\} \subseteq W.$$

The kernel is a subspace of the domain V, and the range is a subspace of W. Moreover, the dimension of the kernel is called the **nullity** of T, and the dimension of the range is called the **rank** of T. (Notice the nomenclature overlaps with language we use for matrices!)

We stated some important properties that all linear transformations satisfy:

- $T(\mathbf{0}) = \mathbf{0}$
- For any $\mathbf{v} \in V$, $T(-\mathbf{v}) = -T(\mathbf{v})$

- For any $\mathbf{v}_1, \mathbf{v}_2 \in V$, $T(\mathbf{v}_1 - \mathbf{v}_2) = T(\mathbf{v}_1) - T(\mathbf{v}_2)$
- For any scalars c_i and vectors $\mathbf{v}_i \in V$, for $1 \le i \le n$,

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_nT(\mathbf{v}_n)$$

We found the kernel, range, nullity, and rank of the two "boring" linear transformations: the zero and identity transformations.

Next, we described one of the most important types of linear transformations, those defined by matrices: Given an $m \times n$ matrix, we checked that the function

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

given by $T(\mathbf{v}) = A\mathbf{v}$ (where $\mathbf{v} \in RR^n$ is a column vector, and so is $T(\mathbf{v}) \in RR^m$).

Using the example of the linear transformation

$$\begin{aligned} T :&\mathbb{R}^3 \to \mathbb{R}^3 \text{ given by} \\ &T(\mathbf{v}) = A\mathbf{v} \end{aligned}$$

where $A = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 0 & -4 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$.

We calculated that in this example, the kernel of T in $\mathbb{R}^3$ is the span of the vectors $\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and the range of T is the one-dimensional subspace of $\mathbb{R}^2$ spanned by $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$. Therefore, the rank of T is two, and the nullity is one; their sum is the number of columns of A, three!

We pointed out that rotation of vectors by a fixed angle θ about the origin in $\mathbb{R}^2$ is a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$, given by multiplication by the matrix

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

We found that its kernel is $\{\mathbf{0}\}$, so the nullity is zero, and the range is all of $\mathbb{R}^2$, so the rank is two.

This is an example of an **isomorphism**: a linear transformation $T : V \to W$ that is both one-to-one and onto:

- **One-to-one**: No two elements have the same output, or equivalently, if $T(\mathbf{v}_1) = T(\mathbf{v}_2)$, then $\mathbf{v}_1 = \mathbf{v}_2$.
- **Onto**: Every element of W is an output value; i.e., if $\mathbf{w} \in W$, then $T(\mathbf{v}) = \mathbf{w}$ for some $\mathbf{v} \in V$.

We will finish our initial discussion of linear transformations next time, and then start our final topic of the semester: **eigenvalues** and **eigenvectors**.

Lecture 25: Tuesday, November 27. Today we pieced together some work from last time to extend our equivalence of conditions for a square matrix of order n to be invertible:

$$\begin{aligned} A \text{ invertible} &\iff \det(A) \neq 0 \\ &\iff \operatorname{rank} A = n \\ &\iff \text{the } n \text{ rows of } A \text{ are linearly independent} \\ &\iff \text{the } n \text{ columns of } A \text{ are linearly independent.} \end{aligned}$$

Next, we defined a **linear transformation** between vector spaces V and W as a function

$$T : V \to W$$

(i.e., V is the domain, and the output values of T lie in W) that satisfy the following two conditions, for all $\mathbf{v}, \mathbf{u} \in V$, and all scalars c:

1. $T(\mathbf{v} + \mathbf{u}) = T(\mathbf{v}) + T(\mathbf{u})$.
2. $T(c\mathbf{v}) = cT(\mathbf{v})$.

We first investigated the linear transformations from $\mathbb{R}$ to $\mathbb{R}$. Unfortunately, the functions $T(x) = x^2$ and $T(x) = \sin x$ fail property (1), but we figured out that the identity map $T(x) = x$, and the line $T(x) = 2x$ are linear transformations! However, the functions $T(x) = 1$ and $T(x) = 2x + 1$ are not linear transformations, since they fail, for instance, property (1), even though they are also lines. We concluded that lines through the origin, i.e., functions of the form $T(x) = ax$ for a constant a are linear transformations.

Next, we turned to considering a linear transformation

$$T : \mathbb{R}^2 \to \mathbb{R}^2$$

given by, for $\mathbf{v} = (v_1, v_2)$,

$$T(\mathbf{v}) = (2v_1 - v_2, v_1 + v_2).$$

We verified the two properties that ensure this is a linear transformation.

Next, we noticed that for any vector space V, the identity function $T(\mathbf{v}) = \mathbf{v}$ is a linear transformation from V to itself. On the other hand, for any two vector spaces V and W, the zero function $T(\mathbf{v}) = \mathbf{0}$ (where $\mathbf{v} \in V$ and $\mathbf{0} \in W$ is a linear transformation from V to W.

We saw that linear transformations can be defined by the image of basis vectors. For instance, we saw in an example that $T(2, 3, 4)$ is defined by where the standard basis vectors map to in a linear transformation from $\mathbb{R}^3$ to itself.

Next time we will consider linear transformations defined by a matrix!

Also, remember to review the notions of **row/columns space**, **rank**, **null space**, and **nullity** for next time!

Lecture 24: Tuesday, November 20. We started class by recalling the definition of the row and column space of a matrix, using the example

$$A = \begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & -2 & 1 & 0 \\ -2 & 2 & -7 & -8 \end{bmatrix}.$$

In this case, the row space of A is the vector subspace

$$\text{span}\left\{\begin{bmatrix}1 & 0 & 3 & 4\end{bmatrix}, \begin{bmatrix}0 & -2 & 1 & 0\end{bmatrix}, \begin{bmatrix}-2 & 2 & -7 & -8\end{bmatrix}\right\} \subseteq \mathbb{R}^4$$

and the column space of A is

$$\text{span}\left\{\begin{bmatrix}1\\0\\-2\end{bmatrix}, \begin{bmatrix}0\\-2\\2\end{bmatrix}, \begin{bmatrix}3\\1\\-7\end{bmatrix}, \begin{bmatrix}4\\0\\-8\end{bmatrix}\right\} \subseteq \mathbb{R}^3.$$

Our first goal is to find a basis for the row and the column spaces.

In fact, A is row equivalent to the matrix

$$B = \begin{bmatrix}1 & 0 & 3 & 4\\0 & 1 & -\frac{1}{2} & 0\\0 & 0 & 0 & 0\end{bmatrix}.$$

and this means that the row vectors here span the same row space! We determined that the nonzero row vectors form a basis, so the row space of B, and of A, is two-dimensional.

This works in general: After turning a matrix A into a matrix B using elementary row operations, if B is in row-echelon form, the nonzero rows of B form a basis for the row space of A!

Next, we turned to the problem of finding a basis for the column space of A. Unfortunately, applying row operations can change the column space (but it does *not* change the dependency among the columns.)

We used two methods. In the first method, we saw that the column space of A is the same as the row space of A^T, so we row-reduced A^T to obtain the new matrix

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

in row-echelon form, so that after transposing back to columns, a basis for the column space of A is

$$\left\{\begin{bmatrix}1\\0\\-2\end{bmatrix}, \begin{bmatrix}0\\1\\-1\end{bmatrix}\right\}.$$

In particular, like the row space of A, it is two-dimensional!

Our second method of finding a basis for the column space of A uses the fact that row operations do not change dependency relations on columns (although they *can* change the column space). If $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ are the columns of A, and $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$ are the columns of B, then by inspecting B, we can see that

$$\mathbf{w}_3 = 3\mathbf{w}_1 - \frac{1}{2}\mathbf{w}_2 \text{ and } \mathbf{w}_4 = 4\mathbf{w}_1.$$

Therefore,

$$\mathbf{v}_3 = 3\mathbf{v}_1 - \frac{1}{2}\mathbf{v}_2 \text{ and } \mathbf{v}_4 = 4\mathbf{v}_1.$$

so that $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix}$ (which are linearly independent) form a basis for the column space of A! Although this basis is different than the first one, we see that the first vector is the same in each, and the second is just a scalar multiple of the other!

Next, we stated a **theorem**: If A is an $m \times n$ matrix, and the dimension of the row space equals the dimension of the column space. We call this the **rank** of A. In our example, $\text{rank}(A) = 2$.

Lecture 23: Thursday, November 15. We started class by looking at a video visualization of the span of vectors in $\mathbb{R}^2$ and in $\mathbb{R}^3$.

We recalled the definition of a *basis* for a vector space, and reminded ourselves of the bases we found last time, for $\mathbb{R}^2$, $M_{2\times 2}(\mathbb{R})$, and $P_3(\mathbb{R})$.

We stated a **theorem**: If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis consisting of n elements of a vector space V, then the following hold:

1. Every element of V can be written in *exactly one way* as a linear combination of S.
2. Every set consisting of *more than n elements* of V are linearly dependent.
3. Every basis for V consists of exactly n elements of V.

We saw a few examples, and then defined the **dimension** of a vector space V as the number of elements in any basis for V. (There is one exception: the vector space consisting only of the zero vector is said to have dimension *zero*, even though its basis consists of the one zero vector.)

We calculated the dimension of some familiar vector spaces: $\dim(\mathbb{R}^n) = n$, $\dim(P_n(\mathbb{R})) = n+1$, and $M_{m\times n}(\mathbb{R}) = mn$.

Then we approached the question of how to find the dimension of a vector subspace. We started with the subspace

$$W = \{(a, b, a+3b) \mid a, b \in \mathbb{R}\}$$

of $\mathbb{R}^2$. We saw that the vectors $(1, 0, 1)$ and $(0, 1, 3)$ are both in W, and that they are linearly independent since they're not multiples of one another. Therefore, to show that

$$S = \{(1, 0, 1), (0, 1, 3)\}$$

is a basis for W, we only need to show that S spans W. This holds, since given an arbitrary element of W, $(a, b, a+3b)$ for some real numbers a, b, we have that

$$(a, b, a+3b) = a(1, 0, 1) + b(0, 1, 3)$$

so that every element of W can be written as a linear combination of the elements of S. Therefore, S is a basis, and since it has two elements, $\dim(W) = 2$.

We then showed the the subspace W' of all *symmetric* 2×2 matrices, of $M_{2\times 2}(\mathbb{R})$ has dimension three by showing that the following is a basis:

$$T = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}.$$

Of course, we needed to show that S spans W', and that S is linearly independent. However, the following **theorem** allows us to skip one of these steps in the future: If V is a vector space of dimension n, then

1. If S is linearly independent, then S is a basis for V.

2. If S spans V, then S is a basis for V.

Next, we defined the **null space** $N(A)$ of an $m \times n$ matrix A as the vector subspace of $\mathbb{R}^n$ consisting of all column vectors $\mathbf{x} = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ for which $A\mathbf{x} = \mathbf{0} \in \mathbb{R}^m$. The dimension of $N(A)$ is called the **nullity** of A.

We already know how to solve the homogeneous system $A\mathbf{x} = \mathbf{0}$, so we focused on finding a basis for the nullspace. In our example of A above, we found the parametric solution to $A\mathbf{x} = \mathbf{0}$ using methods from class:

$$x_1 = -3s - 4t, x_2 = 1/2s, x_3 = s, x_4 = t$$

where s, t are real numbers. Therefore, the null space $N(A)$ is the set of all $\mathbf{x} \in \mathbb{R}^4$ of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3s-4t \\ 1/2s \\ s \\ t \end{bmatrix} = s\begin{bmatrix} -3 \\ 1/2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -4 \\ 0 \\ 0 \\ t \end{bmatrix}.$$

From here, we can determine that the vectors $\begin{bmatrix} -3 \\ 1/2 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -4 \\ 0 \\ 0 \\ t \end{bmatrix}$ form a basis for $N(A)$! Therefore, the nullity of $N(A)$ is two.

Finally, we solved a non-homogeneous linear system $A\mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix}$ like usual, and saw that all solutions look like an element in the null space (determined above) plus a fixed vector. This phenomenon is pervasive, as illustrated by the following **theorem**: If $\mathbf{x}_p$ is a particular solution to a non-homogeneous system $A\mathbf{x} = \mathbf{b}$, then *any* solution has the form $\mathbf{x}_p + \mathbf{x}_h$, where $\mathbf{x}_h$ is an element of the nullspace of A (i.e., a solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$).

Lecture 22: Tuesday, November 13. We started class by recalling the definition of what it means for a set to span a vector space, and gave several examples.

Then, we defined what it means for a set

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$$

of elements in a vector space V to be a **basis** for V:

1. S spans V, and

2. S is linearly independent.

We went back to our original examples, and determined whether a certain spanning set is a basis. Next, we justified the fact that the following set is a basis for $\mathbb{R}^n$:

$$S = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k\}$$

where $\mathbf{e}_i$ is the vector in $\mathbb{R}^n$ whose i-th coordinate is 1, and all other coordinates equal 0. This is called the **standard basis** for $\mathbb{R}^n$.

We showed that there are other bases; for example, the standard basis in $\mathbb{R}^2$ consists of $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$, but we also checked that

$$T = \{(1, 1), (-1, 1)\}$$

is also a basis by verifying that T spans $\mathbb{R}^2$, and is linearly independent in $\mathbb{R}^2$. (This boiled down to computing determinants!)

We then constructed our own basis for $P_2(\mathbb{R})$, the vector space of all polynomials with real coefficients in $\mathbb{R}$ that have degree at most 2:

$$S = \{1, x, x^2, x^3\}.$$

At first, we also included 0 in S, and the set did *span* $P_2(\mathbb{R})$, but we noticed that when the zero vector is included in a set, the set if *never* linearly independent.

We also gave a basis for $M_{2\times 2}(\mathbb{R})$ as the set

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

We saw, intuitively, why S spans $M_{2\times 2}(\mathbb{R})$, and gave an exercise to check that S is linearly independent.

In fact, this **standard basis** for $M_{2\times 2}(\mathbb{R})$ gives the idea for the standard basis for all $M_{m\times n}(\mathbb{R})$, which consists of matrices N_{ij} for $1 \le i \le m$ and $1 \le j \le n$, where N_{ij} has all 0 entries except its (i, j)-entry is 1.

Finally, we noticed that in every case we found, all bases appear to have the same number of elements. This will be important when we discuss the *dimension* of a vector space next time!

Lecture 21: Tuesday, November 6. Today, we defined the **span** of a set of vectors

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$$

as the set of all *linear combinations* of the vectors in S.

The set S is called **linearly independent** if, whenever given a linear combination of vectors that equals the zero vector,

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

then all the scalars must be zero, i.e., $c_1 = c_2 = \cdots = c_k = 0$. The set S is is called **linearly dependent** if it is not linearly independent.

We did several examples of finding the span of a set, and testing linear independence/dependence. Then, we described a method of testing linear independence:

1. From the linear equation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$, write a system of linear equations in $c_1, c_2, \ldots, c_k$.

2. Decide whether the system has only the trivial solution $c_1 = c_2 = \cdots = c_k = 0$. (Notice that this is always a solution!)

3. If the system only has the trivial solution, then S is linearly independent. If there are more solutions, the set is linearly dependent.

We used this method to test linear independence of several sets, one in $\mathbb{R}^3$, one in the vector space of polynomials of degree at most two $P_2(\mathbb{R})$, and one in the set of 2×2 matrices $M_{2\times 2}(\mathbb{R})$.

In the remaining time, we did a short review for the midterm. This included using determinants to tell whether there is a unique solution to a system of equations (without solving the system!) and stating the vector space axioms being used in logical steps.

For your convenience, we list a few other topics that might be important to review (feel free to come to office hours with questions): use elementary row operations to compute determinants, determine collinearity/coplanarity using determinants, find areas/volumes using determinants, determine whether a Markov chain is absorbing, apply Cramer's rule to solve a system of equations.

Lecture 20: Thursday, November 1. We started class today by reviewing the definition of a **linear combination** of vectors in a vector space.

Then we gave examples, illustrating how we can decide whether one vector $\mathbf{w}$ in a vector space V can be written as a linear combination of other vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, using linear algebra (more specifically, setting up and solving a system of linear equations)! We saw that it is possible that $\mathbf{w}$ can be written uniquely as linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, or can be written in infinitely many ways a linear combination, and it is also possible that it *cannot* be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Our examples included a variety of vector spaces V.

Next, given a set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in a vector space V, we defined what it means for S to be a **spanning set** for V: Every vector in V can be written as a linear combination of the elements in S.

If S is a spanning set for V, we also say S **spans** V.

We gave examples of sets that do, and do not, form spanning sets.

Lecture 19: Tuesday, October 30. Today we reviewed the criterion for a subset of a vector space to be a subspace, and then went through many examples, in each, determining whether a given subset is a subspace or not.

We started by looking at two subsets of $M_{2\times 2}(\mathbb{R})$, the vector space of all 2×2 matrices with entries that are real numbers.

- The set W of all symmetric matrices (i.e., matrices A for which $A^T = A$) is a vector subspace, using the criterion.

- The set Z of all singular matrices (i.e., matrices A for which $\det(A) = 0$) is **not** a subspace, since it is not closed under addition: For instance, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ are both singular, but their sum is the identity, which is definitely not singular!

Next, we turned to investigating the subspaces of $\mathbb{R}^2$. We showed that the set Y of all vectors of the form $(0, a)$, where a is any real numbers, is a vector space. It can be identified with the y-axis! In other words, $Y = \{(x, y) \mid x = 0\}$

We then showed that the set U of all vectors of the form (a, a), where a is again any real number, is also a vector space, and can be identified with the line $y = x$; i.e., $U = \{(x, y) \mid x = y\}$.

However, the set T consisting of vectors with *positive* coordinates is not a vector space–for instance, $(0, 0)$ is not in T. But even if we change the set to S, the set of all vectors in $\mathbb{R}^2$ with non-negative coordinates, it is still not a vector space, since it is not closed under scalar multiplication; for instance $-1 \cdot (2, 2) = (-2, -2)$, which is not in S although $(2, 2)$ is in S!

In fact, all vector subspaces of $\mathbb{R}^2$ must be one of the following:

- The set consisting only of $(0, 0)$.
- The set of all points on a fixed line passing through the origin.
- The entire set $\mathbb{R}^2$.

Going along with this, we saw why the set $\{(x, y) \mid x + 5y = 0\}$ is a vector space, while the set $\{(x, y) \mid x + 5y = 1\}$ cannot be one!

In line with this, we made a conjecture about the subspaces of $\mathbb{R}^3$, and we were correct! They are the following:

- The set consisting only of $(0, 0, 0)$.
- The set of all points on a fixed line passing through the origin.
- The set of all points on a fixed plane passing through the origin.
- The entire set $\mathbb{R}^3$.

Using this list, we can show that the set of all vectors of the form $(x_1, x_2, 1)$ (where x_1, x_2 are real numbers) does not form a vector space, but the set of vectors of the form $(x_1, x_1 + x_3, x_3)$ (where x_1, x_3 are real numbers) does form a vector space.

Finally, we defined a **linear combination** of vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ as a vector of the form

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdot + c_k\mathbf{u}_k$$

where the c_i are scalars.

We tied this to common calculations in physics class!

Lecture 18: Thursday, October 25. Today, we reviewed the definition of a vector space, and stated some properties of vector space addition, which required us to use the properties defining a vector space! More precisely, if $\mathbf{v}$ is any element of a vector space V, and c is a real number, then:

- $0 \cdot \mathbf{v} = \mathbf{0}$
- $c \cdot \mathbf{0} = \mathbf{0}$, and
- If $c\mathbf{v} = \mathbf{0}$, then $c = 0$ or $\mathbf{v} = \mathbf{0}$.

Next, we defined a **subspace** W of a vector space V as a subset of V that is itself a vector space. We saw that the set containing only the element $(0, 1)$ in $\mathbb{R}^2$ is *not* a vector space, but the set containing only $(0, 0)$ is a vector space! In general the set containing only the zero element, and a vector space itself, are always subspaces.

Finally, we stated a simple criterion to verify whether a subset W of a vector space V is a subspace: W is a subspace if the following two conditions hold.

- $\mathbf{u}+\mathbf{u} \in W$ for any $\mathbf{v}, \mathbf{u} \in W$, and
- $c\mathbf{u} \in W$ for any $\mathbf{u} \in W$ and scalar c.

Lecture 17: Tuesday, October 23. Our focus today was on the notion of a **vector space**: A set V with two operations:

- **Addition**: For any $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u}+\mathbf{v} \in V$.
- **Scalar multiplication**: For any $\mathbf{v} \in V$ and real number c, $c\mathbf{v} \in V$.

Moreover, these operations are required so satisfy the following properties: **Properties of addition.**

- **Commutativity**: For any $\mathbf{u}, \mathbf{u} \in V$,
$$\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}.$$
- **Associativity**: For any $\mathbf{u}, \mathbf{u}, \mathbf{w} \in V$,
$$(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w}).$$
- **Additive identity**: There is a special element in V called the **zero element**, and denoted $\mathbf{0}$, such that for any $\mathbf{v} \in V$,
$$\mathbf{v}+\mathbf{0} = \mathbf{v}.$$
- **Negatives**: For any $\mathbf{v} \in V$, there is another element in V called the **negative** of $\mathbf{v}$, and denoted $-\mathbf{v}$, for which
$$\mathbf{v}+(-\mathbf{v}) = \mathbf{0}.$$

Properties of scalar multiplication.

- **Distributivity**: For any $\mathbf{u}, \mathbf{v} \in V$ and real numbers c, d,
$$c(\mathbf{u}+\mathbf{v}) = c\mathbf{u}+c\mathbf{v}\mathbf{v} \text{ and } (c+d)\mathbf{u} = c\mathbf{u}+d\mathbf{u}.$$
- **Associativity**: For any $\mathbf{u} \in V$ and real numbers c, d,
$$c(d\mathbf{u}) = (cd)\mathbf{u}.$$
- **Multiplication by** 1: For any $\mathbf{u} \in V$, $1\mathbf{u} = \mathbf{u}$.

We then spent most of the class discussing examples, and non-examples, of vectors spaces!

1. The first examples is the set of all vectors in $\mathbb{R}^2$; we use all of the axioms when doing arithmetic of vectors! Of course, the same ideas make the set of all vectors with three coordinates, $\mathbb{R}^3$, into a vector space, or even $\mathbb{R}^n$ for any integer $n \geq 1$.
2. The previous examples includes $\mathbb{R}$ as a vectors space: notice that "addition" is just regular addition of real numbers, and "scalar multiplication" is just regular multiplication of real numbers!

3. The set of integers $\mathbb{Z}$ is **not** a vector space, because you can find real number scalar, which after multiplication by an integer, results in a number that is *not* an integer: e.g., $\frac{1}{2} \cdot 3 = \frac{3}{2} \notin \mathbb{Z}$.

4. The subset of $\mathbb{R}^2$ consisting of all vectors with *positive* coordinates is **not** a vector space: For example, there is no zero element!

5. The set of all 2×2 matrices, $M_{2\times 2}(\mathbb{R})$ is a vectors space, with zero element the zero matrix. In fact, the set of all $n \times m$ matrices, $M_{n\times m}(\mathbb{R})$, is also a vector space!

6. The set $P(\mathbb{R})$ of all polynomials is a vectors space, with the usual addition and scalar multiplication, and $0 = 0 \cdot x^0$ the zero element. The subset $P_2(\mathbb{R})$ of all degree 2 polynomials (i.e., all parabolas) is **not** a vector space, since we don't have well-defined addition: for instance,

$$(x^2 + 1) + (-x^2 + x) = x + 1$$

which is *not* a parabola. However, you can check that the set $P_{\leq 2}(\mathbb{R})$ of all polynomials of degree *at most* 2 is a vector space! So is $P_{\leq n}$, the set of polynomials of degree at most n, for any $n \geq 0$.

7. The set $C(\infty, \infty)$ of all continuous functions from the real numbers to the real numbers is a vector space, with regular addition and scalar multiplication, and zero element the function sending every element to $0 \in \mathbb{R}$. So is $C[a, b]$, the set of all continuous functions with domain an interval $[a, b]$ to the real numbers.

We showed that $0 \cdot \mathbf{v} = \mathbf{0}$ for any $\mathbf{v}$ in a vector space V. This required the application on several of the properties defining a vector space!

Lecture 16: Thursday, October 18. Today, we reviewed and illustrated **Cramer's rule** for solving systems of linear equations using determinants.

Next, we gave a **formula for the area of a triangle** with vertices $(x_1, y_1), (x_2, y_2)$, and (x_3, y_3) in terms of determinants: it is the absolute value of

$$\frac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.$$

We showed why this holds! Then we did an example to verify the formula we know (when the triangle has a horizontal base).

Next, we saw that these ideas give us a **test for three collinear points**: $(x_1, y_1), (x_2, y_2)$, and (x_3, y_3) line on one line if and only if

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \neq 0.$$

We can analogously determine a **test for four coplanar points** in three dimensions using the determinant of a larger matrix!

The test for collinear points also gives us a **formula for the equation of a line through two points**:

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0.$$

We saw what happens here if the two points do not pass the horizontal line test!

Next, we gave a formula for the **volume of a tetrahedron** using a matrix of order 4!

Finally, we quickly reviewed the definition of a vector in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Lecture 15: Thursday, October 11. Today, we defined the *adjoint* $\operatorname{adj}(A)$ of a square matrix A as the *transpose* of the **cofactor matrix** $[C_{ij}]$. The significance of this matrix is that if A is invertible, then

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A).$$

We did an example of computing the adjoint of a 3×3 matrix, and after pointing out that its determinant is nonzero, we used the above fact to find its inverse. We pointed out that we can always check whether our computation of the inverse of a matrix is correct, by multiplying it by the original matrix (in both orders), and verifying that we obtain the identity matrix. Then we showed why the above fact holds, for *any* matrix of order 2.

We then turned to explaining **Cramer's rule**, first for systems of two equations in two variables, but there is an analogous rule for any system of equations with the same number of variables as equations. Given a linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

let A denote its coefficient matrix $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Then if A_i is the matrix obtained from A by replacing the i-th column of A with the constant vector $\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, if a solution exists (i.e., $\det(A) \neq 0$), it is

$$x_1 = \frac{|A_1|}{|A|}, x_2 = \frac{|A_2|}{|A|}.$$

In general, with n equations and n unknowns, the same definition of A_i (replacing the i-th column of the $n \times n$ coefficient matrix A with the $n \times 1$ matrix of constants) yields the solution (if it exists) given by

$$x_1 = \frac{|A_1|}{|A|}, x_2 = \frac{|A_2|}{|A|}, \ldots, x_n = \frac{|A_n|}{|A|}.$$

Lecture 14: Tuesday, October 9. We started class by reviewing what happened to a given 2×2 matrix after an elementary row operation was applied. The results aligned with the following **theorem**: If A and B are square matrices, and

- B is obtained from A by interchanging two rows, then $|B| = -|A|$.
- B is obtained from A by scaling a row by scalar c, then $|B| = c|A|$.
- B is obtained from A by adding a multiple of a row of A to another row of A, then $|B| = |A|$.

We saw how this theorem is useful, by performing row operations until we obtain an upper-triangular, and tracking what happens to the determinant with each operation.

Next, we noted that we can do the analogous *column operations* to a matrix, and the determinant will change in the same way as the theorem indicates, where the word "row" is replaced with "column".

We argued why any of the following conditions guarantee a **zero determinant**:

- An entire row (or column) consists of 0s.
- Two rows (or columns) are equal.
- One row (or column, respectively) is a multiple of another row (or column).

We then turned to other properties of determinants. In fact, if A and B are square matrices of the same size, then

$$\det(AB) = \det(A)\det(B).$$

If A has order n, then

$$\det(cA) = c^n \det(A),$$

and we showed why this holds for any 2×2 matrix.

In fact, it is *always* true that

$$A \text{ is nonsingular} \iff \det(A) \neq 0,$$

and if A is nonsingular, then we argued why

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Using cofactors, we also argued that for any square matrix A,

$$\det(A^T) = \det(A).$$

Remember to study Markov Chains and determinants for a potential quiz on Thursday!

Lecture 13: **Thursday, October 4**. Today, we recalled the definitions of the **minors** and **cofactors** of a matrix, and the sign pattern for cofactors.

After reviewing the formula for the determinant of a 1×1 and 2×2 matrix, we defined the **determinant** of a square matrix of order $n \geq 2$ (i.e., a $n \times n$ matrix) recursively, as the sum of the entries in its first row, multiplied by the respective cofactors.

We calculated an example explicitly of the determinant of a 3×3 matrix using this definition.

Next, we stated the fact that in the definition of the determinant, we can expand along the first *column* instead of the first row, or along *any row* or *any column*! We must just be careful with the signs of the cofactors.

We went back to our initial example, and calculated the determinant using the second column instead of the first row, and got the same answer!

Next we turned to a matrix with several zero entries, and saw how to choose the row/column over which to expand, based on the placement of the zero entries. We calculated the determinant of a matrix of order 4 using the definition, after carefully choosing the row/column to expand over.

We investigated the determinant of upper and lower triangular matrices, and stated the **fact** that the determinant of any such matrix is simply the product of the entries along the diagonal.

Finally, after noticing that calculating the determinant of two row-equivalent matrices can be easier or harder that one another, depending on the form of the matrices, we saw what happens to the determinant in an example, after each kind of elementary row operation is performed. Next time we will make this precise!

Lecture 12: Tuesday, October 2. We started class by reviewing the definitions of a **stochastic matrix**, **matrix of transition probabilities**, **initial state matrix**, **Markov chain**, and **steady state matrix**.

We gave an example of a Markov chain, using the matrix

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

of transition probabilities of smartphone users owning an Android (first state) or iPhone (second state). Our initial state matrix was $X_0 = \begin{bmatrix} .4 \\ .6 \end{bmatrix}$, representing the fact that at this point in time, 40% of users with an Android phone, and 60% an iPhone.

We calculated the percentage of users with each type of smartphone after 1 and 2 units of time, and then calculated the **steady state matrix** $\overline{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ using the system of linear equations coming from the equality

$$P\overline{X} = \overline{X}$$

along with the equation $x_1 + x_2 = 1$. In fact, $\overline{X} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$.

We defined a **regular stochastic matrix** as a stochastic matrix P for which some power P^n has all positive entries; we gave a couple examples of regular and non-regular stochastic matrices. The corresponding Markov chain is called a **regular Markov chain**, and we state the **theorem** that any regular Markov chain has a **unique steady state**.

Next, we call the i-th state of a stochastic matrix $P = [p_{ij}]$ an **absorbing state** if $p_{ii} = 1$ (so all other entries in the i-th column must equal 0). We call the associated Markov chain an **absorbing Markov chain** if

- There is at least one absorbing state, and
- A member of the population can move from any non-absorbing state to any absorbing state in a finite number of transitions.

Finally, we stated the following **theorem**: An absorbing Markov chain has either a *unique steady state*, or *infinitely many steady states* (in the sense that $P\overline{X} = \overline{X}$).

Next, we switched gears completely. We recalled the definition of the **determinant** of a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ as

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

We know that this value is useful since it tells us whether A is nonsingular: $\det(A) \neq 0$ if and only if A is nonsingular.

Our next goal will be to define the determinant of a 3×3 matrix

$$A = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$$

although the definitions we will now give apply to *any* square matrix.

The **minor** M_{ij} **of** a_{ij} is obtained by first deleting the i-th row and j-th column of A, and then taking the determinant of the remaining matrix. From here, the **cofactor of** a_{ij} is $C_{ij} = (-1)^{i+j} M_{ij}$.

We gave several examples computing minors and cofactors, and then described the general sign pattern for cofactors.

Make sure to know these two last definitions for next time!

Lecture 11: Tuesday, September 25. After discussing Miterm 1 and Quiz 4, we started discussing Markov Chains.

A **stochastic matrix** is a square matrix, with every entry between 0 and 1 (inclusive), and each column sums to 1. We gave several examples of stochastic matrices.

One example of a stochastic matrix is a **matrix of transition probabilities** associated to n states $S_1, \dots, S_n$, which is an $n \times n$ matrix $P = [p_{i,j}]$, where p_{ij} is the probability that the j-th state will change to the i-th state.

We did an example of weather prediction, where the first state is "dry," and the second is "wet."

$$P = \begin{bmatrix} \frac{3}{4} & \frac{2}{3} \\ \frac{1}{4} & \frac{1}{3} \end{bmatrix}.$$

For example, there is $\frac{1}{4}$ chance of weather changing from dry to wet. We drew a diagram representing the states and matrix of transition probabilities.

A **Markov chain** is a sequence of state matrices X_n for which $X_{k+1} = PX_k$.

Here, our **initial state matrix**, which represents the current state, is $X_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ since the initial state is "dry." To find the state matrix after 1 unit of time, we calculate

$$X_1 = PX_0 = \begin{bmatrix} \frac{3}{4} & \frac{2}{3} \\ \frac{1}{4} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} \\ \frac{1}{4} \end{bmatrix}.$$

Iterating this process, we find that after 2 units of time, we have

$$X_2 = PX_1 = P^2 X_0 = \begin{bmatrix} \frac{105}{144} \\ \frac{39}{144} \end{bmatrix}.$$

and

$$X_3 = PX_2 = P^3 X_0 = \begin{bmatrix} \frac{1257}{1728} \\ \frac{471}{1728} \end{bmatrix}.$$

The **n-th state matrix** is $X_n = P^n X_0$, which satisfies the recursion $X_{n+1} = PX_n$.

In fact, in our case, $X_{100} \approx \begin{bmatrix} .727 \\ .272 \end{bmatrix}$.

A **steady state matrix** is a matrix $\overline{X}$ for which $P\overline{X} = \overline{X}$. If the limit $P^n X$ exists, its limit is a steady state. We used the formula $P\overline{X} = \overline{X}$ to find the steady state in our case to be $\begin{bmatrix} \frac{8}{11} \\ \frac{3}{11} \end{bmatrix}$.

Lecture 10: Thursday, September 20. We started class by recalling the definition of **upper-triangular** and **lower-triangular** matrices. Next, we defined an LU-factorization of a square matrix as a product

$$A = LU$$

where L is a lower-triangular, and U is an upper-triangular matrix.

In fact, an LU-factorization exists if A can be transformed into an upper-triangular matrix using only the elementary row operations of adding a multiple of one row to another row below it.

We did an example of finding an LU-factorization of a matrix A. Then we applied our result to find a solution to an equation of the form $A\mathbf{x} = \mathbf{b}$, using the following steps: Note that we want to solve $\mathbf{b} = A\mathbf{x} = (LU)\mathbf{x} = L(U\mathbf{x})$; i.e.,

$$L(U\mathbf{x}) = \mathbf{b}.$$

- **Step 1.** Let $\mathbf{y} = U\mathbf{x}$ and solve the system $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$ (using forward-substitution).
- **Step 2.** Now solve the system $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$ using back-substitution.

Lecture 9: Tuesday, September 18. Today, we discussed the notion of "if and only if," and used it throughout class today.

We defined an **elementary matrix** as an $n \times n$ matrix that can be obtained from the identity matrix I_n by **one** elementary row operation. We gave several examples of elementary matrices (and matrices that are not elementary!).

We saw, through an example, that the following holds: If E is an elementary matrix obtained by performing a specific elementary row operation on I_n, then EA is the same elementary row operation applied to an $n \times m$ matrix A.

Next, we started with a matrix A, and saw that row reduction corresponds to multiplying A on the left by a series of elementary matrices. We say that matrices A and B of the same size are **row equivalent** if there are elementary matrices $E_1, E_2, \ldots, E_k$ for which

$$B = E_k E_{k-1} \cdots E_2 E_1 A.$$

We stated the fact that if E is an elementary matrix, then it is invertible, and E^{-1} is also an elementary matrix. We gave several examples of this, recalling our method of finding inverses (and the formula for the inverse of a 2×2 matrix.

In fact, a square matrix A is invertible if and only if it can be written as a product of elementary matrices! (This can be the "product" of only one matrix.) We went through an example of performing Gauss-Jordan elimination to a matrix, and using each step to write the matrix as a product of elementary matrices. In this example, we got that our matrix A could be written as

$$E_3 E_2 E_1 A = I_2$$

where E_1, E_2, E_3 are elementary matrices, and then iteratively multiplying by E_3^{-1}, then E_2^{-1}, and then E_1^{-1}, that

$$A = E_1^{-1} E_2^{-1} E_3^{-1}.$$

This fact will be very useful later on when discussing "determinants"!

Finally, for what we will start next time, we defined an **upper-triangular matrix** as a square matrix with only 0 entries below the diagonal, a and a **lower-triangular matrix** as a square matrix with only 0 entries below the diagonal.

Lecture 8: Thursday, September 13. Today, we continued to discuss **invertible** matrices, and the **inverse** of an invertible matrix. We are motivated by the fact that for real numbers a, b, we can always solve for the real number x in the equation

$$ax = b$$

by letting $x = \frac{1}{a} \cdot b$, *unless* $a = 0$! This is in analogy with solving a matrix equation

$$AX = B$$

for X, where A and B are given square matrices with the same size. If A is invertible, then $AX = B$ precisely if (*if and only if*) $A^{-1}(AX) = A^{-1}B$, or equivalently, $(A^{-1}A)X = A^{-1}B$, or $X = I_nX = A^{-1}B$. So the unique solution if A is invertible is $X = A^{-1}B$!

We set up an example where such a solution is **not possible**: If $C = \begin{bmatrix} 0 & 0 \\ -11 & -1 \end{bmatrix}$ and $D = \begin{bmatrix} 5 & 1 \\ 0 & 7 \end{bmatrix}$, then the equation

$$CX = D$$

cannot have a solution, since the upper-left corner entry of CX will always be 0, while that of D is 5.

Next, we reviewed the method we developed last time of how to find the inverse of the matrix $A = \begin{bmatrix} 1 & 5 \\ -1 & -3 \end{bmatrix}$: First, we ajoin A with the identity matrix I_2 to obtain the matrix

$$\left[\begin{array}{cc|cc} 1 & 5 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{array}\right]$$

and then we perform elementary row operations until the left-hand matrix (which was A) becomes I_2. In this case, we obtained

$$\left[\begin{array}{cc|cc} 1 & 0 & -\frac{3}{2} & -\frac{5}{2} \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array}\right]$$

The right-hand matrix $\begin{bmatrix} -\frac{3}{2} & -\frac{5}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ is then A^{-1}!

In general, the same method works: To find the inverse of an $n \times n$ matrix A, we adjoin A and I_n to get a matrix $A \mid I_n$, and then perform elementary row operations until the left-hand matrix is I_n; then the right-hand matrix is A^{-1}. In other words, we perform Gauss-Jordan elimination to $A \mid I_n$ to obtain $I_n \mid A^{-1}$.

However, this does *not* work if A is not invertible (since A^{-1} does not exist). What happens here is that when A is turned into reduced row-echelon form, it is *not* the identity matrix. For example, if we try to find the inverse of $B = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 6 & 0 & 1 \end{array}\right]$$

and after replacing $R_2 - 3 \cdot R_1 \to R_2$, we have

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

and the left-hand side is already in reduced row-echelon form. Therefore, we can conclude that the matrix B has no inverse!

We stated the following important **Fact**: A 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible precisely if $ad - bc \neq 0$ (and we refer to the value $ad - bc$ as the *determinant* of A). Moreover, if A is invertible, then its inverse is

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

We checked that our matrix $A = \begin{bmatrix} 1 & 5 \\ -1 & -3 \end{bmatrix}$ from earlier satisfies $ad - bc = 1 \cdot -3 - 5 \cdot (-1) = 2 \neq 0$, and that its inverse satisfies the formula given.

Finally, we stated some important properties and conclusions about inverses of matrices. We verified or gave evidence toward each:

Theorem: If A is an invertible matrix, and c is a scalar, the following hold.

1. $(A^{-1})^{-1} = A$
2. $(A^k)^{-1} = (A^{-1})^k$
3. $(cA)^{-1} = \frac{1}{c} A^{-1}$
4. $(A^T)^{-1} = (A^{-1})^T$

Theorem: If A and B are invertible matrices with the same size, then

$$(AB)^{-1} = B^{-1} A^{-1}.$$

Theorem: If C is an invertible $n \times n$ matrix, then the **cancellation property** with respect to C holds: If A and B are $m \times n$ matrices and $AC = BC$, then $A = B$. Moreover, if D and E are $n \times p$ matrices and $CD = CE$, then $D = E$.

Theorem: If A is an invertible $n \times n$ matrix, and $\mathbf{b}$ is an $n \times 1$ vector, then there exists a solution to the equation

$$A\mathbf{x} = \mathbf{b}$$

where $\mathbf{x}$ is a $n \times 1$ vector.

Lecture 7**:** **Tuesday, September 11**. Today, we started with an example showing that **cancellation** does not always hold for matrices: If $AB = AC$, then it is necessarily true that $A = B$.

We recalled the definition of the **transpose** of a matrix A, and described it three different ways. We defined a matrix A to be **symmetric** if it is a square matrix, and $A = A^T$.

We stated a **theorem** on propertes of transpose matrices:

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(cA)^T = c(A^T)$

4. $(AB)^T = B^T A^T$

and did an example verifying the last statement.

After seeing an examples, and then verified that given a matrix A, AA^T is always symmetric. We also did an example verifying this.

Next, given matrices A and B, and motivated by finding a solution to

$$AX = B$$

for some matrix X, we gave the following definition:

A square $n \times n$ matrix A is **invertible** (or **nonsingular**) if there exists a matrix B for which

$$AB = BA = I_n$$

If A is invertible, then the matrix B above is called the (**multiplicative**) **inverse** of A, and is often denoted A^{-1}.

A matrix that is not invertible is called **non-invertible** or **singular**.

We showed that the inverse of $\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$ is $\begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$, and then turned the the question of finding the inverse of the matrix

$$A = \begin{bmatrix} 1 & 5 \\ -1 & -3 \end{bmatrix}$$

If $X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$ is the inverse of A, then in particular, $AX = I_2$; i.e.,

$$\begin{bmatrix} 1 & 5 \\ -1 & -3 \end{bmatrix} \cdot \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We turned this product requirement into a system of equations in the variables x_{ij}, and found the solution. Next, we saw that this was equivalent to starting with the matrix obtained by **adjoining** A and I_2:

$$\left[\begin{array}{cc|cc} 1 & 5 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{array}\right]$$

and performing elementary row operations to the left-hand "A" portion until it is transformed into the identity matrix. In this case, we got

$$\left[\begin{array}{cc|cc} 1 & 0 & -\frac{3}{2} & -\frac{5}{2} \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array}\right]$$

The right-hand matrix $\begin{bmatrix} -\frac{3}{2} & -\frac{5}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ is A^{-1}!

This generalizes to a method that can be used to find all inverses: Adjoin the matrix and the identity matrix, and perform Gauss-Jordan elimination until the left-hand side becomes the identity matrix. Then the right-hand side is the inverse of the original matrix!

However, a word of **warning**: If the matrix you begin with is not invertible, it will not be possible to perform elementary row operations to obtain the identity matrix on the left-hand side!

Lecture 6: Thursday, September 6. Today, we first reviewed how to turn a system of equations

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

can be written as *one* equation of matrices,

$$A\mathbf{v} = \mathbf{b},$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \ddots & & \vdots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{bmatrix}, \mathbf{v} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

We then discussed how to **partition** the product of matrices, to write the left-hand side of the system of linear equations as a **linear combination** of column matrices: If, for $1 \leq i \leq n$, $\mathbf{a}_i$ denotes the column matrix $\begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix}$, then the system can be written as

$$\begin{aligned} & x_1\mathbf{a_1} + x_2\mathbf{a_2} + \cdots x_n\mathbf{a_n} \\ =& x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}. \end{aligned}$$

We next began to discuss **properties of matrix operations**. We started with some basic properties: If A, B, C are $m \times n$ matrices, and c, d are real scalars, then

- **Commutativity of addition.** $A + B = B + A$
- **Associativity of addition.** $A + (B + C) = (A + B) + C$
- **Associativity of iterated scalar multiplication.** $(cd)A = c(dA)$
- **Scalar identity.** $1 \cdot A = A$
- **Distributivity of scalar multiplication.** $c(A + B) = cA + cB$ and $(c + d)A = cA + dA$

We defined the **negative** of a matrix $A = [a_{ij}]$ as the matrix $-A = [-a_{i,j}]$. We also defined the $m \times n$ **zero matrix 0** or $0_{m\times n}$ as the $m \times n$ matrix in which every entry of 0.

We stated a **theorem** on properties of zero matrices: If A is $m \times n$, and c is a scalar, then

1. $A + \mathbf{0} - A$
2. $A + (-A) = \mathbf{0}$

3. If $cA = \mathbf{0}$, then $c = 0$ or $A = \mathbf{0}$

We did an example of using the properties of matrix operations so far to solve a matrix equation of the form $2X + A = B$ for X.

Next, we discussed some **properties of matrix multiplication**: Given a scalar c, and matrices A, B, C for which the following products defined,

- **Associativity of multiplication.** $A(BC) = (AB)C$
- **Distributivity.** $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$
- **Associativity/Distributivity of scalar multiplication.** $c(AB) = (cA)B = A(cB)$

We did an example in which we verified that associativity of multiplication held, and stated that we can't **cancel** matrices; i.e., if $AC + BC$, then it is not necessarily true that $A = B$.

We defined the $n \times n$ **identity matrix** I or I_n as the matrix with 1 along the diagonal, and zeros everywhere else. We then stated a **theorem** on its properties: If A is an $m \times n$ matrix, then

1. $AI_n = A$
2. $I_m A = A$

We then did some examples using identity matrices, then defined powers of matrices, and gave an example.

Finally, we defined the **transpose** of an $m \times n$ matrix $A = [a_{ij}$ as an $n \times m$ matrix A^T that has the rows of A as its column. In other words, the entry in the i-th row and j-th column is a_{ji}.

Lecture 5: Tuesday, September 4. After our quiz on row-echelon and reduced row-echelon form, we discussed matrix operations.

If we denote a matrix by a variable, it is typically a capital letter, and if A is an $m \times n$ matrix, then

$$A = [a_{ij}] = [a_{ij}]_{m \times n}$$

is shorthand for

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Two matrices are **equal** if the have the same *size*, and the same *entries*.

If one of the dimensions of the matrix is 1, then we sometimes call it a vector; an $m \times 1$ matrix is called a **column vector** or **column matrix**, and a $n \times 1$ matrix is a **row vector/matrix**. These matrices are often called lower case bold letters; e.g., $\mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} -2 & 5 & 0 \end{bmatrix}$.

Given two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$, we can only **add** (or subtract) the matrices if they have the same size (i.e., they are both $m \times n$ matrices, for some m and n), and their sum (or different) is given by the sum (or difference) of the two coordinates:

$$A + B = [a_{ij} + b_{ij}].$$

For example,

$$\begin{bmatrix} 2 & 7 \\ 5 & -1 \end{bmatrix} + \begin{bmatrix} -2 & 2 \\ 0 & 10 \end{bmatrix} = \begin{bmatrix} 0 & 9 \\ 5 & 9 \end{bmatrix}.$$

We can multiply any matrix by a scalar (any real number) by multiplying each of its entries by that real number; e.g.,

$$3 \cdot \begin{bmatrix} 2 & 7 \\ 5 & -1 \end{bmatrix} = \begin{bmatrix} 6 & 21 \\ 15 & -3 \end{bmatrix}.$$

We can only multiply matrices A and B to obtain AB if the number of columns of A equals the number of rows of B, i.e., if A is an $m \times n$ matrix, then B must be an $n \times p$ matrix, for some p. The resulting matrix AB will then be an $m \times p$ matrix. If $AB = [c_{i,j}]_{m \times p}$, then c_{ij} is defined by taking the dot product of the i-th column of A and the j-th row of B:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj.}$$

(Make sure you can draw a picture to explain this!)

For example, the product of a 2×2 and a 2×2 matrix is defined, and is a 2×2 matrix. Moreover,

$$\begin{pmatrix} 2 & 4 \\ 5 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ 0 & 8 \end{pmatrix} = \begin{pmatrix} 2 \cdot 1 + 4 \cdot 0 & 2 \cdot -3 + 4 \cdot 8 \\ 5 \cdot 1 + -1 \cdot 0 & 5 \cdot -3 + -1 \cdot 8 \end{pmatrix} = \begin{pmatrix} 2 & 26 \\ 5 & -23 \end{pmatrix}.$$

As another example, the product of a 2×3 and a 3×4 matrix is defined and is a 2×4 matrix. For instance,

$$\begin{bmatrix} 7 & 0 & -3 \\ 5 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -11 & 5 & 1/7 \\ 1 & 1 & 1 & 1 \\ 100 & \pi & 0 & 0 \end{bmatrix}$$

equals

$$\begin{bmatrix} 7 \cdot 0 + 0 \cdot 1 + -3 \cdot 100 & 7 \cdot -11 + 0 \cdot 1 + -3 \cdot \pi & 7 \cdot 5 + 0 \cdot 1 + -3 \cdot 0 & 7 \cdot 1/7 + 0 \cdot 1 + -3 \cdot 0 \\ 5 \cdot 1 + 1 \cdot 1 + -1 \cdot 100 & 5 \cdot -11 + 1 \cdot 1 + -1 \cdot \pi & 5 \cdot 5 + 5 \cdot 1 + -3 \cdot -1 & 5 \cdot 1/7 + 1 \cdot 1 + -1 \cdot 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 100 & -77 - 3\pi & 35 & 1 \\ -99 & -54 - \pi & 26 & 12/7 \end{bmatrix}.$$

Finally, we discussed turning a system of linear equations into an equality of matrices. To check your knowledge, verify that the system

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

can be written as *one* equation of matrices,

$$A\mathbf{v} = \mathbf{b},$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \ddots & & \vdots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{bmatrix}, \mathbf{v} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Lecture 4: Thursday, August 30. We started class by discussing **free variables** in more detail, and writing the solution set of a system of linear equations parametrically, when there is more than one parameter. For example, a system in variables x, y, z, w could be transformed, using Gauss-Jordan elimination, into the augmented matrix

$$\begin{pmatrix} 1 & 0 & 0 & 6 & -1 \\ 0 & 0 & 1 & 7 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

corresponding to the system of equations

$$\begin{aligned} x + 6z &= -1 \\ z + 7w &= 1 \end{aligned}$$

Since there are four variables and two unknowns, we can start by assigning z to be free; let $z = t$ for t any real number. Then $7w = 1 - z$, so that $w = \frac{1}{7} - \frac{1}{7}t$. Moreover, $x = -1 - 6z = -1 - 6t$ Finally, y can be free as well; let $y = s$, where s is a (possibly different than t) real number. Our solutions are given parametrically by

$$\begin{aligned} x &= -1 - 6t \\ y &= s \\ z &= t \\ w &= \frac{1}{7} = \frac{1}{7}t. \end{aligned}$$

Next, we defined a **homogeneous** system of linear equations as a system that have the constant 0 on the right-hand side of every equation:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= 0 \end{aligned}$$

We noticed immediately that a homogeneous system is always consistent, since there is always a solution

$$x_1 = x_2 = \cdots = x_n = 0,$$

which is sometimes called the **trivial** (or **obvious**) solution.

Therefore, we can use our methods to determine whether this is the only solution, or whether the system has infinitely many solutions. We also noticed that elementary row operations never change the last column of the augmented matrix; it continues to have all zeros.

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} & 0 \\ \vdots & \ddots & & \vdots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} & 0 \end{bmatrix}.$$

Next, we discussed applications of solving systems of linear equations. The first was finding the equation of a polynomial that passes through given points. For example, we asked whether there is a parabola going through the points

$$(-3, 1), (1, -2), \text{ and } (5, 5).$$

If we write the parabola as

$$p(x) = a_0 + a_1 x + a_2 x^2,$$

then

$$\begin{aligned} 1 &= p(-3) = a_0 - 3a_1 + 9a_2 \\ -2 &= p(1) = a_0 + a_1 + a_2 \\ 5 &= p(5) = a_0 + 5a_1 + 25a_2 \end{aligned}$$

This is a system of linear equations in the variables a_0, a_1, a_2, and we can solve it using our methods, to find that the only solution is

$$a_0 = -\frac{35}{16}, a_1 = -\frac{1}{8} a_2 = \frac{5}{16}.$$

We also discussed changing the degree of the polynomial (highest power of x appearing), and shifting the polynomial to move to an easier system of equations.

Our second application was to Network Analysis. A network is given by branches and junctions, where the total flow into a junction must equal the total flow out. Networks are used, for example, in economics, in traffic analysis, and in electrical engineering.

We did an examples of transforming a network, with some values of flow variable, into a system of linear equations. One specific application was to an electrical network, where Kirchoff's Laws can transform the network into a system of equations.

For Tuesday, make sure to know how to perform Gaussian elimination and Gauss-Jordan elimination–these might pop up on a quiz!

Lecture 3**: Tuesday, August 28**. We started class by defining what it means for a matrix to be in **row-echelon form**. There are three requirements:

1. Any row consisting only of 0's (called a **zero row**) must be at the bottom of the matrix.
2. The first nonzero entry in a row that does not consist only of 0's (called a **nonzero row**) is 1, which we call a **leading** 1.
3. The leading 1 in a higher nonzero row is to the left of the leading 1 in any lower row.

In addition, if there are only 0's above and below all leading 1's, then we say that the matrix has **reduced row-echelon form**.

We gave several examples of matrices, and decided whether each has row-echelon form, and if so, whether it has reduced row-echelon form.

Next, we described **Gaussian elimination with back-substitution**: Starting with a system of equations,

1. Write the augmented matrix corresponding to the system.

2. Use elementary row operations to rewrite the matrix in row-echelon form.
3. Write out the system of equations corresponding to the new matrix, and use back-substitution to determine the solutions.

We did two examples of this process, noticing that a matrix in row-echelon form is amenable to the back-substitution process. One of our examples ended up having exactly one solution, and one had no solutions (so was inconsistent).

Next, we described the method of **Gauss-Jordan elimination** to turn a matrix in row-echelon form into one in reduced row-echelon form, through an example. We noticed that this process can lead to fractional coefficients in the final system of equations, but that it is very easy to solve a system of equations coming from a matrix in reduced row-echelon form. We did an extra example, where there were infinitely many solutions, and we wrote the solution set parametrically.

Lecture 2: Thursday, August 23. Today, we recalled the examples of an arbitrary system of linear equations, an a system in **row-echelon form**, from last class. Here, the two systems of equations are **equivalent**, meaning that they have exactly the same solution set. We saw that row-echelon form has the big advantage that we can easily determine all solutions (or that there is no solution) using back-substitution, and we did this using our example.

We turned an arbitrary system of linear equation into row-echelon form in three settings, and then used the solution to determine all solutions. In the first case, there was exactly one solution, in the second, there were no solutions, and in the third, there were infinitely many solutions. We recalled that there precisely cover the three possibilities for the number of solutions in a solution set of a system of linear equations. We also described geometrically what each of these mean.

In the second case, when there were no solutions, two of the equations in row-echelon form were **inconsistent**, meaning that both cannot hold at once. Here, we call the system **inconsistent** as well. In the third case, where there were infinitely many solutions, we could eliminate the third equation and get an equivalent system of equations. In this case, we described the solution set parametrically .

Next, we defined a $m \times n$ **matrix**, and gave several examples. Here, the matrix has m rows and n columns. We defined the **(main) diagonal** of a matrix. From here, we described how any system of equations

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

has an associated **coefficient matrix**

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

and **augmented matrix**

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \ddots & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix}.$$

Finally, we saw that the operations we had been doing to turn a system of equations into row-echelon form can also be performed on a matrix. These are called **elementary row operations**:

1. Interchange two rows.
2. Multiply a row by a nonzero constant.
3. Add a multiple of one row to another.

Before Tuesday's class, make sure to check out examples in §1.2 of implementing these operations to solve a system of linear equations, with an eye toward our computations earlier today!

Lecture 1: Tuesday, August 21. We began our course by detailing the course syllabus.

Next, we began to discuss **linear equations**. Take any linear equation $y = Mx + B$ in one variable, x, where M and B are constant real numbers–note that we will often use the notation $M, B \in \mathbb{R}$ to indicate this. We can easily transform this equation into one of the form

$$a_1x + a_2y = b$$

(what are a_1, a_2, and b in terms of M, and B?).

This generalizes to the definition of a **linear equation in n variables** $x_1, x_2, \dots, x_n$:

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b$$

where $a_1, a_2, \dots, a_n$, and b are constant real numbers, called **coefficients**; in our new notation, $a_1, a_2, \dots, a_n \in \mathbb{R}$.

So, for example, a linear equation in three variables x, y, and z, has the form

$$a_1x + a_2y + a_3z = b.$$

We went through several examples of linear and non-linear equations. Notably, $e^x + 5y = 1$ is not linear, while $ex + 5y = 1$ is linear, and $\sin(\pi)x - y = 0$ is linear, while $\sin(\pi x) - y = 0$ is not. Neither is $xy = -5$.

We defined a **solution** to a linear equation in the natural way, and defined the **solution set** as the set of all solutions. We described how to describe the solution set of a linear equation **parametrically**, through two examples.

Next, we defined a system of linear equations a finite collection of linear equation. A **solution** of a system is a value that is a solution to each equation, and the **solution set** is defined analogously. A system of m linear equations in n variables is often written using the following convention for subscripts of the coefficients:

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

Considering the equations and graphs of two lines in two variables (what we're used to), we noticed that two linear can either intersect at exactly one point, be equations for the same line (so intersect at infinitely many points), or be parallel (so never intersect). The same phenomenon happens for any system of equations, in any number of variables! In other words, the solution set either consists of **exactly one point**, **infinitely many points**, or **no points** at all.

Finally, we wrote a system of equations in two ways, one which we will call **Row-Echelon form**, and we noticed this form has several advantages, one being that it is much easier to determine how many solutions there are! We will start here next time, showing how to turn any system of equations into this special form.