Lecture 1

January 19, 2022

Linear Equations

A linear equation is an equation of the form

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b \tag{1}$$

where the b is a constant, and each x_i is a different variable with a coefficient a_i.
A system of linear equations is a collection of linear equations. The solution of a system is the ordered list $(s_1, s_2, \ldots, s_n)$ of numbers that satisfy each equation when we plug in $x_1 = s_1, x_2 = s_2, x_3 = s_3, \ldots$. For example, (0, 3, -9) is a solution of the system:

$$\begin{cases} 2x + y - z = 12 \\ 4x - 3z = 27 \end{cases} \tag{2}$$

Because when we plug in the values $(x = 0, y = 3, z = -9)$, both equations are true.

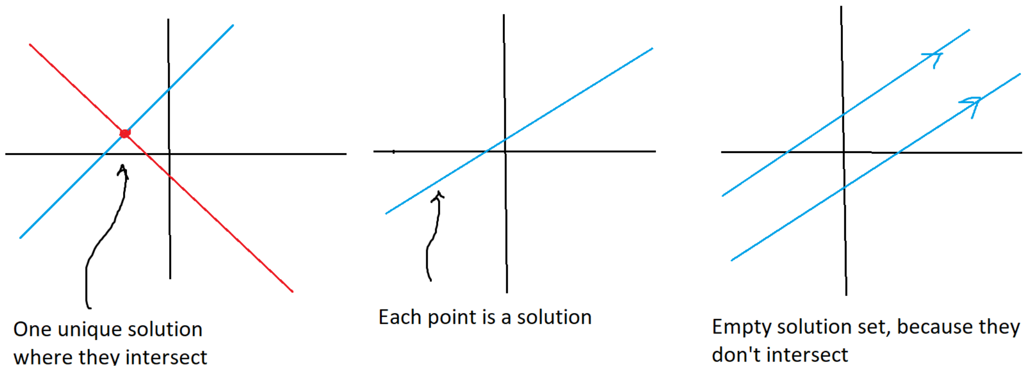
$$\begin{cases} 2(0) + (3) - (-9) = 12 \\ 4(0) - 3 * (-9) = 27 \end{cases} \tag{3}$$

The solution set is the set of all possible solutions to a system.
Two systems of linear equations are equivalent if and only if they have the same solution set. For example,

$$\begin{cases} 2x + 5y = 15 \\ 3x - y = 31 \end{cases} \quad \begin{cases} x = 10 \\ y = -1 \end{cases} \tag{4}$$

are equivalent because both systems have the solution (10,-1).

Geometric picture



The solution set of a single linear equation in two variables is a line. The solution set of a linear system can be a line, a point, or empty.

In general, a system can have zero, one or infinitely many solutions.

Solving Linear Systems

We can abbreviate a system of linear equations with a matrix. For example, the system

$$\begin{aligned} 2x_1 - x_2 + 3x_3 &= 4 \\ 3x_2 - x_3 &= 0 \\ 4x_1 + x_2 - 6x_3 &= 10 \end{aligned}$$

has a coefficient matrix of $\begin{bmatrix} 2 & -1 & 3 \\ 0 & 3 & -1 \\ 4 & 1 & -6 \end{bmatrix}$ and an augmented matrix of $\begin{bmatrix} 2 & -1 & 3 & 4 \\ 0 & 3 & -1 & 0 \\ 4 & 1 & -6 & 10 \end{bmatrix}$
Where, in the coefficient matrix we just take the coefficients on each variable and put them in the matrix in order, so that each equation becomes a row. For the augmented matrix we add a column of the constants from the right side of each equation.
To find a solution set, we use the **Row Operations**:

1. Replace one row by the sum of itself and a multiple of another. $r_i \rightarrow r_i + c * r_j$
2. Interchange two rows. $r_i \longleftrightarrow r_j$

3. Multiply a row by a nonzero constant. $r_i \to c * r_i$

where r_i =row i and r_j =row j.
These operations give an equivalent system of linear equations. For example, we solve the following system using the row operations.

$$2x + 5y = 15$$
$$3x - y = 31$$

$$\begin{bmatrix} 2 & 5 & 15 \\ 3 & -1 & 31 \end{bmatrix}$$
$$r_2 \to r_2 - r_1 \begin{bmatrix} 2 & 5 & 15 \\ 1 & -6 & 16 \end{bmatrix}$$
$$r_1 \longleftrightarrow r_2 \begin{bmatrix} 1 & -6 & 16 \\ 2 & 5 & 15 \end{bmatrix}$$
$$r_2 \to r_2 - 2 * r_1 \begin{bmatrix} 1 & -6 & 16 \\ 0 & 17 & -17 \end{bmatrix}$$
$$r_2 \to 1/17 * r_2 \begin{bmatrix} 1 & -6 & 16 \\ 0 & 1 & -1 \end{bmatrix}$$
$$r_1 \to r_1 + 6 * r_2 \begin{bmatrix} 1 & 0 & 10 \\ 0 & 1 & -1 \end{bmatrix}$$

This corresponds to the system of linear equations

$$x_1 = 10$$
$$x_2 = -1$$

So the solution is (10,-1).

Lecture 2

January 24, 2022

Echelon Form

For a matrix in **echelon form**, each row has a non-zero leading number, which we call the **pivot**. In order for a matrix to be in echelon form, it must satisfy the following conditions:

1. There can never be two pivots in the same column.
2. There are zeros below each pivot.

For example, $\begin{bmatrix} 1 & -6 & 16 \\ 0 & 17 & -17 \end{bmatrix}$ is in echelon form.

Reduced Echelon Form

A matrix is in reduced echelon form if it meets the following conditions:

1. The leading entries are ones.
2. There are zeros above and below each pivot.

For example, $\begin{bmatrix} 1 & 0 & 10 \\ 0 & 1 & -1 \end{bmatrix}$ is in reduced echelon form. This form is the easiest to solve, and our goal is to get any given matrix into this form.

Variables

For a coefficient matrix $\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$ each column represents a variable. The first column gives the coefficients on x_1, the second column is x_2 and the third column is x_3. Given an augmented matrix of a linear system, we call the variables that have a pivot in their column **basic variables**. All other variables are **free variables**. Free variables can take any value, and basic variables are dependent on free variables. For example, in the matrix

$$\begin{bmatrix} \mathbf{1} & 2 & 4 \\ 0 & 0 & 0 \end{bmatrix} \tag{1}$$

The pivot is in the first column so x_1 is the basic variable, and x_2 is the free variable. This matrix corresponds to the equation

$$\begin{aligned} x_1 + 2x_2 &= 4 \\ \implies x_1 &= 4 - 2x_2 \end{aligned}$$

So the solution set is $\{\, (4 - 2x_2, x_2) : x_2 \in \mathbb{R} \,\}$.

Consistency

A linear system is **inconsistent** if it has no solutions, otherwise it is consistent. A linear system is inconsistent if and only if the constant column has a pivot, i.e. there is a row in echelon form that looks like $\begin{pmatrix} 0 & 0 & \dots & 0 & b \end{pmatrix}$. There needs to be a pivot in one of the columns *before* the constant column. For example, the matrix

$$\begin{bmatrix} 1 & 0 & 2 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & 4 \end{bmatrix} \tag{2}$$

has no solutions, because the pivot is in the constant column, so the last row is saying that 0=4.

Unique solutions

A solution is **unique** if it is the only solution that exists. A linear system has a unique solution if and only if it is consistent, and every variable is a basic variable. For example, in

$$\begin{bmatrix} \mathbf{1} & 0 & 3 \\ 0 & \mathbf{1} & 4 \end{bmatrix} \tag{3}$$

x_1 and x_2 are basic variables. So the unique solution is

$$\begin{aligned} x_1 &= 3 \\ x_2 &= 4 \end{aligned}$$

Every matrix is row equivalent to exactly one matrix in reduced echelon form, (the reduced echelon form is unique).

Summary of steps

1. Start with linear system
2. Turn into augmented matrix
3. Use row operations
4. Get to reduced echelon form
5. Solve each equation for the basic variables in terms of free variables.

For example, we will solve the following matrix. We will want the higher pivots to be on the left of the lower pivots.

$$\begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

$$\text{r}_1 \longleftrightarrow r_4 \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & -1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

$$r_2 \to r_2 + r_1,\ r_3 \to r_3 + 2r_1 \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

$$r_3 \to r_3 - 5/2 * r_2,\ r_4 \to r_4 + 3/2 * r_2 \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix}$$

$$r_3 \longleftrightarrow r_4 \begin{bmatrix} \mathbf{1} & 4 & 5 & -9 & -7 \\ 0 & \mathbf{2} & 4 & -6 & -6 \\ 0 & 0 & 0 & \mathbf{-5} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (4)$$

This is in Echelon form. Now to get it into reduced echelon form, we use division to turn the pivots into one, and try to get zeros above each pivot.

$$r_2 \to 1/2 * r_2,\ r_3 \to -1/5 * r_3 \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 1 & 2 & -3 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathrm{r}_1 \to r_1 + 9r_3,\ r_2 \to r_2 + 3r_3 \begin{bmatrix} 1 & 4 & 5 & 0 & -7 \\ 0 & 1 & 2 & 0 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$r_1 \to r_1 - 4r_2 \begin{bmatrix} \mathbf{1} & 0 & -3 & 0 & 5 \\ 0 & \mathbf{1} & 2 & 0 & -3 \\ 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (5)$$

This matrix corresponds to the linear system

$$\begin{cases} x_1 - 3x_3 = 5 \\ x_2 + 2x_3 = -3 \\ x_4 = 0 \end{cases} \tag{6}$$

Which is equivalent to

$$\begin{cases} x_1 = 5 + 3x_3 \\ x_2 = -3 - 2x_3 \\ x_4 = 0 \end{cases} \tag{7}$$

Since each column is a variable, the basic variables are x_1, x_2 and x_4. The free variable is x_3. Each solution is of the form $(x_1,\ x_2,\ x_3,\ x_4)$ so the solution set is $\{\,(5 + 3x_3, -3 - 2x_3, x_3, 0) : x_3 \in \mathbb{R}\,\}$.

We look at another example of using the row operations:

$$\begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ -2 & 4 & 5 & -5 & 3 \\ 3 & -6 & -6 & 8 & -3 \end{bmatrix}$$

$$\text{r}_2 \to r_2 + 2r_1,\ r_3 \to r_3 - 3r_1 \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & -3 & -1 & -3 \end{bmatrix}$$

$$r_3 \to r_3 + r_2 \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$r_2 \to 1/3r_2 \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 1 & 1/3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (8)$$

This is in echelon form. To get it into reduced echelon form, we need to get a zero above the pivot in the second row.

$$r_1 \to r_1 + r_2 \begin{bmatrix} \mathbf{1} & -2 & 0 & 10/3 & 1 \\ 0 & 0 & \mathbf{1} & 1/3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{9}$$

Since the first and third columns have pivots, the basic variables are x_1 and x_3. The free variables are x_2 and x_4. The matrix corresponds to the linear system.

$$\begin{cases} x_1 - 2x_2 + 10/3 * x_4 = 1 \\ x_3 + 1/3 * x_4 = 1 \end{cases} \tag{10}$$

Solving for the basic variables gives

$$\begin{cases} x_1 = 1 + 2x_2 - 10/3 * x_4 \\ x_3 = 1 - 1/3 * x_4 \end{cases} \tag{11}$$

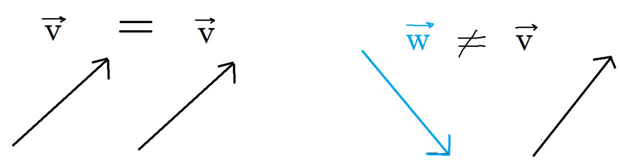
So the solution set is $\{\, (1 + 2x_2 - 10/3 * x_4, x_2, 1 - 1/3 * x_4, x_4) : x_2, x_4 \in \mathbb{R} \,\}$

Lecture 3

January 26, 2022

Vectors and Vector spaces

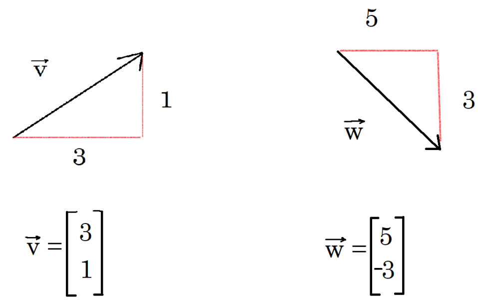
A **vector** is an object that has a magnitude and direction. The **zero vector** has magnitude zero, and is the only vector with no direction. We represent vectors with arrows, and in writing we use the notation $\vec{v}$. The position of the vector is unimportant, two vectors are equal if and only if they have the same magnitude and direction.

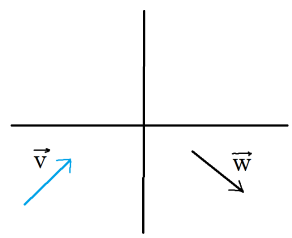


If we draw a vector on a coordinate plane, we can assign it an ordered list of numbers:

$$\vec{v} = \begin{bmatrix} \text{difference in x coordinates} \\ \text{difference in y coordinates} \end{bmatrix} \tag{1}$$

The zero vector = $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$





Two vectors are equal if and only if the corresponding entries are equal. That is,

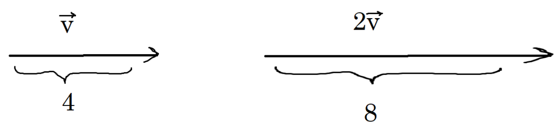
$$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} \tag{2}$$

if and only if $a_1 = a_2$ and $b_1 = b_2$. In n-dimensional space, denoted $\mathbb{R}^n$, we can write a vector as $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ or $(a_1, a_2, \cdots, a_n)$.

Arithmetic with vectors

Scalar multiplication: Let $c > 0$ and $\vec{v}$ be a vector. Then $c\vec{v}$ is a vector in the same direction as $\vec{v}$ and with a magnitude $c|\vec{v}|$, where $|\vec{v}|$ = magnitude of $\vec{v}$.

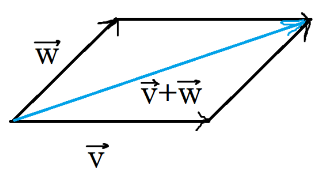
If $c < 0$, then $c|\vec{v}|$ has the opposite direction of $\vec{v}$ and a magnitude of $|c| \cdot |\vec{v}|$. If $c = 0$ then $c\vec{v} = \vec{0}$. Algebraically, if $c \in \mathbb{R}$ and $\vec{v} = \begin{bmatrix} v_1 \\ \dots \\ v_n \end{bmatrix}$ then $c\vec{v} = \begin{bmatrix} c \cdot v_1 \\ \dots \\ c \cdot v_n \end{bmatrix}$.



For example, if $\vec{v} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$ and $c = -2$ then $-2\vec{v} = \begin{bmatrix} -8 \\ 0 \end{bmatrix}$.

Vector addition

The Parallelogram rule tells us that when we add two vectors, the resulting sum is the vector joining the opposite vertices of a parallelogram.



Algebraically, addition is done 'component-wise', that is, we just add the components. So if $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ \cdots \\ w_n \end{bmatrix}$ then

$$\vec{v} + \vec{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \cdots \\ v_n + w_n \end{bmatrix} \tag{3}$$

For example,

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix} + \begin{bmatrix} -6 \\ 7 \end{bmatrix} = \begin{bmatrix} -4 \\ 6 \end{bmatrix} \tag{4}$$

Properties of vectors in $\mathbb{R}^n$

Let $\vec{u}, \vec{v}, \vec{w}$ be vectors in $\mathbb{R}^n$, and $c, d \in \mathbb{R}$. Then,

1. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
2. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
3. $\vec{u} + \vec{0} = \vec{0} + \vec{u} = \vec{u}$
4. $\vec{u} + (-\vec{u}) = (-\vec{u}) + \vec{u} = \vec{0}$
5. $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
6. $(c + d)\vec{u} = c\vec{u} + d\vec{u}$
7. $c(d\vec{u}) = (cd)\vec{u}$
8. $1 \cdot \vec{u} = \vec{u}$

Vector space

In general, a **vector space** is a non-empty collection of objects, called vectors, on which two operations are defined, called scalar multiplication and vector addition, such that the following holds:

1. Properties 1-8 above are true.
9. $\vec{u} + \vec{v}$ should be an element of the vector space.
10. $c\vec{u}$ is an element in the space.

An example of a vector space is $\mathbb{R}^n$. We define addition as $\begin{bmatrix} v_1 \\ \dots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ \dots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ \dots \\ v_n + w_n \end{bmatrix}$. Then we define scalar multiplication as $c \begin{bmatrix} v_1 \\ \dots \\ v_n \end{bmatrix} = \begin{bmatrix} c \cdot v_1 \\ \dots \\ c \cdot v_n \end{bmatrix}$. These definitions satisfy all the properties listed.

Interesting example

An interesting example of a vector space, though not clearly related to vectors, is given:
For $n \geq 0$, let P_n be the set of polynomials of degree at most n with coefficients in $\mathbb{R}$; then the object of our vector space is a polynomial of the form

$$\vec{p}(t) = a_0 + a_1 t + \cdots + a_n t^n \tag{5}$$

where $a_i \in \mathbb{R}$. Addition is defined as follows. Given another vector $\vec{q}(t) = b_0 + b_1 t + \cdots + b_n t^n$,

$$\vec{p}(t) + \vec{q}(t) = (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n \tag{6}$$

And scalar multiplication can be defined as

$$c\vec{p} = ca_0 + ca_1 t + \cdots + ca_n t^n \tag{7}$$

This also satisfies all 10 properties!

Example of something that fails to be a vector space:
If P_n were the set of all polynomials of exactly degree$= n$, it would not be a vector space. An example of how it fails is given:

$$\begin{aligned} p(t) &= 2 + 3t + t^n \\ q(t) &= -1 + 2t^2 - t^{n-1} - t^n \\ p(t) + q(t) &= 1 + 3t + 2t^2 - t^{n-1} \end{aligned}$$

This sum is not an element of the space, because the final term does not have degree n, and so it fails property 9.

Lecture 4

January 31, 2022

Linear Combinations

Let V be a vector space. Given $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_p} \in V$ and $c_1, c_2, \ldots, c_p \in \mathbb{R}$, the vector

$$\vec{y} = c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_p\vec{v_p} \tag{1}$$

is called the **linear combination** of $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_p}$ with **weights** $c_1, c_2, \ldots, c_p$. Note that $\vec{y}$ is in the vector space V. Examples of linear combinations:

$$\sqrt{3}\vec{v_1} + \vec{v_2}, \quad \pi\vec{v_1} + 0 \cdot \vec{v_2}, \quad \vec{0} = 0 \cdot \vec{v_1} + 0 \cdot \vec{v_2} \tag{2}$$

We look at another example, with vectors $\vec{a_1} = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$ $\vec{a_2} = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$.

Question: is $\vec{b}$ a linear combination of $\vec{a_1}$ and $\vec{a_2}$? In other words, do there exist $x_1, x_2 \in \mathbb{R}$ such that

$$x_1\vec{a_1} + x_2\vec{a_2} = \vec{b} \tag{3}$$

$$x_1 \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \tag{4}$$

Using scalar multiplication and vector addition gives

$$\begin{bmatrix} x_1 \\ -2x_1 \\ -5x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 5x_2 \\ 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \tag{5}$$

$$\begin{bmatrix} x_1 + 2x_2 \\ -2x_1 + 5x_2 \\ -5x_1 + 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \tag{6}$$

This is equivalent to the system

$$\begin{cases} x_1 + 2x_2 = 7 \\ -2x_1 + 5x_2 = 4 \\ -5x_1 + 6x_2 = -3 \end{cases} \tag{7}$$

Which has an augmented matrix of

$$\begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \tag{8}$$

Using the row operations, we get the reduced echelon form

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \tag{9}$$

which implies that

$$\begin{cases} x_1 = 3 \\ x_2 = 2 \end{cases} \tag{10}$$

So $\vec{b}$ **is** a linear combination:

$$\vec{b} = 3\vec{a_1} + 2\vec{a_2} \tag{11}$$

In general, to check if $\vec{b}$ is a linear combination of $\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}$ we use the matrix $\begin{bmatrix} \vec{a_1} & \vec{a_2} & \ldots & \vec{a_n} & \vec{b} \end{bmatrix}$, that is, a matrix whose first column is $\vec{a_1}$, second column is $\vec{a_2}$, and so on. The equation $x_1\vec{a_1} + \cdots + x_n\vec{a_n} = \vec{b}$ has the same solution set as the linear system whose augmented matrix is $\begin{bmatrix} \vec{a_1} & \ldots & \vec{a_n} & \vec{b} \end{bmatrix}$.

Span of vectors

If $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n} \in$ V then the set of all linear combinations of $\vec{v_1}, \ldots, \vec{v_n}$ is called the subset of V **spanned** by $\vec{v_1}, \ldots, \vec{v_n}$. We write this set of all linear combinations as

$$\text{span}\{\, \vec{v_1}, \ldots, \vec{v_n} \,\} = \{\, c_1\vec{v_1} + \cdots + c_n\vec{v_n} \mid c_1, \ldots, c_n \in \mathbb{R} \,\} \tag{12}$$

The vector $\vec{b}$ belongs to span$\{\vec{v_1}, \dots, \vec{v_n}\}$ if and only if $\vec{b}$ is a linear combination of $\vec{v_1}, \dots, \vec{v_n}$. The zero vector $\vec{0}$ is always in the span,

$$\vec{0} \in \text{span}\{\vec{v_1}, \dots, \vec{v_n}\} \tag{13}$$

because $\vec{0} = 0 \cdot \vec{v_1} + 0 \cdot \vec{v_2} + \dots$, which is a linear combination.

The following example is given.

$$\vec{a_1} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \quad \vec{a_2} = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix} \tag{14}$$

Question: is $\vec{b} \in \text{span}\{\vec{a_1}, \vec{a_2}\}$? We use these vectors as the columns of an augmented matrix:

$$\begin{bmatrix} 1 & 5 & -3 \\ -2 & -13 & 7 \\ 3 & -3 & 1 \end{bmatrix} \tag{15}$$

and use the row operations to get to

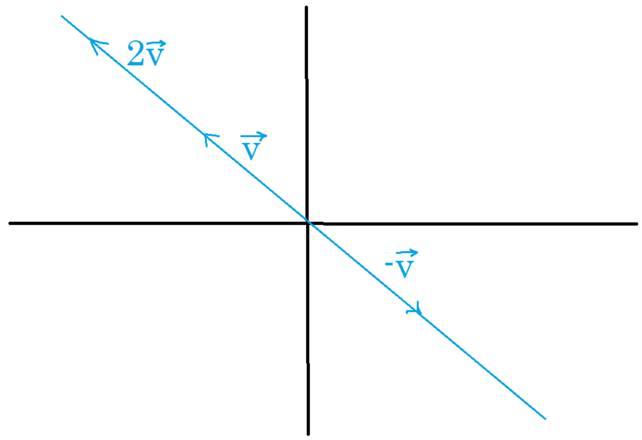
$$\begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & 0 & -2 \end{bmatrix} \tag{16}$$

This system is inconsistent because the last row is saying that $0 = -2$, so there is no solution and $\vec{b}$ is not in span$\{\vec{a_1}, \vec{a_2}\}$.

In general, if there are no solutions to the linear system $\begin{bmatrix} \vec{a_1} & \vec{a_2} & \dots & \vec{a_n} & \vec{b} \end{bmatrix}$ then $\vec{b}$ is **not** in span$\{\vec{a_1}, \dots, \vec{a_n}\}$.

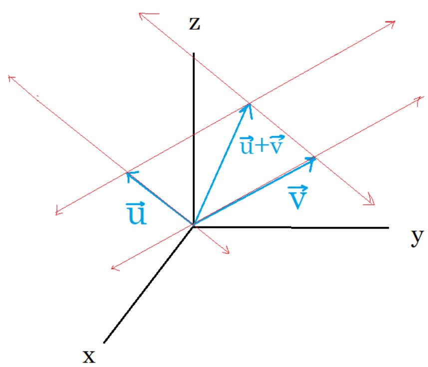
Geometry of span$\{\vec{v}\}$ and span$\{\vec{v}, \vec{u}\}$

In $\mathbb{R}^2$, span$\{\vec{v}\}$ is a line through the origin, for all $\vec{v}$ except $\vec{0}$ because span$\{\vec{0}\} =$ 0.



The vector reaches each point on the line by scalar multiplication, $c \cdot \vec{v}$.

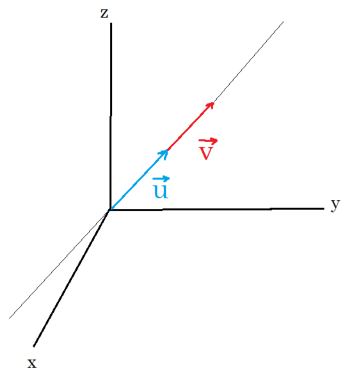
In $\mathbb{R}^3$, span$\{\vec{v}, \vec{u}\}$ is a plane through the origin, but only if $\vec{u}$ and $\vec{v}$ are not linear combinations of each other, $\vec{u} \neq c \cdot \vec{v}$ for some c.



All linear combinations of $\vec{u}$ and $\vec{v}$ are in a plane.

The two vectors are able to reach every point in a plane through vector addition and scalar multiplication.

For the case when $\vec{u} = c \cdot \vec{v}$, the span$\{\vec{v}, \vec{u}\}$ is just a line.



Example problem: Let

$$W = \{ \begin{bmatrix} 5b+2c \\ b \\ c \end{bmatrix} : b, c \in \mathbb{R} \} \tag{17}$$

Find vectors $\vec{u}$ and $\vec{v}$ so that $W = \text{span}\{ \vec{v}, \vec{u} \}$. The vectors are in $\mathbb{R}^3$, and the variables b, c are the weights.

$$\begin{bmatrix} 5b+2c \\ b \\ c \end{bmatrix} = b \cdot \vec{u} + c \cdot \vec{v} \tag{18}$$

We can rewrite the left side as

$$\begin{bmatrix} 5b+2c \\ b \\ c \end{bmatrix} = \begin{bmatrix} 5b \\ b \\ 0 \end{bmatrix} + \begin{bmatrix} 2c \\ 0 \\ c \end{bmatrix} \tag{19}$$

$$\begin{bmatrix} 5b+2c \\ b \\ c \end{bmatrix} = b \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \tag{20}$$

So the vectors are

$$\vec{u} = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix} \quad \vec{v} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \tag{21}$$

Subspaces

A **subspace** of a vector space V is a subset H of V that satisfies:

1. $\vec{0} \in$ H
2. H is closed under addition: if $\vec{u}, \vec{v} \in$ H then $\vec{u} + \vec{v} \in$ H.
3. H is closed under scalar multiplication: if $\vec{u} \in$ H and $c \in \mathbb{R}$ then $c \cdot \vec{u} \in$ H.

These three properties guarantee that H itself is a vector space. (All the other conditions are inherited from V). The following examples are given,

V is a subspace of itself.

$\{\vec{0}\}$ is a subspace of any V.

If $\vec{v_1}, \ldots, \vec{v_n} \in$ V then span$\{\vec{v_1}, \ldots, \vec{v_n}\}$ is a subspace of V, because

a) $\vec{0} \in$ span$\{\vec{v_1}, \vec{v_2}\}$

b) $\vec{v_1} + \vec{v_2} \in$ span$\{\vec{v_1}, \vec{v_2}\}$ since $\vec{v_1} + \vec{v_2}$ is a linear combination.

c) $c \cdot \vec{v_1} \in$ span$\{\vec{v_1}, \vec{v_2}\}$ since $c \cdot \vec{v_1}$ is also a linear combination.

Lecture 5

February 2, 2022

Definition: If A is an $m \times n$ matrix $\begin{bmatrix} \vec{a_1} & \vec{a_2} & \dots & \vec{a_n} \end{bmatrix}$ and $\vec{x} \in \mathbb{R}^m$, then the product of A and $\vec{x}$, written $A\vec{x}$, is the linear combination of the columns of A with weights given by $\vec{x}$:

$$A\vec{x} = \begin{bmatrix} \vec{a_1} & \vec{a_2} & \cdots & \vec{a_n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} \tag{1}$$

$$= x_1\vec{a_1} + x_2\vec{a_2} + x_3\vec{a_3} + \cdots + x_n\vec{a_n} \tag{2}$$

For example,

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \cdot \begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7 \cdot \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$
$$= \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix}$$
$$= \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

Note that

$$\begin{bmatrix} 3 \\ 6 \end{bmatrix} \in \text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -5 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \end{bmatrix} \right\}$$

because it is a linear combination of those vectors.

For matrix multiplication, we always need the number of columns in the first matrix to match the number of rows in the second matrix, otherwise it won't work.

In the example above, the left matrix has 3 columns and the right matrix has 3 rows, so it works.

Equivalent statements: Each of these is true if and only if the others are true.

1. $\vec{b} \in \text{Span}\{\, \vec{a_1}, \cdots, \vec{a_n} \,\}$
2. There exist $x_1, x_2, \cdots, x_n \in \mathbb{R}$ such that $\vec{b} = x_1\vec{a_1} + x_2\vec{a_2} + \cdots + x_n\vec{a_n}$.
3. There is a solution $(x_1, x_2, \cdots, x_n)$ to the linear system whose augmented matrix is $\begin{bmatrix} \vec{a_1} & \vec{a_2} & \cdots & \vec{a_n} & \vec{b} \end{bmatrix}$.
4. The matrix equation $A\vec{x} = \vec{b}$ has a solution.

This means that $A\vec{x} = \vec{b}$ has a solution if and only if $\vec{b} \in \text{Span}\{\, \vec{a_1}, \cdots, \vec{a_n} \,\}$.

Column spaces

The **column space** of an $m \times n$ matrix, written Col A, is the set of all linear combinations of the columns of A. So if $A = \begin{bmatrix} \vec{a_1} & \cdots & \vec{a_n} \end{bmatrix}$ then

$$\text{Col A} = \text{Span}\{\, \vec{a_1}, \cdots, \vec{a_n} \,\}$$

.

Note that Col A is a subspace of $\mathbb{R}^m$, (because $\vec{a_1}, \cdots, \vec{a_n} \in \mathbb{R}^m$). Also note that we can write the definition of a column space as

$$\text{Col A} = \left\{ \vec{b} \in \mathbb{R}^m \colon A\vec{x} = \vec{b} \text{ for some } \vec{x} \in \mathbb{R}^n \right\} \tag{3}$$

The following example is given. Find a matrix A such that W$= \left\{ \begin{bmatrix} 6a - b \\ a + b \\ -7a \end{bmatrix} : a, b \in \mathbb{R} \right\}$ is equal to Col A. Since a column space is a linear combination of vectors, we're looking for the vectors that we can use as the columns in A. So we separate a and b, and use them as weights:

$$\begin{bmatrix} 6a-b \\ a+b \\ -7a \end{bmatrix} = \begin{bmatrix} 6a \\ a \\ -7a \end{bmatrix} + \begin{bmatrix} -b \\ b \\ 0 \end{bmatrix} \tag{4}$$

$$a \cdot \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix} + b \cdot \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \tag{5}$$

This is a linear combination, so $W = \text{Span}\left\{ \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$ which is the column space of the matrix

$$A = \begin{bmatrix} 6 & -1 \\ 1 & 1 \\ -7 & 0 \end{bmatrix} \tag{6}$$

That's our answer.

Question: when is Col A all of $\mathbb{R}^m$? This happens when the equation $A\vec{x} = \vec{b}$ has a solution for *all* $\vec{b} \in \mathbb{R}^m$.

Theorem: Let A be an $m \times n$ matrix, then the following are logically equivalent statements (if one is true, they're all true).

1. For each $\vec{b} \in \mathbb{R}^m$, $A\vec{x} = \vec{b}$ has a solution.
2. For each $\vec{b} \in \mathbb{R}^m$, $\vec{b}$ is a linear combination of the columns of A.
3. Col A $= \mathbb{R}^m$. (The column space fills the whole space).
4. The columns of A span $\mathbb{R}^m$, (so if A$=\begin{bmatrix} \vec{a_1} & \dots & \vec{a_n} \end{bmatrix}$, then Span$\{\vec{a_1}, \dots, \vec{a_n}\} = \mathbb{R}^m$).
5. The matrix A has a pivot position in every row. (This is the easiest one to check, and we can use it to verify the others).

The following example is given. Let $A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix}$. Is Col A $= \mathbb{R}^3$? Is it true that A$\vec{x} = \vec{b}$ has a solution for every $\vec{b} \in \mathbb{R}^3$? The augmented matrix here for an unknown $\vec{b}$ is

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix}$$

Using the row operations we get this to

$$\begin{bmatrix} \mathbf{1} & 3 & 4 & b_1 \\ 0 & \mathbf{14} & 10 & b_2 + 4b_1 \\ 0 & 0 & 0 & b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) \end{bmatrix}$$

This does not have a pivot position in row 3, so by the theorem, Col A $\neq \mathbb{R}^3$. This means there must be a vector $\vec{b} \in \mathbb{R}^3$ such that A$\vec{x} = \vec{b}$ has no solution, because from the augmented matrix of A, if the last element $b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1)$ is non-zero then the last row would be saying that $0 = c$ for a nonzero c. One example of $\vec{b}$ not in the column space is $\vec{b} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, because $2 + 3(1) - \frac{1}{2}(0 + 4(1)) = 2 + 3 - 2 = 3 \neq 0$, so the system is inconsistent.

Properties of $A\vec{x}$

Let A be an $m \times n$ matrix, with $\vec{u}, \vec{v} \in \mathbb{R}^m$ and $c \in \mathbb{R}$:

1. $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$
2. $A(c\vec{u}) = c(A\vec{u})$

Computing $A\vec{x}$:

$$
\begin{aligned}
A\vec{x} &= \begin{bmatrix} 1 & 3 & 7 \\ 2 & 4 & -1 \\ 0 & -1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix} \\
= 2 \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + 1 \cdot \begin{bmatrix} 3 \\ 4 \\ -1 \end{bmatrix} &+ 5 \cdot \begin{bmatrix} 7 \\ -1 \\ 3 \end{bmatrix} \\
= \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \\ -1 \end{bmatrix} &+ \begin{bmatrix} 35 \\ -5 \\ 15 \end{bmatrix} \\
&= \begin{bmatrix} 40 \\ 3 \\ 14 \end{bmatrix}
\end{aligned}
$$

This works, but a more efficient way is just to distribute each entry in $\vec{x}$ to the columns of A in your head.

$$
\begin{bmatrix} 1 \cdot 2 & + & 3 \cdot 1 & + & 7 \cdot 5 \\ 2 \cdot 2 & + & 4 \cdot 1 & + & (-1) \cdot 5 \\ 0 \cdot 2 & + & (-1) \cdot 1 & + & 3 \cdot 5 \end{bmatrix} = \begin{bmatrix} 40 \\ 3 \\ 14 \end{bmatrix}
$$

Lecture 6

February 7, 2022

Null space of a matrix

A linear system is **homogenous** if it can be written as $A\vec{x} = \vec{0}$, (where A is an $m \times n$ matrix, and $\vec{0} \in \mathbb{R}^n$). The vector $\vec{x} = \vec{0}$ is always a solution, so we call it the **trivial solution**.

The **null space** of an $m \times n$ matrix is the set of all solutions to the equation $A\vec{x} = \vec{0}$. It's denoted

$$\text{Nul A} = \left\{ \vec{x} \in \mathbb{R}^n \colon A\vec{x} = \vec{0} \right\} \tag{1}$$

where A is $m \times n$ and $\vec{x}$ is $n \times 1$, so that the number of rows in $\vec{x}$ match the number of columns in A. Basically, it's the set of vectors $\vec{x}$ where the linear system is equal to zero.

For example, let $A = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix}$. Is $\vec{x} = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$ in Nul A?

$$A\vec{x} = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$$

$$= 5 \cdot \begin{bmatrix} 1 \\ -5 \end{bmatrix} + 3 \cdot \begin{bmatrix} -3 \\ 9 \end{bmatrix} - 2 \cdot \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 5 - 9 + 4 \\ -25 + 27 - 2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}$$

So $\vec{x} \in$ Nul A.

Theorem: Nul A is a subspace of $\mathbb{R}^n$, (where n is the number of columns). Why? Check the 3 conditions for a subspace:

1. $\vec{0} \in$ Nul A because $A\vec{0} = \vec{0}$.

2. If $\vec{x}, \vec{y} \in$ Nul A, then $\vec{x}+\vec{y} \in$ Nul A. This is true because $A(\vec{x}+\vec{y}) = A\vec{x}+A\vec{y} = \vec{0}+\vec{0} = \vec{0}$, which is a given since $\vec{x}, \vec{y} \in$ Nul A means that both $A\vec{x}$ and $A\vec{y}$ are zero.

3. If $\vec{x} \in$ Nul A and $c \in \mathbb{R}$, we want to check if $c\vec{x} \in$ Nul A.
$A(c\vec{x}) = c(A\vec{x}) = c \cdot \vec{0} = \vec{0}$. So it works.

How do we find Nul A?
We describe Nul A by writing it as the span of a collection of vectors, (just like we did for Col A).
For example, find a spanning set of Nul A, where

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

Note that $\vec{x}$ will be in $\mathbb{R}^5$ because there are 5 columns. Now, take the augmented matrix for $A\vec{x} = \vec{0}$

$$\begin{bmatrix} -3 & -6 & -1 & 1 & -7 & 0 \\ 1 & -2 & 2 & 3 & -1 & 0 \\ 2 & -4 & 5 & 8 & -4 & 0 \end{bmatrix}$$

$$\text{row operations} \rightarrow \begin{bmatrix} \mathbf{1} & -2 & 0 & -1 & -3 & 0 \\ 0 & 0 & \mathbf{1} & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Columns 1 and 3 have pivots, so x_1 and x_3 are basic variables, and x_2, x_4, x_5 are free. This corresponds to the system

$$\begin{cases} x_1 - 2x_2 - x_4 - 3x_5 = 0 \\ x_3 + x_4 - 2x_5 \end{cases}$$

$$\implies \begin{cases} x_1 = 2x_2 + x_4 + 3x_5 \\ x_3 = 2x_5 - 2x_4 \end{cases}$$

The vectors $\vec{x}$ will be of the form $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$. So

$$\vec{x} \in \left\{ \begin{bmatrix} 2x_2 + x_4 + 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} : x_2, x_4, x_5 \in \mathbb{R} \right\}$$

This is Nul A. We want to put this in the form of a linear combination, with the weights given by the free variables, so we'll look at the coefficents on each variable and separate the vector out.

$$\begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \tag{2}$$

So Nul A is the set of all linear combinations of these vectors, which means

$$\text{Nul A} = \text{Span} \left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\} \tag{3}$$

Steps to find Nul A:

1. Write the augmented matrix of $A\vec{x} = \vec{0}$.

2. Use the row operations

3. Find free and basic variables

4. Solve for the basic variables in terms of free variables

5. Express the solution set as a linear combination of vectors.

Remarks

1. The number of vectors in a spanning set for Nul A is the number of free variables.
2. $\vec{0}$ is always in Nul A.
3. Nul A contains non-trivial (nonzero) vectors if and only if the equation has at least one free variable, (because otherwise all variables are equal to zero and the only solution is $\vec{0}$).

Distinguishing Nul A vs. Col A

Let $A = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$. The questions are

1. Find k such that
   (a) Nul A is a subspace of $\mathbb{R}^k$.
   (b) Col A is a subspace of $\mathbb{R}^k$.

   The answers are

   (a) $k = 4$ because there are 4 columns in A, which means there are 4 variables, so $\vec{x}$ has 4 entries and $\vec{x} \in \mathbb{R}^4$. In general, the number of dimensions k in Nul A is the number of **columns** in A.
   (b) $k = 3$, because in the column space, each vector is just one of the columns and each column has 3 coordinates (one for each row). So in general, the number of dimensions k in Col A is the number of **rows** in A.

2. Find a non-zero vector in:
   (a) Nul A. The null space is Nul A$=\left\{ \vec{x} \in \mathbb{R}^4 \colon A\vec{x} = \vec{0} \right\}$. We start by using the augmented matrix.

$$\begin{bmatrix} 2 & 4 & -2 & 1 & 0 \\ -2 & -5 & 7 & 3 & 0 \\ 3 & 7 & -8 & 6 & 0 \end{bmatrix}$$
$$\text{row reduce} \rightarrow \begin{bmatrix} \mathbf{1} & 0 & 9 & 0 & 0 \\ 0 & \mathbf{1} & -5 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 \end{bmatrix}$$

Columns 1, 2 and 4 have pivots so x_1, x_2 and x_4 are basic variables. x_3 is free. Just to find one solution, choose $x_3 = 1$,

$$\begin{cases} x_1 + 9x_3 = 0 \\ x_2 - 5x_3 \\ x_4 = 0 \end{cases} \implies \begin{cases} x_1 = -9x_3 = -9 \\ x_2 = 5x_3 = 5 \\ x_4 = 0 \end{cases} \tag{4}$$

So the solution is $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -9 \\ 5 \\ 1 \\ 0 \end{bmatrix} \in \text{Nul A}.$

(b) Col A is just the space spanned by the columns, so any linear combination of the column vectors is in Col A, and we can just pick any column.

$$\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix} \in \text{Col A} \tag{5}$$

Note: this solution is not unique.

3. If $\vec{u} = \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$.

(a) Is $\vec{u} \in$ Nul A? Is $\vec{u} \in$ Col A?
Answer: $\vec{u} \notin$ Col A because it's the wrong size ($\vec{u} \in \mathbb{R}^4$, but Col A is a subspace of $\mathbb{R}^3$).
Check Nul A:

$$A\vec{u} = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & 5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 3 \end{bmatrix} \neq \vec{0} \tag{6}$$

So $\vec{u} \notin$ Nul A.

(b) Is $\vec{v} \in$ Nul A? Is $\vec{v} \in$ Col A?
(left as an exercise)

Lecture 7

February 9, 2022

Review

A homogenous equation is one of the form $A\vec{x} = \vec{0}$. Warm up exercise:

If $A = \begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & 8 \end{bmatrix}$ solve $A\vec{x} = \vec{0}$. Since the null space is the set of all vectors $\vec{x}$ that solve $A\vec{x} = \vec{0}$, we just need to find Nul A.

$$\begin{bmatrix} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & 8 & 0 \end{bmatrix} \to \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & -9 & 16 & 0 \end{bmatrix}$$
$$\to \begin{bmatrix} \mathbf{1} & 0 & -4/3 & 0 \\ 0 & \mathbf{1} & 0 & 0 \end{bmatrix}$$

This matrix corresponds to the linear system

$$\begin{cases} x_1 - \frac{4}{3}x_3 = 0 \\ x_2 = 0 \end{cases} \implies \begin{cases} x_1 = \frac{4}{3}x_3 \\ x_2 = 0 \end{cases}$$

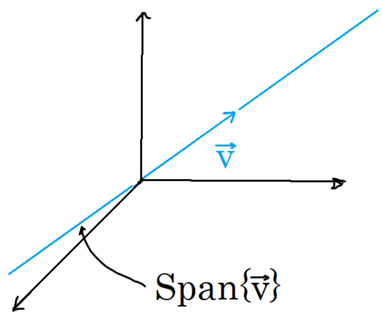
There is no pivot in the third column, so x_3 is free. These are the entries in the vector $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. So the solution set is

$$\left\{ \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} : x_3 \in \mathbb{R} \right\} \tag{1}$$

If we want to write this in terms of a span, we factor out the variables from $\vec{x}$.

$$\begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = x_3 \cdot \begin{bmatrix} 4/3 \\ 0 \\ 1 \end{bmatrix}$$

This is just a linear combination of a single vector, so Nul A=Span$\left\{ \begin{bmatrix} 4/3 \\ 0 \\ 1 \end{bmatrix} \right\}$. Different multiples of this vector give examples of $\vec{x}$ that solve $A\vec{x} = \vec{0}$. Geometrically, this is a line in $\mathbb{R}^3$.



Non-homogenous system of equations
A linear system $A\vec{x} = \vec{b}$ is **non-homogenous** if $\vec{b} \neq \vec{0}$. For example, if $A = \begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & 8 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 7 \\ 1 \\ -4 \end{bmatrix}$. Describe all solutions to $A\vec{x} = \vec{b}$. The augmented matrix is

$$\begin{bmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & 1 \\ 6 & 1 & 8 & -4 \end{bmatrix}$$

$$\text{row reduce} \rightarrow \begin{bmatrix} \mathbf{1} & 0 & -4/3 & -1 \\ 0 & \mathbf{1} & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\text{This corresponds to } \begin{cases} x_1 - \frac{4}{3}x_3 = -1 \\ x_2 = 2 \end{cases} \implies \begin{cases} x_1 = \frac{4}{3}x_3 - 1 \\ x_2 = 2 \\ x_3 \text{ is free} \end{cases}$$

So the solution set is
$$\left\{ \begin{bmatrix} \frac{4}{3}x_3 - 1 \\ 2 \\ x_3 \end{bmatrix} : x_3 \in \mathbb{R} \right\} \tag{2}$$
Then we factor out the variables in $\vec{x}$.
$$\vec{x} = \begin{bmatrix} \frac{4}{3}x_3 - 1 \\ 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 4/3 \\ 0 \\ 1 \end{bmatrix}$$
and the solution set can be written as
$$\left\{ \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 4/3 \\ 0 \\ 1 \end{bmatrix} : x_3 \in \mathbb{R} \right\} \tag{3}$$

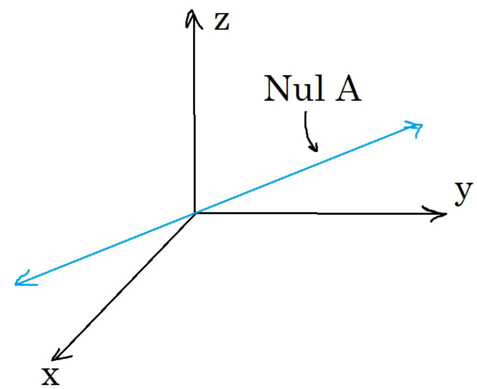
Notice that the only difference between this solution and the solution from the homogenous case (1) is the constant vector. So the right vector represents the null space. The left vector is one particular solution, because x_3 can be any real number, so if we look at $x_3 = 0$ then the right vector becomes zero and we're left with one example of a solution: $\vec{x} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$ such that $A\vec{x} = \vec{b}$.

So to find all the solutions to a non-homogenous system, we just need to find one particular solution and then add Nul A.

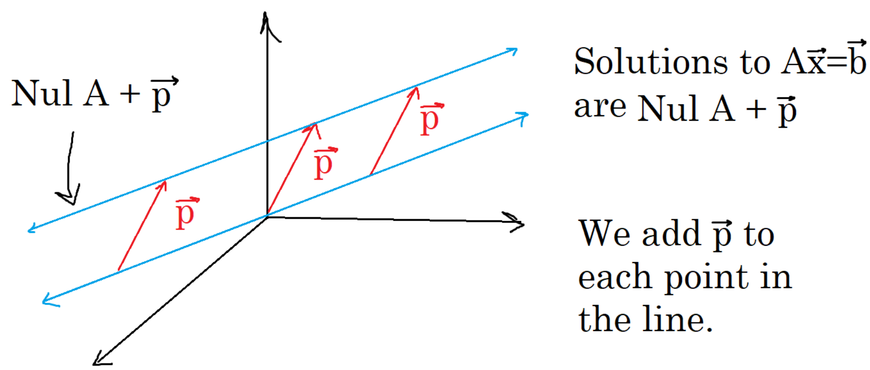
Theorem: Suppose $A\vec{x} = \vec{b}$ is consistent (has a solution) for a given $\vec{b}$ and let $\vec{p}$ be a solution. Then the solution set to $A\vec{x} = \vec{b}$ is the set of all vectors $\vec{w} = \vec{p} + \vec{v_h}$, where $\vec{v_h}$ is a solution to the homogenous equation $A\vec{x} = \vec{0}$, which is the same as saying $\vec{v_n} \in$ Nul A. Equivalently, we can write the solution set as
$$\vec{p} + \text{Nul A} \tag{4}$$
or
$$\{\, \vec{p} + \vec{v_h} : \vec{v_h} \in \text{Nul A} \,\} \tag{5}$$
We can use any $\vec{p}$ that happens to be a solution, but one sure way to get it is to follow the method above, where $\vec{p}$ is the constant vector.

Geometric picture

Nul A can be a line (if there's one free variable) or a plane (if there are 2 free variables).



So the solution to $A\vec{x} = \vec{b}$ looks like



The solutions to a non-homogenous equation are just the solutions to the homogenous equation translated by a vector $\vec{p}$. This applies to planes as well as lines. For example, let

$$A = \begin{bmatrix} 1 & 3 & -5 \\ 1 & 4 & -8 \\ -3 & -7 & 9 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} \tag{6}$$

Find one solution to $A\vec{x} = \vec{b}$. The vector $\vec{b}$ is exactly the first column, so one solution is $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Because then

$$\begin{aligned} A\vec{x} &= \begin{bmatrix} 1 & 3 & -5 \\ 1 & 4 & -8 \\ -3 & -7 & 9 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= 1 \cdot \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} + 0 \cdot \begin{bmatrix} 3 \\ 4 \\ -7 \end{bmatrix} + 0 \cdot \begin{bmatrix} -5 \\ -8 \\ 9 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} = \vec{b} \end{aligned}$$

This is one particular solution that happened to be easy to find.

Linear Independence

A set of vectors $\{\vec{v_1}, \cdots, \vec{v_p}\}$ in a vector space V is **linearly independent** if

$$c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_p\vec{v_p} = \vec{0} \tag{7}$$

only has the trivial solution

$$c_1 = c_2 = \cdots = c_p = 0$$

The set $\{\vec{v_1}, \cdots, \vec{v_p}\}$ is **linearly dependent** if equation (7) has a non-trivial solution, (at least one coefficent is nonzero).

The following example is given. Consider a set containing a single vector $\{\vec{v}\}$. Then $c\vec{v} = \vec{0}$ means that

$$\begin{cases} \text{If } \vec{v} \neq \vec{0} \text{ then c has to be zero} \\ \text{If } \vec{v} = \vec{0} \text{ then c is any real number} \end{cases}$$

So $\{\vec{v}\}$ is linearly independent if $\vec{v} \neq \vec{0}$, because then the coefficient needs to be zero, and the only solution is the trivial solution.

Another example: If $V = \mathbb{R}^n$ and $A = \begin{bmatrix} \vec{v_1} & \cdots & \vec{v_p} \end{bmatrix}$, then the set $\{\vec{v_1}, \cdots, \vec{v_p}\}$

being linearly independent is equivalent to saying $A\vec{x} = \vec{0}$ only has the solution $\vec{x} = \vec{0}$. We can see this as follows. If $A = \begin{bmatrix} \vec{v_1} & \cdots & \vec{v_p} \end{bmatrix}$ and $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_p \end{bmatrix}$ then

$$A\vec{x} = x_1\vec{v_1} + x_2\vec{v_2} + \cdots + x_p\vec{v_p} \tag{8}$$

So $A\vec{x} = \vec{0}$ is the same thing as

$$x_1\vec{v_1} + \cdots + x_p\vec{v_p} = 0 \tag{9}$$

This meets the form of equation (7), so these vectors are linearly independent if the only solution to this is $x_1 = x_2 = \cdots = x_p = 0$, which means that $\vec{x} = \vec{0}$.
In general, given a matrix A, the columns of A are linearly independent exactly when Nul A= $\{\vec{0}\}$.

Theorem: A set $\{\vec{v_1}, \cdots, \vec{v_p}\}$ of two or more vectors with $\vec{v_1} \neq \vec{0}$, is linearly dependent if and only if some $\vec{v_j}$ with $j > 0$ is a linear combination of the preceding vectors $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_{j-1}}$.
For example, are the following vectors lienarly indendent?

1. $\vec{v_1} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\vec{v_2} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$. Well $\vec{v_2} = 2\vec{v_1}$ so the answer is no.

2. $\vec{v_1} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\vec{v_2} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$. The answer is yes, because $\vec{v_2} \neq c\vec{v_1}$ for any $c \in \mathbb{R}$.

Lecture 8

February 14, 2022

Linear Transformations

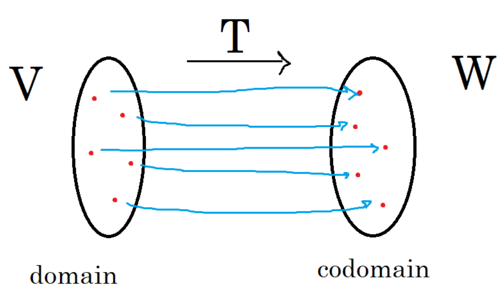
A linear transformation T from a vector space V to a vector space W is a rule (or map) that assigns to each vector $\vec{x} \in V$ a unique vector $T(\vec{x}) \in W$ such that the following conditions are true.

1. $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$ for all $\vec{u}, \vec{v} \in V$.
2. $T(c\vec{u}) = c \cdot T(\vec{u})$ for all $\vec{u} \in V, c \in \mathbb{R}$.

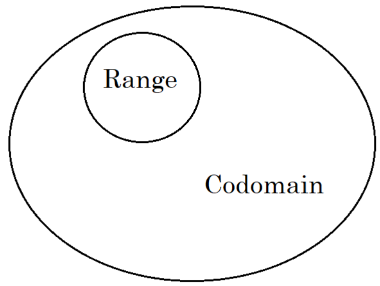
Informally, we can say that T respects the vector space's structure.

The vector space V is the **domain** of T, and W is the **codomain**. The **kernel** of T is the set of all vectors $\vec{x} \in V$ such that $T(\vec{x}) = \vec{0}$. (where $\vec{0}$ is in W). (The kernel is sometimes called the null space of the map).

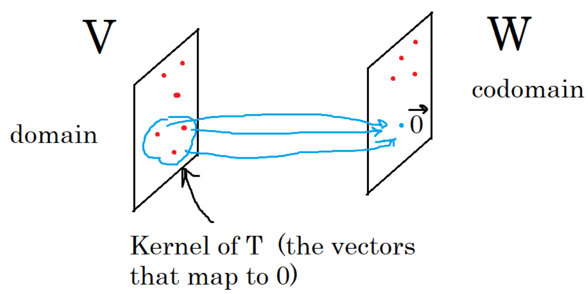
The **range** of T is the set of all $\vec{y} \in W$ such that there exists $\vec{x} \in V$ for which $T(\vec{x}) = \vec{y}$.



The range is the set of vectors that V gets mapped to, and it is a subspace of the codomain.



Using these pictures, we can imagine the kernel of T as the set of points from V that map to the zero vector in W:



Analogy
For the function $f(x) = x^2$,

1. The Domain is $\mathbb{R}$. (This is the set of all values x could be)
2. The Codomain is $\mathbb{R}$ (This is the set of all y)
3. The Range is $[0, +\infty)$ since x^2 is positive for all $x \in \mathbb{R}$.
4. The Kernel is $\{0\}$ since $0^2 = 0$.

It's the same for linear transformations: we take a vector from the domain, map it to a vector in the codomain, and the set of all possible vectors we could get is the range. The kernel is just the vectors that get mapped to $\vec{0}$.

A matrix as a transformation
Given any $m \times n$ matrix A, the map from $\mathbb{R}^n$ to $\mathbb{R}^m$, written $T : \mathbb{R}^n \to \mathbb{R}^m$, and given by $T(\vec{x}) = A\vec{x}$ is a linear transformation. Why? Check the conditions.

1. Let $\vec{u}, \vec{v} \in \mathbb{R}^n$, then

$$\begin{aligned} T(\vec{x}+\vec{y}) &= A(\vec{x}+\vec{y}) \\ &= A\vec{x}+A\vec{y} \\ &= T(\vec{x})+T(\vec{y}) \end{aligned}$$

   So this meets the first condition for a linear transformation.

2. Let $\vec{x} \in \mathbb{R}^n$, and $c \in \mathbb{R}$, then

$$\begin{aligned} T(c\vec{x}) &= A(c\vec{x}) \\ &= c \cdot A\vec{x} \\ &= c \cdot T(\vec{x}) \end{aligned}$$

   So the second condition is also met.

Now we want to look at some of the properties of T.

1. The kernel is the set of vectors that get mapped to zero, so it is

$$\begin{aligned} \text{Kernal of T} &= \left\{ \vec{x} \in \mathbb{R}^n \colon T(\vec{x}) = \vec{0} \right\} \\ &= \left\{ \vec{x} \in \mathbb{R}^n \colon A\vec{x} = \vec{0} \right\} \\ &= \text{Nul A} \end{aligned}$$

   Because a null space is the set of all $\vec{x}$ such that $A\vec{x} = \vec{0}$. So the kernel is the null space.

2. The range is the set of all vectors equal to $A\vec{x}$. So

$$\text{Range of T} = \{ \vec{y} \in \mathbb{R}^m \colon \text{ there exists } \vec{x} \in \mathbb{R}^n \text{ such that } A\vec{x} = \vec{y} \} \tag{1}$$

   Since $A = \begin{bmatrix} \vec{a_1} \dots \vec{a_n} \end{bmatrix}$ and $\vec{x} = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}$, we can write $\vec{y}$ as the linear combination

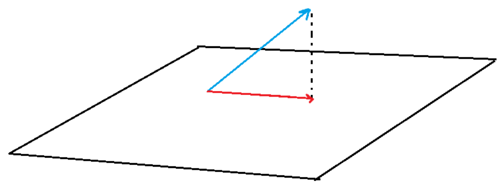
$$A\vec{x} = x_1\vec{a_1} + x_2\vec{a_2} + \cdots + x_n\vec{a_n} \tag{2}$$

   So we're looking at the set of $\vec{y}$ such that we can write it as a linear combination of the columns of A. This is the column space.

$$\text{Range of T} = \text{Col A} \tag{3}$$

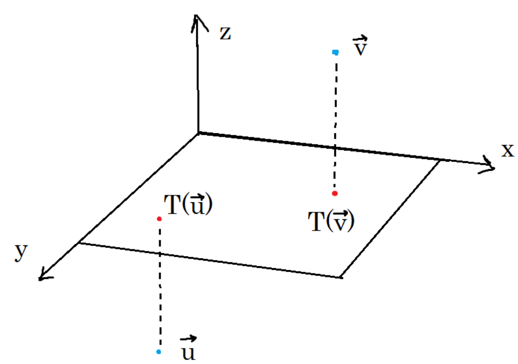
Linear transformation example: "Projection"

A projection is a map from $\mathbb{R}^3$ to $\mathbb{R}^3$, $T\colon \mathbb{R}^3 \to \mathbb{R}^3$, where the range is a plane.



For the map, we just get rid of the z-value.

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \tag{4}$$



We can do this linear transformation using a matrix. Specifically,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ such that } T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = A \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{5}$$

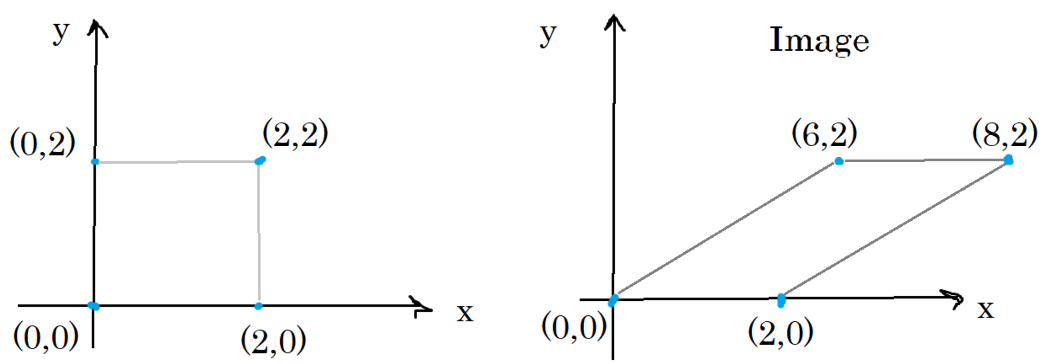
because then

$$\begin{aligned} A \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\ &= \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \end{aligned}$$

Another example: "Shear":
One shear map from $\mathbb{R}^2$ to $\mathbb{R}^2$. $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ is defined by

$$A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \text{ such that } T(\vec{u}) = A\vec{u} \tag{6}$$

The map is $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 3y \\ y \end{bmatrix}$. Geometrically, for a square it looks like



because when we apply the linear transformation to each of the points in the square we get

$$\begin{aligned}T((0,0)) &= (0,0)\\T((2,0)) &= (2,0)\\T((2,2)) &= (8,2)\\T((0,2)) &= (6,2)\end{aligned}$$

Another example: the "Derivative"

Let V be the vector space of all real-valued functions defined on an interval [a,b] that are differentiable and have continuous derivatives. Let W be the vector space of continuous real-valued functions from [a,b] to $\mathbb{R}$. Let $D\colon V \to W$ be defined by

$$D(f) = f' \tag{7}$$

Then, D is a linear transformation. Why? Check the conditions.

1.

$$\begin{aligned}D(f+g) &= (f+g)'\\&= f' + g'\\&= D(f) + D(g)\end{aligned}$$

2.

$$\begin{aligned}D(cf) &= (cf)'\\&= c \cdot f'\\&= c \cdot D(f)\end{aligned}$$

So this is a linear transformation.

Properties of Linear Transormations

Let $T\colon V \to W$ be a linear transformation, then

1. $T(\vec{0}) = \vec{0}$

   Because $\vec{0} = 0 \cdot \vec{u}$ for any $\vec{u} \in$ V, so $T(\vec{0}) = T(0 \cdot \vec{u}) = 0 \cdot T(\vec{u}) = \vec{0}$.

2. $T(c\vec{u} + d\vec{v}) = c \cdot T(\vec{u}) + d \cdot T(\vec{v})$

Example problem

Let T be the linear transformation defined by $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$ such that $T(\vec{x}) = A\vec{x}$.

a. Find m and n so that $T\colon \mathbb{R}^n \to \mathbb{R}^m$.
   We need an $\vec{x}$ with 2 entries to match the number of columns.

$$\begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix} \tag{8}$$

   So $n = 2$, because that's our starting vector. Then the final vector has 3 rows, so $m = 3$.

b. Find $T(\vec{u})$ if $\vec{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

$$T(\vec{u}) = A\vec{u} \quad \text{so it is} \quad \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 - 3 \cdot (-1) \\ 3 \cdot 2 + 5 \cdot (-1) \\ (-1) \cdot 2 + 7 \cdot (-1) \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$$

c. if $\vec{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, find $\vec{x} \in \mathbb{R}^2$ so that $T(\vec{x}) = \vec{b}$.
   The augmented matrix of $A\vec{x} = \vec{b}$ is

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & -5 \end{bmatrix} \text{ row reduce } \to \begin{bmatrix} 1 & 0 & 3/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\text{This corresponds to } \begin{cases} x_1 = 3/2 \\ x_2 = -1/2 \end{cases} \quad \text{so } \vec{x} = \begin{bmatrix} 3/2 \\ -1/2 \end{bmatrix}$$

Lecture 9

February 16, 2022

More on linear transformations

Given the linear transformation $T(\vec{x}) = A\vec{x}$, where

$$A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$$

If $A\vec{x} = \vec{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$ what is the **preimage** of $\vec{b}$ under T? That is, what is the collection of vectors $\vec{x} \in \mathbb{R}^2$ such that $A\vec{x} = \vec{b}$, (we're looking for *all* $\vec{x}$). To do this, we just solve for $\vec{x}$ in the linear system $A\vec{x} = \vec{b}$.

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & -5 \end{bmatrix} \to \begin{bmatrix} \mathbf{1} & 0 & 3/2 \\ 0 & \mathbf{1} & -1/2 \\ 0 & 0 & 0 \end{bmatrix} \tag{1}$$

So $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3/2 \\ -1/2 \end{bmatrix}$. This solution is unique, so the preimage only consists of this solution. The preimage of $\vec{b}$ is $\left\{ \begin{bmatrix} 3/2 \\ -1/2 \end{bmatrix} \right\}$.

We could also ask the question, if $\vec{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$, is $\vec{c}$ in the range of T? Equivalently, is $\vec{c} \in$ Col A? (This is equivalent because the range of T is the set of all $\vec{b}$ you could get from $A\vec{x}$, which is just the column space). To see if $\vec{b}$ is in the column space, we put it in the augmented matrix

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & 5 \end{bmatrix} \text{ row reduce } \rightarrow \begin{bmatrix} \mathbf{1} & -3 & 3 \\ 0 & \mathbf{1} & 2 \\ 0 & 0 & -35 \end{bmatrix} \tag{2}$$

This system is inconsistent because the last row has no pivot, so there is no solution, and there is no way to get the last column as a combination of the first two. So it must not be in the column space.

$$\begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} \notin \text{ Col A} \tag{3}$$

which also implies that $\begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$ is not in the range of T.

Example problem

$$\text{If} \quad A = \begin{bmatrix} 1 & 1 & 2 & -1 \\ 1 & 1 & 0 & 1 \end{bmatrix} \tag{4}$$

Let T be the linear transformation defined by A. Then, $T\colon \mathbb{R}^4 \rightarrow \mathbb{R}^2$, because a vector multiplied by A would need four entries, and the result would have two. The problems are,

1. Find vectors that span the kernel of T.
2. Describe the range of T.

The answers to each are

1. Since the kernel is the set of $\vec{x}$ such that $A\vec{x} = \vec{0}$, it is just the null space. We write the kernel of T as Ker T, so

$$\text{Ker T} \;=\; \text{Nul A} \tag{5}$$

To find Nul A, we use the augmented matrix of $A\vec{x} = \vec{0}$.

$$\begin{bmatrix} 1 & 1 & 2 & -1 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$
$$\text{row reduce } \rightarrow \begin{bmatrix} \mathbf{1} & 1 & 0 & 1 & 0 \\ 0 & 0 & \mathbf{1} & -1 & 0 \end{bmatrix}$$

There are pivots in columns 1 and 3, so x_1 and x_3 are basic variables. x_2 and x_4 are free variables. The matrix corresponds to the system

$$\begin{cases} x_1 + x_2 + x_4 = 0 \\ x_3 - x_4 = 0 \end{cases}$$

$$\implies \begin{cases} x_1 = -x_2 - x_4 \\ x_2 \text{ is free} \\ x_3 = x_4 \\ x_4 \text{ is free} \end{cases}$$

So the solution is $\begin{bmatrix} -x_2 - x_4 \\ x_2 \\ x_4 \\ x_4 \end{bmatrix}$, and we want to express this as a linear combination of some vectors, so we write it as a sum. (Look at the coefficients on the variables).

$$\begin{bmatrix} -x_2 - x_4 \\ x_2 \\ x_4 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

x_2 and x_4 are the weights, and they can change freely. The span of these vectors gives the null space, so

$$\text{Ker T} = \text{Span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\} \tag{6}$$

This is how we find the kernel of a T that is defined by a matrix.

2. For the next question, since the range of T and the column space of A are the same,

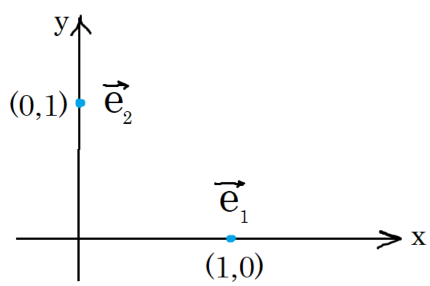
$$\text{Range of T} = \text{Col A} = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\} \tag{7}$$

because the column space is the span of the individual columns. Recall the theorem which states that Col A $= \mathbb{R}^m$ if A has a pivot in every row. In our case, there is a pivot in every row, so Col A $= \mathbb{R}^2$, which implies that

$$\text{Range of T } = \mathbb{R}^2 \tag{8}$$

It was already given that the codomain is $\mathbb{R}^2$ at the beginning of this problem, because of the statement $T\colon \mathbb{R}^4 \to \mathbb{R}^2$, so the codomain and range happen to be the same in this example.

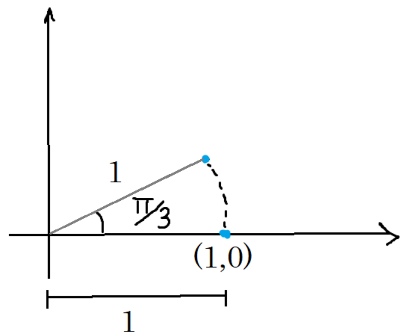
Finding the matrix associated with a linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$
We want to express a given T as a matrix such that $T(\vec{x}) = A\vec{x}$. For example, suppose T is a rotation of $\frac{\pi}{3}$ radians in the counterclockwise direction about the origin in $\mathbb{R}^2$. Find A so that $T(\vec{x}) = A\vec{x}$. We start by defining two basic vectors.
Let $\vec{e_1} = \begin{bmatrix}1\\0\end{bmatrix}$ and $\vec{e_2} = \begin{bmatrix}0\\1\end{bmatrix}$.



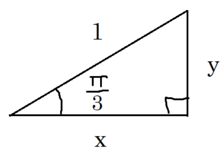
What are $T(\vec{e_1})$ and $T(\vec{e_2})$? Later, we will find that they're useful in the matrix

$$A = \begin{bmatrix} T(\vec{e_1}) & T(\vec{e_2}) \end{bmatrix} \tag{9}$$

such that $T(\vec{e_1})$ and $T(\vec{e_2})$ are the columns. To find these, we're really asking where is $\vec{e_1} = \begin{bmatrix}1\\0\end{bmatrix}$ is sent?



The new coordinates are found by



So $x = \cos(\frac{\pi}{3}) = \frac{1}{2}$ and $y = \sin(\frac{\pi}{3}) = \frac{\sqrt{3}}{2}$. These are the new coordinates of $\vec{e_1}$, so

$$T(\vec{e_1}) = \begin{bmatrix} 1/2 \\ \frac{\sqrt{3}}{2} \end{bmatrix}$$

Similar reasoning can show that

$$T(\vec{e_2}) = \begin{bmatrix} -\frac{\sqrt{3}}{2} \\ 1/2 \end{bmatrix}$$

The **claim** is that $T(\vec{x}) = A\vec{x}$ for every $\vec{x}$, where

$$A = \begin{bmatrix} T(\vec{e_1}) & T(\vec{e_2}) \end{bmatrix} \tag{10}$$

$$= \begin{bmatrix} 1/2 & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 1/2 \end{bmatrix} \tag{11}$$

Why does this work? Any vector $\vec{u} = \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$ can be written as

$$\vec{u} = x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{12}$$

which is equal to $\vec{u} = x \cdot \vec{e_1} + y \cdot \vec{e_2}$. So $\vec{u}$ is a linear combination of $\vec{e_1}$ and $\vec{e_2}$ with weights x and y. To see where all the points get mapped to, we want to find $T(\vec{u})$.

$$\begin{aligned} T(\vec{u}) &= T(x\vec{e_1} + y\vec{e_2}) \\ &= x \cdot T(\vec{e_1}) + y \cdot T(\vec{e_1}) \end{aligned}$$

by the properties of linear transformations. This is a linear combination of vectors, so we can form a matrix whose columns are $T(\vec{e_1})$ and $T(\vec{e_2})$. The equation becomes

$$T(\vec{u}) = \begin{bmatrix} T(\vec{e_1}) & T(\vec{e_2}) \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{13}$$

$$= A \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{14}$$

And that's why it works. Also, recall that $\vec{u} = \begin{bmatrix} x \\ y \end{bmatrix}$, then

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = A \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{15}$$

To find the linear transformation of any vector $\vec{x}$ in $\mathbb{R}^2$, we can use this equation.

Summary: We took the vectors $\vec{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and found the transformations on those, then since any vector $\vec{x}$ can be written as a linear combination of $\vec{e_1}$ and $\vec{e_2}$, the transformation of $\vec{x}$ can be written as a linear combination of $T(\vec{e_1})$ and $T(\vec{e_2})$. Then, by the definition of a matrix product,

$$x \cdot T(\vec{e_1}) + y \cdot T(\vec{e_1}) = \begin{bmatrix} T(\vec{e_1}) & T(\vec{e_2}) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{16}$$

This matrix is our A!

Theorem:
If $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, and $\vec{e_1}, \vec{e_2}, \dots, \vec{e_n}$ are defined as

$$\vec{e_1} = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \vec{e_2} = \begin{bmatrix} 0 \\ 1 \\ \dots \\ 0 \end{bmatrix}, \cdots, \vec{e_n} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 1 \end{bmatrix} \tag{17}$$

then the matrix associated with T is

$$A = \begin{bmatrix} T(\vec{e_1}) & T(\vec{e_2}) & \cdots & T(\vec{e_n}) \end{bmatrix} \tag{18}$$

where $T(\vec{e_1})$ is the first column, $T(\vec{e_2})$ is the second column, and so on.
This matrix is called the **Standard matrix of T**.

Definitions:

1. The image of a vector is the vector that it gets mapped to.
2. A linear transformation from V to W, $T\colon V \to W$ is **one-to-one** (or injective) if each $\vec{b} \in W$ is the image of at most one $\vec{x} \in V$. (i.e. $T(\vec{x}) = T(\vec{y})$ implies $\vec{x} = \vec{y}$).
3. T is **onto** (or surjective) if each $\vec{b}$ is the image of at least one $\vec{x} \in V$. That is,

$$\text{Range of T } = W$$

4. If T is both injective and surjective, then T is an **isomorphism**, and V and W are **isomorphic vector spaces**.

Lecture 10

February 28, 2022

Composition

We can compose two linear transformations just as we can compose two functions, where the codomain of the first matches the domain of the second, (so that any result from the first transformation is in the domain of the second transformation). For example, if $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ and $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ and we have the standard matrices

$$A = \begin{bmatrix} 2 & 0 & 4 \\ 1 & 2 & 7 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 0 \\ 4 & -5 \\ 2 & 0 \end{bmatrix} \tag{1}$$

associated with T and S respectively, find $S(T(\vec{u})) = (S \circ T)(\vec{u})$ where $\vec{u} = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix}$.

First, we apply T to $\vec{u}$, and then we apply S to that.

$$T(\vec{u}) = A\vec{u} = \begin{bmatrix} 2 & 0 & 4 \\ 1 & 2 & 7 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$$

Then since B is the standard matrix of S,

$$\begin{aligned} S(T(\vec{u})) = S\left(\begin{bmatrix} 8 \\ 6 \end{bmatrix}\right) &= B \cdot \begin{bmatrix} 8 \\ 6 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 4 & -5 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 8 \\ 6 \end{bmatrix} \\ &= \begin{bmatrix} 8 \\ 2 \\ 16 \end{bmatrix} \end{aligned}$$

Summary: we started with a vector in three coordinates, then went to two, and back to three. So we went through

$$\mathbb{R}^3 \to \mathbb{R}^2 \to \mathbb{R}^3 \tag{2}$$

The first arrow represents T, and the second arrow represents S. The end result is

$$\begin{aligned}(S \circ T)(\vec{u}) &= S(T(\vec{u})) \\ &= S(A\vec{u}) \\ &= B(A\vec{u}) \\ &= (BA)\vec{u}\end{aligned}$$

This means we can express $S \circ T$ as $(BA)\vec{u}$. So we can think of BA as the standard matrix of $S \circ T$.

Matrix definitons:
Let A be an $m \times n$ matrix. We can write it as

$$A = (a_{ij}) \tag{3}$$

where a_{ij} is the entry in row i and column j. This notation for a matrix is useful. The **diagonal entries** of A are a_{ii}, (that is a_{11}, a_{22}, a_{33}, and so on). These entries form the **main diagonal** of A. A **diagonal matrix** is a square matrix (i.e. m=n) whose non-diagonal entries are all zero. The following examples are given.

$$\text{For } \begin{bmatrix} \mathbf{2} & 4 & 7 \\ 9 & \mathbf{3} & 1 \end{bmatrix} \text{ the main diagonal is 2 3.}$$

$$\text{Diagonal matrix: } \begin{bmatrix} \mathbf{2} & 0 & 0 \\ 0 & \mathbf{7} & 0 \\ 0 & 0 & \mathbf{100} \end{bmatrix} \text{ this is also a square matrix}$$

A **zero matrix** is a matrix with all entries zero. For example, $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. We can write the zero matrix as any of the following

$$0, (0), [0], 0_{m\times n} \tag{4}$$

The **identity matrix** is a square diagonal matrix with $a_{ii} = 1$ for all i. For example,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

We write an $n \times n$ identity matrix as I or I_n. Two matrices are **equal** if they have the same size and the same entries.

Operations on matrices:

1. **Sum:** If A,B are $m \times n$ matrices, then
$$A + B = (a_{ij} + b_{ij}) \tag{6}$$
We just add the corresponding entries. For example, if
$$A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} \text{ then } A + B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$$
If $C = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$, then A+B and B+C are undefined because C is not the same size as either A or B.

2. **Scalar multiplication:**
If A is an $m \times n$ matrix, and $r \in \mathbb{R}$, then
$$rA = (ra_{ij}) \tag{7}$$
we multiply each entry by the real number r. For example, if $A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}$ then $2A = \begin{bmatrix} 8 & 0 & 10 \\ -2 & 6 & 4 \end{bmatrix}$.

Theorem: Let A,B,C be matrices of the same size and $r, s \in \mathbb{R}$. Then

(a) $A + B = B + A$

(b) $(A + B) + C = A + (B + C)$

(c) $A + 0_{m \times n} = A$

(d) $r(A + B) = rA + rB$

(e) $(r + s)A = rA + sA$

(f) $r(sA) = (rs)A$

3. **Matrix multiplication:** If A is $m \times n$ and B is $n \times p$, (the number of columns in the first matrix should match the number of rows in the second), such that $B = \begin{bmatrix} \vec{b_1} & \vec{b_2} & \cdots & \vec{b_p} \end{bmatrix}$ then

$$AB = A \begin{bmatrix} \vec{b_1} & \vec{b_2} & \cdots & \vec{b_p} \end{bmatrix} \tag{8}$$

$$= \begin{bmatrix} A\vec{b_1} & A\vec{b_2} & \cdots & A\vec{b_p} \end{bmatrix} \tag{9}$$

We multiply the matrix A by each column vector in B and put it in a new matrix. Remarks:

   (a) AB is $m \times p$ because each vector has m rows, and there are p columns in the final matrix.

   (b) The number of columns in A must match the number of rows in B for AB to be defined.

$$(m \times \mathbf{n})(\mathbf{n} \times p) = m \times p \tag{10}$$

   (c) Each column of AB is a linear combination of columns of A with weights given by the appropriate column of B. We see this as follows

$$A\vec{b_1} = \begin{bmatrix} \vec{a_1} & \vec{a_2} & \cdots & \vec{a_n} \end{bmatrix} \vec{b_1} = \begin{bmatrix} \vec{a_1} & \vec{a_2} & \cdots & \vec{a_n} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{12} \\ \cdots \\ b_{1n} \end{bmatrix}$$

$$= b_{11}\vec{a_1} + b_{12}\vec{a_2} + \cdots + b_{1n}\vec{a_n}$$

   So $A\vec{b_1}$ is a linear combination of $\vec{a_1}, \cdots, \vec{a_n}$ with weights given by $b_{11}, b_{12}, \cdots, b_{1n}$.

The following example of matrix multiplication is given. Let

$$A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \text{ and } B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$$

Note that AB is defined, but BA is undefined because the columns and rows wouldn't match. The product is

$$AB = \begin{bmatrix} A\vec{b_1} \cdots A\vec{b_p} \end{bmatrix} = \begin{bmatrix} A \begin{bmatrix} 4 \\ 1 \end{bmatrix} & A \begin{bmatrix} 3 \\ -2 \end{bmatrix} & A \begin{bmatrix} 6 \\ 3 \end{bmatrix} \end{bmatrix}$$

And we evaluate the column vectors

$$A\begin{bmatrix}4\\1\end{bmatrix}=\begin{bmatrix}2&3\\1&-5\end{bmatrix}\begin{bmatrix}4\\1\end{bmatrix}=\begin{bmatrix}2\cdot 4+3\cdot 1\\1\cdot 4-5\cdot 1\end{bmatrix}=\begin{bmatrix}11\\-1\end{bmatrix}$$
$$A\begin{bmatrix}3\\-2\end{bmatrix}=\cdots=\begin{bmatrix}0\\13\end{bmatrix}$$
$$A\begin{bmatrix}6\\3\end{bmatrix}=\cdots=\begin{bmatrix}21\\-9\end{bmatrix}$$

So $AB=\begin{bmatrix}11&0&21\\-1&13&-9\end{bmatrix}$.

Row-column rule for matrix multiplication
If $(AB)_{ij}$ is the entry of AB in the i^{th} row and the j^{th} column, then

$$(AB)_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj} \tag{11}$$

(when A is $m\times n$). This formula is another way to find the product of two matrices. Using the earlier example,

$$AB=\begin{bmatrix}2&3\\1&-5\end{bmatrix}\begin{bmatrix}4&3&6\\1&2&3\end{bmatrix}$$

The first entry is $(AB)_{11}=a_{11}b_{11}+a_{12}b_{21}$, the second entry is $(AB)_{21}=a_{21}b_{11}+a_{22}b_{21}$, and so on. So

$$\begin{bmatrix}2&3\\1&-5\end{bmatrix}\begin{bmatrix}4&3&6\\1&2&3\end{bmatrix}=\begin{bmatrix}2\cdot 4+3\cdot 1&\cdots\\1\cdot 4-5\cdot 1&\cdots\end{bmatrix}=\begin{bmatrix}11&0&21\\-1&13&-9\end{bmatrix}$$

The first entry is the result of multiplying the first row of A with the first column of B.

$$\begin{bmatrix}2&3\end{bmatrix}\begin{bmatrix}4\\1\end{bmatrix}=2\cdot 4+3\cdot 1$$

Then the entry $(AB)_{21}$ is the product of the second row of A with the first column of B, and so on. To memorize this, think: the row that the entry is in tells you which row from A to use, and the column that the entry is in tells you which column from B to use, then to calculate the entry, multiply the row by column. Notice the size of the matrices:

$$(2\times 2)(2\times 3)=(2\times 3)$$

Theorem: If A is $m \times n$ and B,C are matrices of the appropriate sizes so that the operations are defined, then

(a) $A(BC) = (AB)C$

(b) $A(B + C) = AB + AC$

(c) $(B + C)A = BA + CA$

(d) $I_m A = A = AI_n$ where I is the identity matrix, and the number of columns on the first matrix matches the number of rows on the second.

Lecture 11

March 2, 2022

Remarks on matrix multiplication

1. $AB \neq BA$
2. Cancellation law does not hold. So $AB = CB$ does not imply $A = C$.
3. If $AB = 0$, it is not true that $A = 0$ or $B = 0$.

Powers of a matrix

If A is $n \times n$, (it is a square matrix), then

$$A^k = AA \dots A \tag{1}$$

where A is multiplied by itself k times, and

$$A^0 = I_n \tag{2}$$

where I_n is the $n \times n$ identity matrix.

Transpose

A is $m \times n$. The **transpose** of A, written as A^T, is an $n \times m$ matrix such that the i^{th} row of A^T is the i^{th} column of A, and the j^{th} column of A^T is the j^{th} row of A. (Basically, we just swap the rows and columns). The following example is given,

$$A = \begin{bmatrix} 2 & 3 & 7 \\ 0 & -1 & 1 \end{bmatrix}$$
$$A^T = \begin{bmatrix} 2 & 0 \\ 3 & -1 \\ 7 & 1 \end{bmatrix}$$

Theorem: If A,B are matrices of the appropriate sizes, then:

1. $(A^T)^T = A$

2. $(A+B)^T = A^T + B^T$

3. For all $r \in \mathbb{R}$, $(rA)^T = rA^T$

4. $(AB)^T = B^T A^T$, similarly, $(ABC)^T = C^T B^T A^T$

Inverse of a matrix

An $n \times n$ matrix is **invertible** if there exists an $n \times n$ matrix C such that $AC = I_n$ and $CA = I_n$. Here, C is called the **inverse** of A, and we write it $C = A^{-1}$.

The inverse matrix C is uniquely determined by A. *Proof:*

If $BA = I_n$ and $AB = I_n$, then

$$B = BI = B(AC) = (BA)C = IC = C$$

So $B = C$, and the inverse is unique.

Definitions:

A matrix that is not invertible is called **singular**, and an invertible matrix is called **non-singular**.

Example of an inverse matrix:

$$\text{If } A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} \text{ and } C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} \text{ show that } C = A^{-1}.$$

Check that $AC = I$ and $CA = I$:

$$AC = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 2 \cdot (-7) + 5 \cdot 3 & 2 \cdot (-5) + 5 \cdot 2 \\ (-3) \cdot (-7) + 3 \cdot (-7) & (-3)(-5) + (-7)(2) \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

This is the identity matrix.

$$CA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} (-7) \cdot 2 + (-5)(-3) & (-7) \cdot 5 + (-5)(-7) \\ 3 \cdot 2 + 2 \cdot (-3) & 3 \cdot (5) + 2 \cdot (-7) \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Therefore $C = A^{-1}$.
One simple example of an inverse is the identity matrix:

If $I = I_n$ then $I \cdot I = I$ implies $I = I^{-1}$. So I is the inverse of itself.

Theorem: If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $ad - bc \neq 0$ then A is invertible and

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{3}$$

If $ad - bc$ is zero, then A is singular (not invertible). The quantity $ad - bc$ is called the **determinant** of A, written det A. So a 2×2 matrix A is invertible if and only if

$$\text{det A } \neq 0 \tag{4}$$

The following example is given.
If $A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$ such that $a = 3, b = 4, c = 5, d = 6$, then

$$\begin{aligned} \text{det A } &= 3 \cdot 6 - 4 \cdot 5 \\ &= -2 \neq 0 \end{aligned}$$

So A is invertible, and

$$A^{-1} = \frac{1}{-2} \cdot \begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}$$

We can check that this works by the definition of an inverse matrix:

$$\begin{aligned} \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix} &= \begin{bmatrix} 3 \cdot (-3) + 4 \cdot (5/2) & 3 \cdot 2 + 4 \cdot (-3/2) \\ 5 \cdot (-3) + 6 \cdot (5/2) & 5 \cdot 2 + 6 \cdot (-3/2) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned}$$

So it is the inverse.

Theorem:
If A is an invertible matrix, then for each vector $\vec{b} \in \mathbb{R}^m$, the equation $A\vec{x} = \vec{b}$ has a unique solution that is

$$\vec{x} = A^{-1}\vec{b} \tag{5}$$

Proof: If $A\vec{x} = \vec{b}$ then

$$A^{-1}\left(A\vec{x} = \vec{b}\right)$$
$$A^{-1}A\vec{x} = A^{-1}\vec{b}$$

$A^{-1}A = I$ and $I\vec{x} = \vec{x}$ so the equation above becomes

$$\vec{x} = A^{-1}\vec{b} \tag{6}$$

so given that A is invertible, this is the solution.

We can use this theorem to solve the equation $A\vec{x} = \vec{b}$ using the inverse of A. For example, given the linear system

$$\begin{cases} 3x_1 + 4x_2 = 3 \\ 5x_1 + 6x_2 = 7 \end{cases}$$

This is equivalent to the matrix equation

$$\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$
$$\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix} \vec{x} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

which represents our $A\vec{x} = \vec{b}$. The matrix $\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$ is invertible by the previous example, so

$$\begin{aligned} \vec{x} = A^{-1}\vec{b} &= \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 7 \end{bmatrix} \\ &= \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix} \begin{bmatrix} 3 \\ 7 \end{bmatrix} \\ &= \begin{bmatrix} -3 \cdot 3 + 2 \cdot 7 \\ \frac{5}{2} \cdot 3 + (\frac{-3}{2}) \cdot 7 \end{bmatrix} \\ &= \begin{bmatrix} 5 \\ -3 \end{bmatrix} \end{aligned}$$

This is our solution.

The theorem also implies that if A is invertible, then Col A= $\mathbb{R}^m$ and Nul A= $\left\{\vec{0}\right\}$, because it says there exists a unique solution $\vec{x}$ for any given $\vec{b}$, so the column space spans $\mathbb{R}^m$. Also, the null space is the set of $\vec{x}$ such that $A\vec{x} = \vec{0}$, so by the theorem

$$\begin{aligned}\vec{x} &= A^{-1}\vec{b} \\ &= A^{-1}\vec{0} \\ &= \vec{0}\end{aligned}$$

so the null space is $\left\{\vec{0}\right\}$.

Theorem:

1. If A is invertible, then A^{-1} is also invertible, and $(A^{-1})^{-1} = A$.
2. For invertible $n \times n$ matrices A,B (of the same size),

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \\ &= AIA^{-1} \\ &= AA^{-1} = I\end{aligned}$$

   So

$$(AB)^{-1} = B^{-1}A^{-1} \tag{7}$$

   And for three matrices, $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.

3. If A is invertible, then A^T is also invertible, and

$$(A^T)^{-1} = (A^{-1})^T \tag{8}$$

   Proof:

$$\begin{aligned}A^{-1}A &= I \\ (A^{-1}A)^T &= I^T \\ A^T(A^{-1})^T &= I \\ (A^{-1})^T &= (A^T)^{-1}\end{aligned}$$

How do we find A^{-1} if A is not 2×2?

An **elementary** matrix is a matrix obtained by a single row operation on the identity matrix. For example,

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \quad E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

The row operations done on the identity matrix are:

$$\begin{aligned} E_1 &: r_3 \to r_3 - 4r_1 \\ E_2 &: r_1 \longleftrightarrow r_2 \\ E_3 &: r_3 \to 5r_3 \end{aligned}$$

When we multiply one of these elementary matrices by a given A, the result is that the same row operation is applied to A. Basically, the row operation itself is expressed by these matrices.

Theorem: An $m \times n$ matrix A is invertible if and only if A is row equivalent to the identity matrix I_n, (the reduced echelon form of A is I_n). In this case, any sequence of row operations that produces I_n from A will produce A^{-1} from I_n. (If we apply the same steps that we did to row reduce from A to I_n, but starting with I_n then we get A^{-1}). **Procedure:**

1. Augment A with I_n:
$$\begin{bmatrix} A & | & I_n \end{bmatrix}$$

2. Row reduce $\begin{bmatrix} A & | & I_n \end{bmatrix}$ to get $\begin{bmatrix} I_n & | & A^{-1} \end{bmatrix}$ if possible. If not possible, A is singular.

For example, find A^{-1} if $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$

$$\text{First step:} \quad \begin{bmatrix} A & | & I_3 \end{bmatrix} = \left[\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right]$$

$$\text{Row reduce:} \quad r_1 \longleftrightarrow r_2 \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right]$$

$$r_3 \to r_3 - 4r_1 \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{array}\right]$$

$$r_3 \to r_3 + 3r_2 \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{array}\right]$$

$$r_3 \to \frac{1}{2}r_3 \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{array}\right]$$

$$r_2 \to r_2 - 2r_3 \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{array}\right]$$

$$r_1 \to r_1 - 3r_3 \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{array}\right]$$

The reduced echelon form on the left is I_3, so the matrix on the right is A^{-1}.

Lecture 12

March 7, 2022

Elementary matrices

Examples of elementary matrices include

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix} \tag{1}$$

We get these by applying one row operation to the identity matrix.

Recall the theorem which states that A is invertible if and only if A is row equivalent to the identity matrix. Any sequence of row operations that produce I_n from A will produce A^{-1} from I_n. Our goal is to use the elementary matrices to do these row operations. For example, if

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \tag{2}$$

Then look at the following product

$$E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g-4a & h-4b & i-4c \end{bmatrix} \tag{3}$$

Notice that, in effect, multiplying E_1 by A is equivalent to applying the row operation $r_3 \to r_3 - 4r_1$ to A. Now look at the second elementary matrix E_2:

$$E_2 A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix} \tag{4}$$

This represents the row operation $r_2 \longleftrightarrow r_1$. Consider E_3A

$$E_3A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{bmatrix} \tag{5}$$

This is the row operation $r_3 \to 5r_3$.

So we can use elementary matrices to perform the row operations. (Basically, the matrix is a way of encoding the row operation).

Any possible sequence of row operations corresponds to a sequence of elementary matrices we can multiply by A. To row reduce,

$$A \to I_n$$

we apply one row operation at a time, starting with E_1 and going up to however many there are:

$$(E_k \cdots (E_2(E_1A)) \cdots) = I_n$$

which is equivalent to

$$(E_kE_{k-1} \cdots E_2E_1)\, A = I_n \tag{6}$$

where each E is an elementary matrix. Recall the algorithm

$$\begin{bmatrix} A & | & I_n \end{bmatrix} \to \text{row reduce} \to \begin{bmatrix} I_n & | & A^{-1} \end{bmatrix} \tag{7}$$

In this process of row reduction, what we're really doing is

$$E_kE_{k-1} \cdots E_2E_1 \begin{bmatrix} A & | & I_n \end{bmatrix} = \begin{bmatrix} (E_kE_{k-1} \cdots E_2E_1)A & | & (E_kE_{k-1} \cdots E_2E_1)I_n \end{bmatrix}$$

Then since $(E_kE_{k-1} \cdots E_2E_1)A = I_n$ and the inverse of matrix is unique, $E_kE_{k-1} \cdots E_2E_1 = A^{-1}$ because it undoes A. This means the augmented matrix becomes

$$\begin{bmatrix} I_n & | & A^{-1} \end{bmatrix} \tag{8}$$

This is why the algorithm works.

Invertible Matrix Theorem:
Let A be a square $n \times n$ matrix. The following are equivalent:

1. A is invertible.

2. A is row equivalent to the identity matrix I_n.

3. A has n pivot positions.

4. The linear transformation $\vec{x} \to A\vec{x}$ is injective (one-to-one).

5. The linear transformation $\vec{x} \to A\vec{x}$ is surjective (or onto, which means the column space fills the whole range).

6. A^T is invertible.

(Also, statements 4 and 5 are enough to show that the linear transformation $T\colon \vec{x} \to A\vec{x}$ is an isomorphism). An example of the use of this theorem is given:

$$\text{Is } A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix} \quad \text{invertible?}$$

$$\text{row reduce } A \to \begin{bmatrix} \mathbf{1} & 0 & -2 \\ 0 & \mathbf{1} & 4 \\ 0 & 0 & \mathbf{3} \end{bmatrix}$$

There are three pivots, so by part 3 of the theorem, A is invertible.

Definition: A linear transformation $T\colon V \to W$ is **invertible** if there exists a linear transformation $S\colon W \to V$ such that

$$S(T(\vec{x})) = \vec{x} \text{ for all } \vec{x} \in V$$
$$\text{and} \quad T(S(\vec{y})) = \vec{y} \text{ for all } \vec{y} \in W.$$

Also, if the transformation is done using a matrix, and $S \circ T$ means the composition of S and T, then this definition is saying

$$\begin{cases} S \circ T = I \\ T \circ S = I \end{cases} \tag{9}$$

$$
\begin{tikzcd}
V \arrow[r, bend left, "T"] & W \arrow[l, bend left, "S"]
\end{tikzcd}
$$

Theorem:
If $V = W = \mathbb{R}^n$ and A is the standard matrix of T, (so that $T(\vec{x}) = A\vec{x}$), then T is invertible if and only if A is invertible. In this case, the inverse of T is the linear transformation

$$\vec{x} \to A^{-1}\vec{x} \tag{10}$$

Think of it like this:

$$
\begin{aligned}
T &: \vec{x} \to A\vec{x} \\
T^{-1} &: A\vec{x} \to A^{-1}(Ax) = \vec{x}
\end{aligned}
$$

T takes $\vec{x}$ to $A\vec{x}$ and then T^{-1} takes $A\vec{x}$ back to $\vec{x}$. We can also say that

$$T^{-1}(T(\vec{x})) = A^{-1}(A\vec{x}) = \vec{x} \tag{11}$$

Example problem
What can you say about an injective linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^n$? First, we can rewrite the transformation as $T\colon \vec{x} \to A\vec{x}$. Then, by the invertible matrix theorem,

1. T is surjective.
2. A is $n \times n$.
3. A is invertible, (where A is the standard matrix of T), and T is invertible.
4. A has n pivots, (pivot in every row and every column).

Lecture 13

March 14, 2022

Determinants

Let $A = (a_{ij})$ be an $n \times n$ matrix. Then A_{ij} is the matrix formed from A by deleting the i^{th} row and j^{th} column from A. (It will be an $(n-1) \times (n-1)$ matrix). For example, if

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

Then

$$A_{11} = \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix}$$

where we deleted the first row and the first column from A: $\begin{bmatrix} \boldsymbol{1} & \boldsymbol{-2} & \boldsymbol{5} \\ \boldsymbol{2} & 0 & 4 \\ \boldsymbol{0} & 1 & 0 \end{bmatrix}$. Also,

$$A_{23} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

Here we deleted the second row and the third column from A: $\begin{bmatrix} 1 & -2 & \boldsymbol{5} \\ \boldsymbol{2} & \boldsymbol{0} & \boldsymbol{4} \\ 0 & 1 & \boldsymbol{0} \end{bmatrix}$.

Cofactors: The value

$$(-1)^{i+j} \ \det A_{ij} \tag{1}$$

is denoted C_{ij}, and is called the (i, j)-cofactor of A. Note that we're using A_{ij}, the matrix A with row i and column j deleted. Also note that $(-1)^{i+j}$ is only either -1 or $+1$, (because if the power is even then the result is 1, and if the power is odd then the result is -1).

Definition: For $n \geq 2$, the **determinant** of an $n \times n$ matrix $A = (a_{ij})$ is

$$\det\, A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n} \tag{2}$$

This is the cofactor expansion along the first row, but it may be done along any row. The following example is given,

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

The signs of cofactors will alternate along the entries, because $(-1)^{i+j}$ is -1 or $+1$. Do a cofactor expansion along the first row:

$$\det\, A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

$$= (1)(-1)^{1+1} \det\, A_{11} + (-2)(-1)^{1+2} \det\, A_{12} + (5)(-1)^{1+3} \det\, A_{13}$$

$$= 1 \cdot \det \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix} + 2 \cdot \det \begin{bmatrix} 2 & 4 \\ 0 & 0 \end{bmatrix} + 5 \cdot \det \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= (0-4) + 2 \cdot (0-0) + 5 \cdot (2-0)$$
$$= 6$$

where we used the fact that the determinant of a 2×2 matrix $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\det\, B = ad - bc$. Remarks:

1. We can also write
$$\begin{vmatrix} 1 & -2 & 5 \\ 2 & 0 & 4 \\ 0 & 1 & 0 \end{vmatrix} \text{ for det } A$$

2. We can expand along any row and any column. To choose which row, do the one with the most zeros, because then more terms will just be zero and our work will be easier.

Example

Find det A by expanding along the third row.

$$\begin{vmatrix} 1 & -2 & 5 \\ 2 & 0 & 4 \\ 0 & 1 & 0 \end{vmatrix} = a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33}$$

$$= 0 \cdot C_{31} + a_{32} \cdot C_{32} + 0 \cdot C_{33}$$

$$= a_{32} \cdot C_{32}$$

$$= (1)(-1)^{3+2} \cdot \det \begin{bmatrix} 1 & 5 \\ 2 & 4 \end{bmatrix}$$

$$= -1 \cdot (4 - 10)$$

$$= 6$$

Definition

An $m \times n$ matrix $A = (a_{ij})$ is **triangular** if $a_{ij} = 0$ for all $j > i$ or for all $j < i$, (either above or below the diagonal). If there are all zeros below the diagonal, it is called upper triangular. If all the zeros are above the diagonal, then it is lower triangular. Examples:

$$\begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \end{bmatrix} \text{ is upper triangular.}$$

$$\begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix} \text{ is lower triangular.}$$

$$\begin{bmatrix} \mathbf{1} & 3 & 5 & 7 \\ 0 & \mathbf{2} & 4 & 6 \\ 0 & 0 & \mathbf{3} & 1 \end{bmatrix} \text{ is upper triangular, and the diagonal is 1 2 3.}$$

Theorem:

If A is a square triangular matrix, then det A is the product of the entries along the diagonal. For example, find det A if

$$A = \begin{bmatrix} \mathbf{3} & -7 & 8 & 9 & -6 \\ 0 & \mathbf{2} & -5 & 7 & 3 \\ 0 & 0 & \mathbf{1} & 5 & 0 \\ 0 & 0 & 0 & \mathbf{4} & -1 \\ 0 & 0 & 0 & 0 & \mathbf{1} \end{bmatrix}$$

This is upper triangular, so by the theorem,

$$\det A = 3 \cdot 2 \cdot 1 \cdot 4 \cdot 1 = 24$$

Why does the theorem work?

$$\det A = \begin{vmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & 7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 4 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

Take a cofactor expansion along the first column:

$$\begin{aligned} \det A = 3 \cdot &\begin{vmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 4 & -1 \\ 0 & 0 & 0 & 1 \end{vmatrix} \\ = 3 \cdot 2 \cdot &\begin{vmatrix} 1 & 5 & 0 \\ 0 & 4 & -1 \\ 0 & 0 & 1 \end{vmatrix} \\ = 3 \cdot 2 \cdot 1 \cdot &\begin{vmatrix} 4 & -1 \\ 0 & 1 \end{vmatrix} \\ = 3 \cdot 2 \cdot 1 \cdot 4 \cdot 1 \end{aligned}$$

where at each step we took a cofactor expansion along the first column, and since there are all zeros below the diagonal, all terms in the cofactor expansion will be zero except the first one, so we only use the first entry times the determinant of the matrix with the first row and first column deleted. We got the same result, and that is why the end result is just the product of all the diagonal entries.

Row operations:

1. If a multiple of one row of A is added to another row of A to form B, then

$$\det A = \det B \tag{3}$$

2. If two rows of A are interchanged to form B, then

$$\det A = -\det B \tag{4}$$

3. If a row of A is multiplied by a constant k to form B, then

$$\det B = k \cdot \det A \tag{5}$$

The goal is to reduce a complicated matrix to a simpler matrix (ideally a triangular matrix), so that we can just find the determinant of that simpler matrix. For example, find det A if

$$A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix}$$

Perform the row operations $r_2 \to r_2 + 2r_1$, $r_3 \to r_3 + r_1$.

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

This has the same determinant because of Part 1 of the theorem. Then perform the row operation $r_2 \longleftrightarrow r_3$.

$$\begin{bmatrix} \mathbf{1} & -4 & 2 \\ 0 & \mathbf{3} & 2 \\ 0 & 0 & \mathbf{-5} \end{bmatrix}$$

Part 2 of the theorem implies this has the opposite sign of det A.

$$\det \begin{bmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{bmatrix} = (1)(3)(-5) = -15$$

since it is upper triangular. Now we use the theorem to get det A.

$$\det A = \det \begin{bmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{bmatrix} = -\det \begin{bmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{bmatrix} = -(-15)$$

So

$$\det A = 15$$

Another example

Find det A if

$$A = \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}$$

Divide the first row by 2.

$$\det\ A = 2 \cdot \begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix}$$

Then do $r_2 \to r_2 - 3r_1$, $r_3 \to r_3 + 3r_1$, $r_4 \to r_4 - r_1$.

$$\det\ A = 2 \cdot \begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & -12 & 10 & 10 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

Then do $r_3 \to r_3 + 4r_2$.

$$\det\ A = 2 \cdot \begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

Then do $r_4 \to r_4 - \frac{1}{2}r_3$.

$$\det\ A = 2 \cdot \begin{vmatrix} \mathbf{1} & -4 & 3 & 4 \\ 0 & \mathbf{3} & -4 & -2 \\ 0 & 0 & -\mathbf{6} & 2 \\ 0 & 0 & 0 & \mathbf{1} \end{vmatrix}$$

So

$$\det\ A = 2 \cdot 1 \cdot 3 \cdot (-6) \cdot 1 = -36$$

We can use the row operations to simplify the process of finding the determinant.

Lecture 14

March 18, 2022

Suppose A is a square matrix and A has been reduced to echelon form U by row replacement and interchange only. (This is possible because it is just echelon form so we wouldn't need to multiply any row by a constant). If there are r interchanges in reducing A to U, then

$$\det A = (-1)^r \det U \tag{1}$$

This is useful because we can always make the matrix U upper-triangular since it is in echelon form, and then we can calculate its determinant as the product of its diagonal entries. Let $U = (u_{ij})$, where each u_{ij} is an entry, then

$$\det U = u_{11} \cdot u_{22} \cdot \ldots \cdot u_{nn} \tag{2}$$

So

$$\det A = (-1)^r u_{11} \cdot u_{22} \cdot \ldots \cdot u_{nn} \tag{3}$$

If any one of these is zero, then $\det A = 0$, and that corresponds to a diagonal entry that is zero. If this happens, it means one of the columns does not have a pivot, which implies A is not invertible. So if $\det A = 0$, then A is not invertible. We can see this more clearly by looking at the two possibilities:

1. If $\det A = 0$, then $u_{11}u_{22}\cdots u_{nn} = 0$, in which case $u_{ii} = 0$ for some i. So there is no pivot in the i^{th} column. So A is not invertible.

2. If $\det A \neq 0$, then $u_{11}u_{22}\cdots u_{nn} \neq 0$, in which case $u_{ii} \neq 0$ for all i. So there is a pivot in every column. So A is invertible.

This justifies the following theorem.

Theorem: A square matrix A is invertible if and only if $\det A \neq 0$.

For example,

$$A = \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & 8 & 0 & 9 \end{bmatrix}$$

$$r_3 \to r_3 + 2r_1 \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & 8 & 0 & 9 \end{bmatrix}$$

$$r_3 \to r_3 - r_2 \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -5 & 8 & 0 & 9 \end{bmatrix}$$

The row of zeros implies that one of the columns do not have a pivot, so A is not invertible, and det $A = 0$. (Another way to see that det $A = 0$ is to do a cofactor expansion along the 3^{rd} row, and then since it is all zeros, the determinant is also zero). Whenever you see two rows that are the same, you know that det $A = 0$.

Column operations

We can perform column operations just as we can row operations.

$$A = \begin{bmatrix} 3 & 1 & 4 \\ 2 & 7 & 1 \\ 0 & 0 & 3 \end{bmatrix} c_2 \to c_2 + c_1 \begin{bmatrix} 3 & 4 & 4 \\ 2 & 9 & 1 \\ 0 & 0 & 3 \end{bmatrix} = B$$

Given this operation, det A =det B. If you add a multiple of a column to another column, the determinant doesn't change. If you swap two columns, then you multiply the determinant by -1. If you multiply a column by a scalar k, then you multiply the determinant by k. (Note this could not be used to solve systems of linear equations, only determinants). This is true because we can take the transpose of A, and then the row operations correspond to the column operations.

Theorem:

If A is a square matrix, then

$$\det A = \det A^T \tag{4}$$

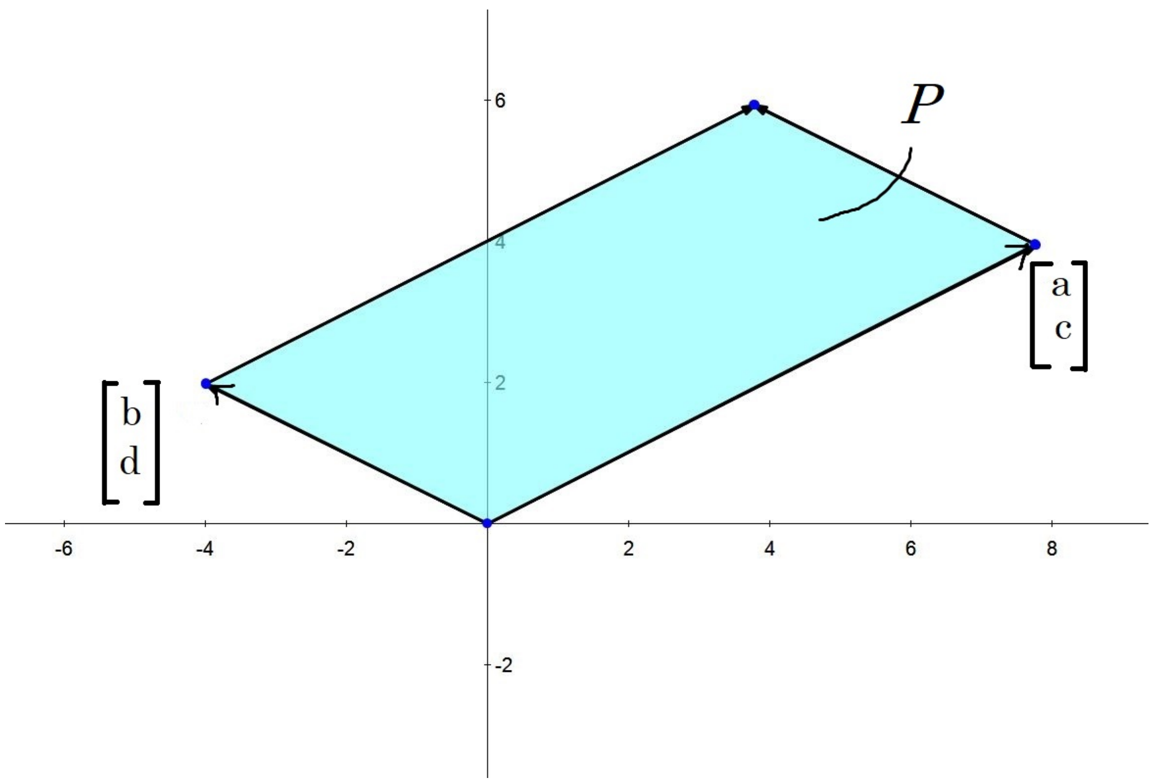
Because of this theorem, column operations have the same effect as row operations on the determinant. For example,

$$\begin{vmatrix} 3 & 3 & 1 \\ 4 & 4 & 0 \\ 7 & 7 & 1 \end{vmatrix} = 0$$

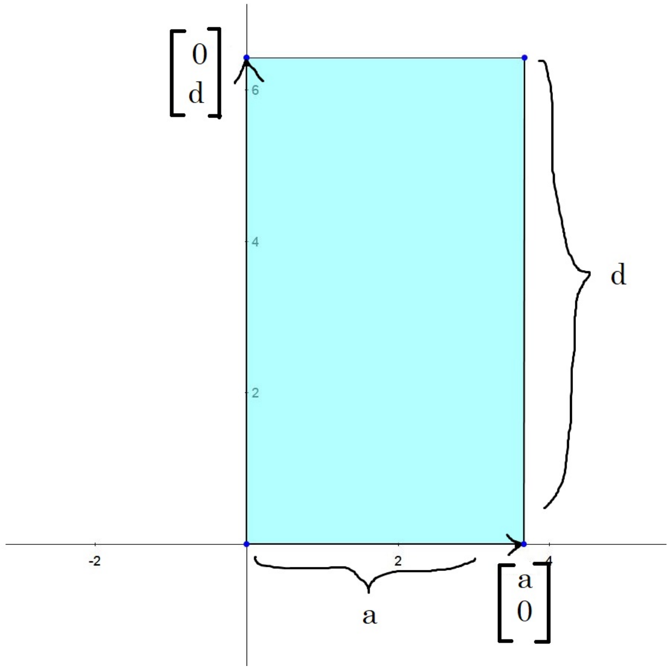
because it has two equal columns, so we could do a column operation that makes a diagonal entry zero, $(c_2 \to c_2 - c_1)$, so det $A = 0$. Another way to see this is to use the theorem: A^T has two equal rows, so det $A^T = 0$ and det $A = 0$.

Determinant as area/volume:
Let A be a 2×2 matrix and consider the parallelogram P spanned by its columns. We are interested in finding the area of P.

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



First we look at the simplest case: if A is diagonal, then $A = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$. The parallelogram that corresponds to this is

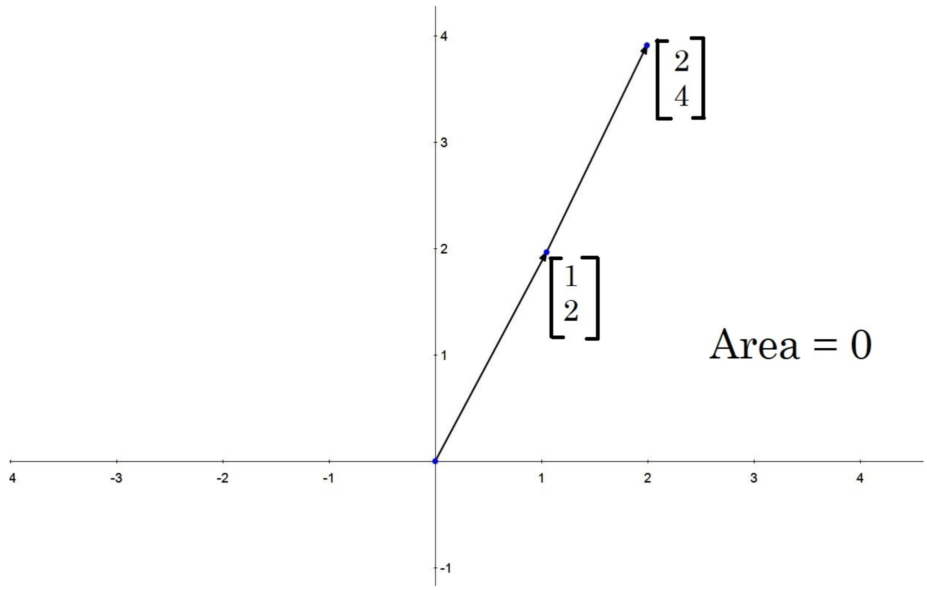


The area of this rectangle is $|ad|$, and the determinant of $\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$ is ad. So

$$\text{Area} = |ad| = |\det\ A| \tag{5}$$

If A is not diagonal, then we can transform A into a diagonal matrix by the elementary row or column operations (swapping 2 columns or rows, or adding a multiple of one column or row to another). These operations will not change the area of the parallelogram. Side-note: If A is not invertible, then one of the columns is a multiple of the other, and so there is no parallelogram, it is just a line, which has an area of zero. For example,

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

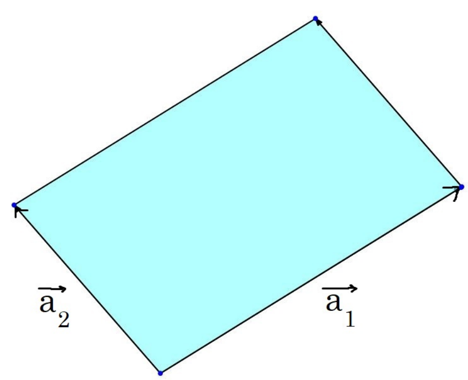


If the columns are linearly dependent, then the area is zero.

So why do the column operations have no effect on the area? If the matrix we're considering is

$$A = \begin{bmatrix} \vec{a_1} & \vec{a_2} \end{bmatrix} \tag{6}$$

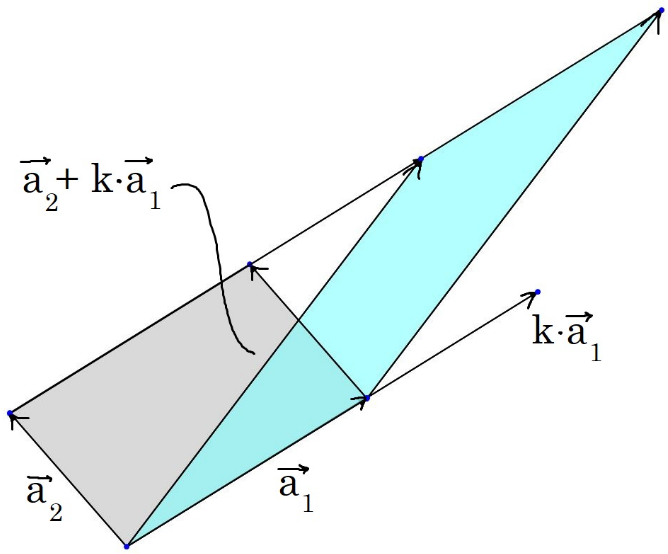
Then, it corresponds to



Now if we add a multiple of one of the columns to the other, to get

$$A' = \begin{bmatrix} \vec{a_1} & \vec{a_2} + k \cdot \vec{a_1} \end{bmatrix} \tag{7}$$

then the parallelogram corresponding to the new matrix is

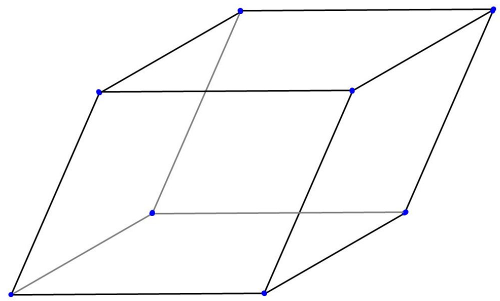


The base of the new parallelogram is still $\vec{a_1}$, and the height is also the same, so this parallelogram has the same area as before. So the column operation did not change the area. This means we can always use the column operations to turn any matrix into a diagonal matrix, and then since we already know that the parallelogram for $\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$ has an area of $|da| = |\det A|$, that is also the area for our original parallelogram.

Theorem: If A is a 2×2 matrix, then the parallelogram spanned by the columns of A has an area of

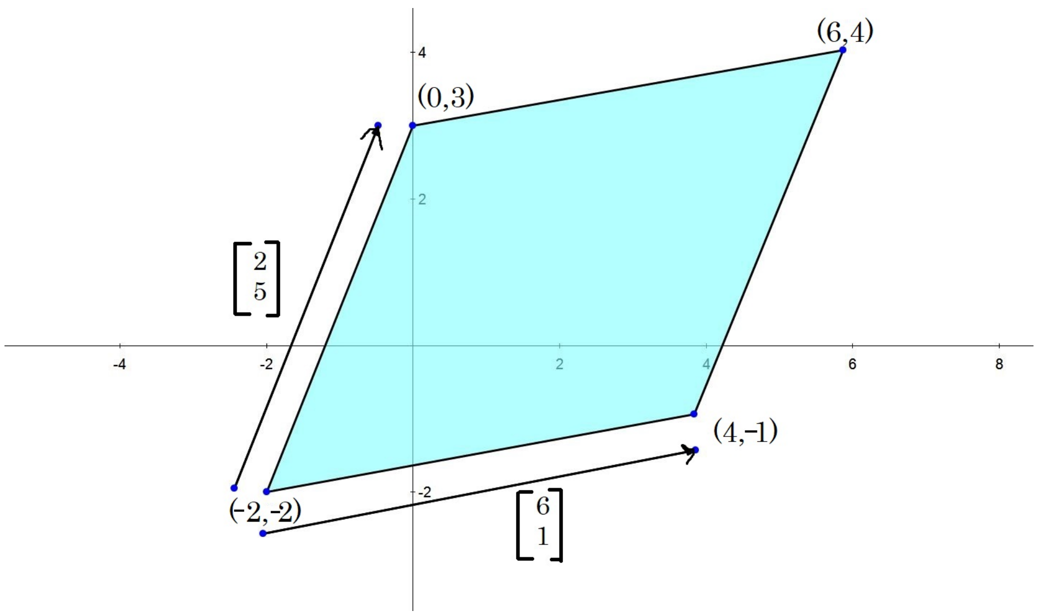
$$|\det A| \tag{8}$$

In three dimensions,



If A is 3×3, then the parallelepiped spanned by its columns has a volume of $|\det A|$.

Example
Find the area of the parallelogram with vectors $(-2,-2), (0,3), (4,-1), (6,4)$.



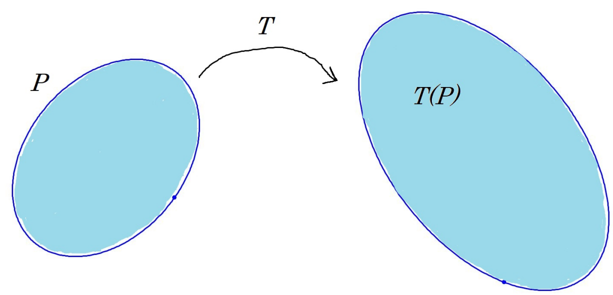
To find the area of this parallelogram, find the vectors that are the sides of the parallelogram, put them in the columns of A, and find $\det A$. So

$$A = \begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} \implies \text{Area } = |\det A| = |6 \cdot 5 - 2 \cdot 1| = 28$$

Linear transformations

Linear transformations can be represented by a standard matrix A, such that $T(\vec{x}) = A\vec{x}$. Given a region P, we can also apply a linear transformation to each point in that region so that

$$T(P) = AP \tag{9}$$



So for the linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^2$,

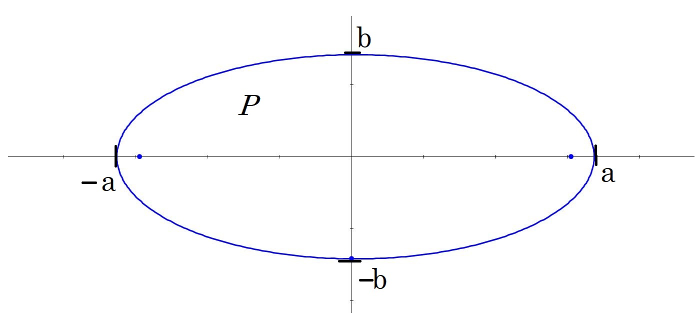
$$\text{Area of } T(P) = |\det\, A| \cdot (\text{Area of P}) \tag{10}$$

And the same thing applies to volume.

Example

Find the area of a region P of $\mathbb{R}^2$ bounded by the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$



Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ so that

$$T(\vec{x}) = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax \\ by \end{bmatrix} \tag{11}$$

This linear transformation maps the unit circle onto the ellipse. So the image of the unit circle under T is P. So the area of the ellipse is

$$\begin{aligned} \text{Area} &= \left| \begin{matrix} a & 0 \\ 0 & b \end{matrix} \right| \cdot (\text{Area of unit circle}) \\ &= ab\pi \end{aligned}$$

This is the area of an ellipse!

Lecture 15

March 21, 2022

Basis of a vector space

Let H be a subspace of a vector space V. An indexed set of vectors

$$B = \left\{\vec{b_1}, \vec{b_2}, \dots, \vec{b_p}\right\}$$

is a **basis** for H if:

1. B is a linearly independent set
2. Span $B = H$

For example, if A is an invertible $n \times n$ matrix, then the columns of A are linearly independent, and the columns span Col $A = \mathbb{R}^n$. So the columns of A form a basis for $\mathbb{R}^n$. Another example is

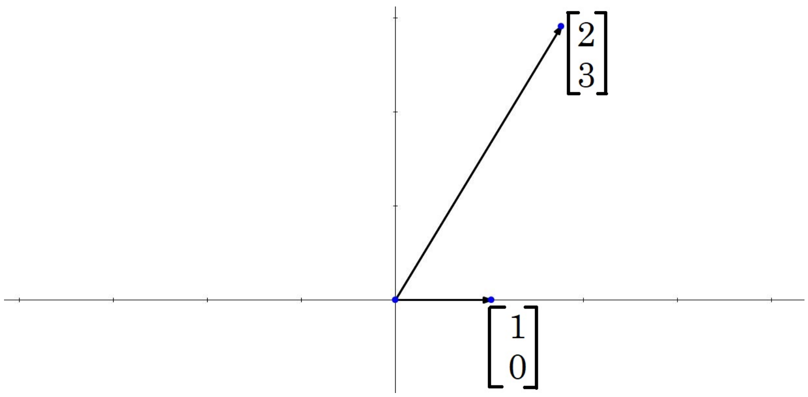
$$I_n = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

The columns are linearly independent, and they span $\mathbb{R}^n$. We call the columns

$$\vec{e_1} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \vec{e_2} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \vec{e_n} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

So $\{\vec{e_1}, \vec{e_2}, \dots, \vec{e_n}\}$ is a basis for $\mathbb{R}^n$, called the **standard basis** for $\mathbb{R}^n$.

Example: Is $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ a basis for $\mathbb{R}^2$?



We could check properties 1 and 2. Or we could put these vectors in the columns of A, and check if it is invertible.

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$$

The determinant of this matrix is det $A = 3 - 0 = 3$, which implies that A is invertible because det $A \neq 0$. So Col $A = \mathbb{R}^2$ and $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^2$.

Another example: Let $\mathbb{P}_n$ = set of all polynomials of degree $\leq n$. This is a vector space. Then

$$S = \{1, t, t^2, \ldots, t^n\}$$

is the **standard basis** for $\mathbb{P}_n$. Check the properties:

1. Span $S = \mathbb{P}_n$ because an arbitrary vector in $\mathbb{P}_n$ has the form:

$$\vec{p}(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$$

So every vector is a linear combination of the elements of S, with weights $a_0, a_1, \ldots, a_n$.

2. Linear independence: We cannot write any of the elements as a linear combination of the others, because no t^k is a linear combination of $\{1, t, t^2, \ldots, t^{k-1}\}$. That is, for any choice of coefficients,

$$t^k \neq c_0 + c_1 t + \cdots + c_{k-1} t^{k-1}$$

So the vectors are linearly independent, and S is a basis for $\mathbb{P}_n$.

Another example: Is $\left\{ \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} \right\}$ a basis for $\mathbb{R}^3$?

$$A = \begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ -6 & 7 & 5 \end{bmatrix} \text{ row reduce} \rightarrow \begin{bmatrix} 3 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

This is invertible because there is a pivot in each column, also because it is row reduce-able to the identity matrix, and also because $\det A = 3 \cdot 2 \cdot 1 = 6$. (Three different ways of showing the same thing). So A is invertible, and the columns of A form a basis for Col $A = \mathbb{R}^3$.

Example of finding a spanning set

Let

$$\vec{v}_1 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 6 \\ 16 \\ -5 \end{bmatrix}$$

and

$$H = \text{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$$

Find a basis for H. We have to check that $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are linearly independent. Are they? No, because

$$\vec{v}_3 = 5\vec{v}_1 + 3\vec{v}_2$$

So we can write $\vec{v}_3$ as a linear combination of $\vec{v}_1$ and $\vec{v}_2$, and $\vec{v}_3 \in \text{Span}\{\vec{v}_1, \vec{v}_2\}$. This means that $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is not a linearly independent set. Since $\vec{v}_3$ is contained in the span of $\vec{v_1}$ and $\vec{v_2}$, the vector $\vec{v}_3$ is redundant.
So $\text{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \text{Span}\{\vec{v}_1, \vec{v}_2\}$, and

$$H = \text{Span}\{\vec{v}_1, \vec{v}_2\}$$

Then, since $\vec{v}_2$ is not a multiple of $\vec{v}_1$, they are linearly independent. So $\{\vec{v}_1, \vec{v}_2\}$ is a basis for H. In general,

1. Start with the spanning set
2. Check if the vectors are linearly independent

3. Remove redundant vectors

4. If the vectors are linearly independent, you have the basis.

Spanning Set Theorem

Let $S = \{\vec{v}_1, \dots, \vec{v}_p\}$ be a set in a vector space V and let

$$H = \text{Span}\,\{\vec{v}_1, \dots, \vec{v}_p\} \tag{1}$$

Then

1. If one vector in S is a linear combination of the others, then the set formed by removing that vector from S spans H.

2. If $H \neq \left\{\vec{0}\right\}$ then some subset of S is a basis for H.

Main idea: Keep removing redundant vectors until the set is linearly independent, then part (1) says that this set still spans H, so part (2) implies that this set is a basis for H.

The basis for a null space

Find a basis for Nul A if

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

The null space is the set of all $\vec{x}$ such that $A\vec{x} = \vec{0}$, so we look at

$$A \to \text{row reduce} \to \begin{bmatrix} \mathbf{1} & -2 & 0 & -1 & 3 \\ 0 & 0 & \mathbf{1} & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \vec{0}$$

The variables x_2, x_4, x_5 are free, and x_1, x_3 are basic variables.

$$\implies \begin{cases} x_1 - 2x_2 - x_4 + 3x_5 = 0 \\ x_3 + 2x_4 - 2x_5 = 0 \end{cases}$$

$$\implies \begin{cases} x_1 = 2x_2 + x_4 - 3x_5 \\ x_3 = -2x_4 + 2x_5 \end{cases}$$

So $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix}$. To find the spanning set, we separate this out in terms of the free variables.

$$\vec{x} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

So $\text{Nul } A = \text{Span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}$.

These vectors are linearly independent if the equation

$$x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} = \vec{0}$$

has only the trivial solution. From the definition of $\vec{x}$ this equation is equivalent to

$$\vec{x} = \vec{0}$$

$$\begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

This requires that $x_2 = 0, x_4 = 0$ and $x_5 = 0$ which also implies that $x_1 = 0$ and $x_3 = 0$. So the trivial solution $\vec{x} = \vec{0}$ is the only possible solution that could satisfy the equation. So the vectors are linearly independent, and

$$\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\} \text{ is a basis for Nul } A.$$

Our standard procedure for finding vectors that span Nul A will always produce a basis for Nul A. (The final set just needs to be linearly independent).

The basis for a column space

Find a basis for Col B if

$$B = \begin{bmatrix} \mathbf{1} & 4 & 0 & 2 & 0 \\ 0 & 0 & \mathbf{1} & -1 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The free variables, (no pivots), imply that these vectors are dependent, because we can write the columns without pivots as linear combinations of the columns with pivots.

$$\begin{aligned} \vec{b}_2 &= 4\vec{b}_1 \\ \vec{b}_4 &= 2\vec{b}_1 - \vec{b}_3 \end{aligned}$$

So $\vec{b}_2$ and $\vec{b}_4$ are redundant, and we can remove them. Then, $\vec{b}_1$, $\vec{b}_3$ and $\vec{b}_5$ are linearly independent because each one has a pivot and so cannot be written as a linear combination of the others. So $\left\{\vec{b}_1, \vec{b}_3, \vec{b}_5\right\}$ is a basis for Col B. (By the spanning set theorem)

Another example: Find a basis for Col A if

$$A = \begin{bmatrix} \mathbf{1} & 4 & 0 & 2 & -1 \\ 3 & 12 & \mathbf{1} & 5 & 5 \\ 2 & 8 & 1 & 3 & \mathbf{2} \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}$$

Note that A row reduces to B from the previous example. Then, linear dependence (or independence) of the columns of A corresponds to solutions of $A\vec{x} = \vec{0}$, where the left side of that equation represents linear combinations of the columns of A with weights given by $\vec{x}$. If there is a nonzero solution $\vec{x}$, then the columns of A are linearly dependent. If the only solution is $\vec{x} = \vec{0}$, then the columns of A are linearly independent. Since $A \sim B$, the equations $A\vec{x} = \vec{0}$ and $B\vec{x} = \vec{0}$ have the same solutions, that is, the same vector $\vec{x}$ and hence the same weights on the columns. The basis for B was $\left\{\vec{b}_1, \vec{b}_3, \vec{b}_5\right\}$. So the basis for Col A is

$$\{\vec{a}_1, \vec{a}_3, \vec{a}_5\} = \left\{ \begin{bmatrix} 1 \\ 3 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 2 \\ 8 \end{bmatrix} \right\}$$

Any linear dependency among columns of B are also there among columns of A. So the vectors used in the basis above are the columns of A, not B. The pivot columns of A will always be linearly independent so the pivot columns of A form a basis for Col A.

Lecture 16

March 23, 2022

Warm up

Let

$$\vec{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \vec{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \text{ and } \vec{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

Write $\vec{x}$ as a linear combination of $\vec{b}_1$ and $\vec{b}_2$. First, $\vec{b}_1$ and $\vec{b}_2$ form a basis for $\mathbb{R}^2$ because they are linearly independent and $\begin{bmatrix} \vec{b}_1 & \vec{b}_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$ has a nonzero determinant $\det A = 2 + 1 = 3$, so it is invertible, and the columns must span $\mathbb{R}^2$. Now we want to find values of c_1, c_2 such that

$$\begin{aligned} c_1\vec{b}_1 + c_2\vec{b}_2 &= \vec{x} \\ \begin{bmatrix} \vec{b}_1 & \vec{b}_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \vec{x} \\ \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} 4 \\ 5 \end{bmatrix} \end{aligned}$$

Option 1: Augment this matrix, row reduce, and solve.
Option 2: Find the inverse matrix and multiply it on both sides of the equation.

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

So

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

Then, we multiply it on each side of the earlier equation to undo that matrix

$$\frac{1}{3}\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 5 \end{bmatrix}$$
$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 9 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

This means that $c_1 = 3, c_2 = 2$. So

$$\vec{x} = 3\vec{b}_1 + 2\vec{b}_2$$

Change in coordinates

A basis is a "coordinate system" for a vector space.

Unique representation therorem: Let $B = \left\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\right\}$ be a basis for a vector space V. Then for each $\vec{x} \in V$, there exists a unique set of scalars $c_1, c_2, \ldots, c_n$ such that

$$\vec{x} = c_1\vec{b}_1 + c_2\vec{b}_2 + \cdots + c_n\vec{b}_n \tag{2}$$

In other words, every vector in V can be written as a linear combination of vectors in B in a unique way.

Definition: $B = \left\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\right\}$ is a basis for V and $\vec{x} \in V$. The **coordinates of $\vec{x}$ relative to** B (also called the B-coordinates of $\vec{x}$), are the weights $c_1, c_2, \ldots, c_n$ such that

$$\vec{x}_1 = c_1\vec{b}_1 + c_2\vec{b}_2 + \cdots + c_n\vec{b}_n \tag{3}$$

We write the coordinate vector of $\vec{x}$ relative to B as

$$\left[\vec{x}\right]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \tag{4}$$

The map $V \to \mathbb{R}^n$ which sends $\vec{x}$ to $\left[\vec{x}\right]_B$ is called the **coordinate mapping** determined by B.

Example: $B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^2$, (because the determinant is nonzero). Suppose $\vec{x} \in \mathbb{R}^2$ has $\left[\vec{x}\right]_B = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Find $\vec{x}$.

$$\begin{aligned} \vec{x} &= c_1\vec{b}_1 + c_2\vec{b}_2 \\ &= -2\vec{b}_1 + 3\vec{b}_2 \\ &= -2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 6 \end{bmatrix} \end{aligned}$$

We call the **standard basis** for $\mathbb{R}^n$,

$$\mathcal{E} = \{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\} \tag{5}$$

For a vector $\vec{x} \in \mathbb{R}^n$,

$$\left[\vec{x}\right]_{\mathcal{E}} = \vec{x} \tag{6}$$

because for example if

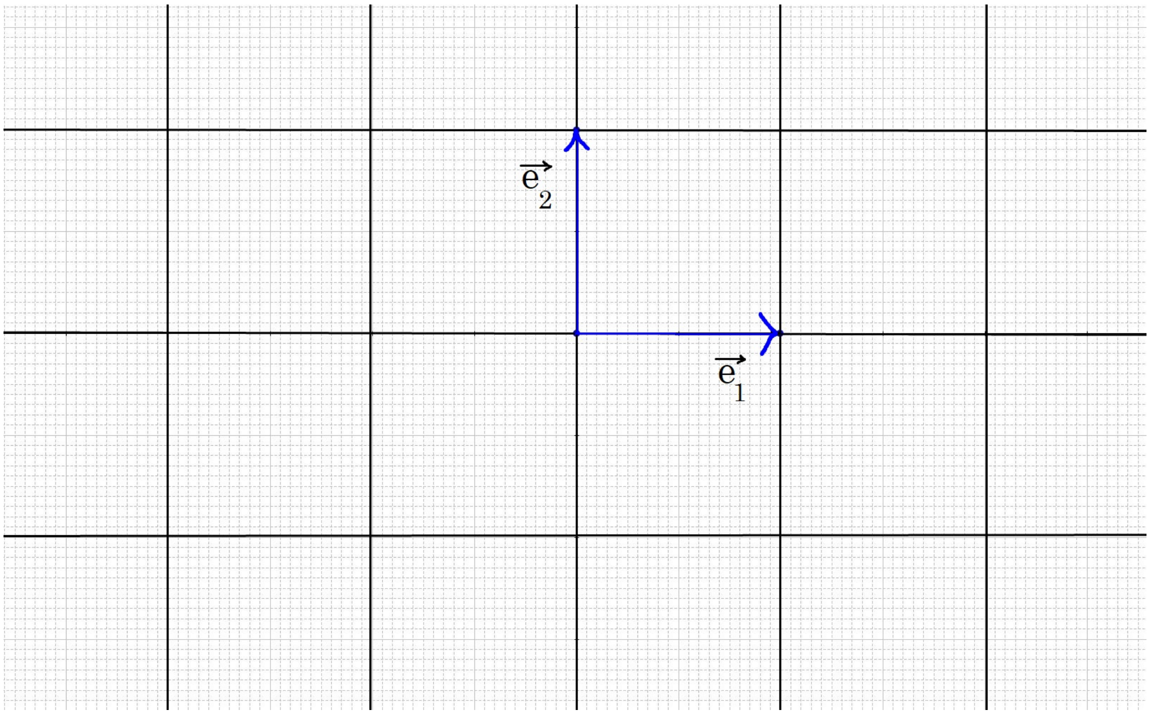
$$\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

We use $\mathcal{E} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$, so $\left[\vec{x}\right]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \vec{x}$.

So for simplicity, we tend not to write $\left[\vec{x}\right]_{\mathcal{E}}$, instead we just write $\vec{x}$.

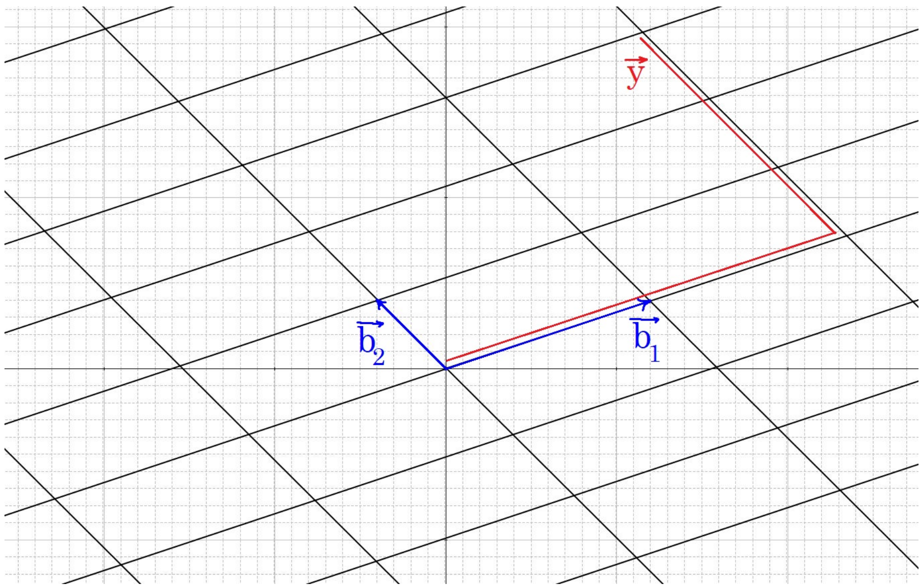
Geometry:

$$\left(\mathbb{R}^2, \mathcal{E}\right)$$



But then if $\vec{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \vec{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ such that $B = \left\{ \vec{b}_1, \vec{b}_2 \right\}$ then

$(\mathbb{R}^2, B)$



The example vector shown in the figure is $\left[\vec{y}\right]_B = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. To find this, move over $2\vec{b}_1$ and move up $3\vec{b}_2$. This still covers the whole space.

From warmup:

$$\begin{bmatrix} \vec{b}_1 & \vec{b}_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \vec{x} \tag{7}$$

We call $P_B = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 \end{bmatrix}$ the **change of coordinates matrix**, where $B = \left\{\vec{b}_1, \vec{b}_2\right\}$. So that

$$P_B \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \vec{x} \tag{8}$$

Then,

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \left[\vec{x}\right]_B$$

So

$$P_B\left[\vec{x}\right]_B = \vec{x} \tag{9}$$

This is why P_B has its name: if you take the coordinate vector of $\vec{x}$ relative to B and multiply it by the **change of coordinates matrix**, you get the standard coordinates of $\vec{x}$. Then, to find $\left[\vec{x}\right]_B$, we multiply both sides by P_B^{-1}:

$$P_B^{-1}P_B\left[\vec{x}\right]_B = P_B^{-1}\vec{x}$$

$$\left[\vec{x}\right]_B = P_B^{-1}\vec{x} \tag{10}$$

If $B = \left\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\right\}$ is a basis for $\mathbb{R}^n$, then $P_B = \begin{bmatrix}\vec{b}_1 & \vec{b}_2 & \ldots & \vec{b}_n\end{bmatrix}$ is the change of coordinate matrix. Again, it changes the B-coordinates of $\vec{x}$ to the standard coordinates of $\vec{x}$ such that $P_B\left[\vec{x}\right]_B = \vec{x} = \left[\vec{x}\right]_{\mathcal{E}}$

$$\begin{tikzcd}(R^n, B) \arrow[rr, bend left, "P_B"] & & (R^n, \mathcal{E}) \arrow[ll, bend left, "P_B^{-1}"]\end{tikzcd}$$

Remarks:

1. P_B is always invertible, (because its columns are linearly independent).

2. P_B^{-1} changes the standard coordinates into B coordinates, so that $\left[\vec{x}\right]_B = P_B^{-1}\vec{x}$.

Example

Consider the vector space $\mathbb{P}_2$, and the basis $B = \{1, t, t^2\}$. The vector

$$\begin{aligned}\vec{p}(t) &= 2 + 3t + 4t^2 \\ &= 2(1) + 3(t) + 4(t^2)\end{aligned}$$

is an element of $\mathbb{P}_2$. We can write the coordinate vector of $\vec{p}$ relative to B as

$$\left[\vec{p}\right]_B = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \in \mathbb{R}^3$$

because those are the coefficients on the elements of the basis.

For any vector space V and any $\vec{x} \in V$, no matter what V is,

$$\left[\vec{x}\right]_B \in \mathbb{R}^n \tag{11}$$

where B is a basis for V. This is an important result because it means that we can use real numbers to express an element of an abstract vector space if we just do a change of coordinates.

Lecture 17

March 28, 2022

More on the Basis of a vector space

Example: The vector space $\mathbb{P}_2$ has a basis $B = \{1, t, t^2\}$. One element of this vector space is

$$\vec{p}(t) = 2 + 3t + 4t^2$$

The B-coordinates of this vector are the coefficients on the elements of the basis.

$$[\vec{p}(t)]_B = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \in \mathbb{R}^3$$

The fact that these coordinates are all real numbers is true in general. For any vector space V, $\vec{x} \in V$,

$$[\vec{x}]_B \in \mathbb{R}^n \tag{1}$$

(If B has n vectors). So the vector $\vec{x}$ gets mapped to $[\vec{x}]_B$, and the vector space V gets mapped to $\mathbb{R}^n$.

Theorem: Let V be a vector space with basis $B = \left\{\vec{b}_1, \ldots, \vec{b}_n\right\}$. The coordinate mapping $\vec{x} \to [\vec{x}]_B$ is an isomorphism (that is, it is a linear transformation that is both injective and surjective). This is visualized in Figure 1.

Example: The vector space $\mathbb{P}_2$ has a basis $B = \{1, t, t^2\}$ and an element $p(t) \in \mathbb{P}_2$. The coordinate mapping

$$p(t) \to [p(t)]_B \tag{2}$$

implies by the theorem that

$$\mathbb{P}_2 \to \mathbb{R}^3 \tag{3}$$

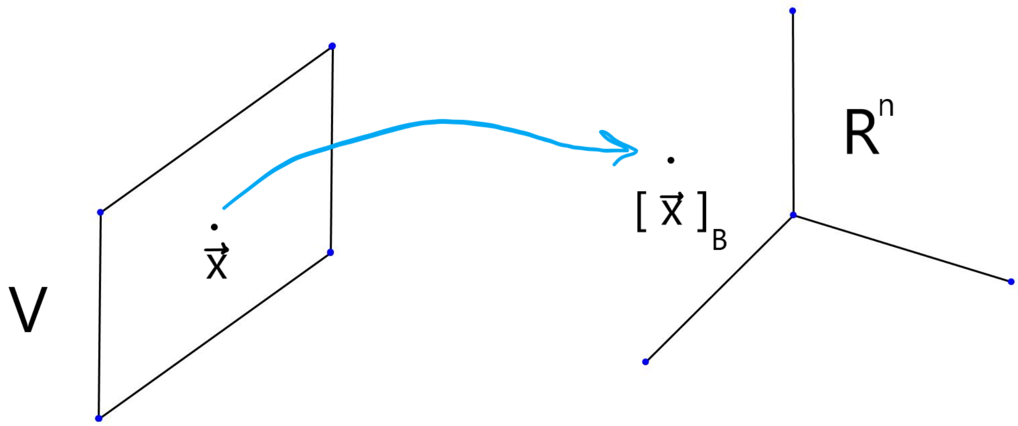


Figure 1: Coordinate mapping is an isomorphism

So, $\mathbb{P}_2$ is isomorphic to $\mathbb{R}^3$. Why is $\vec{x} \to [\vec{x}]_B$ an isomorphism? In general for a linear map T, the following equations are true.

$$T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) \tag{4}$$

$$T(c\vec{u}) = c \cdot T(\vec{u}) \tag{5}$$

So we just need to check that both of these conditions hold for the coordinate mapping.

1. Let $\vec{u}, \vec{v} \in V$ and the basis be $B = \left\{\vec{b}_1, \ldots, \vec{b}_n\right\}$. We want $[\vec{u} + \vec{v}]_B = [\vec{u}]_B + [\vec{v}]_B$. Since the vectors belong to the vector space, we can write them as linear combinations of the elements of the basis, so that

$$\begin{cases} \vec{u} = c_1\vec{b}_1 + c_2\vec{b}_2 + \cdots + c_n\vec{b}_n \\ \vec{v} = d_1\vec{b}_1 + d_2\vec{b}_2 + \cdots + d_n\vec{b}_n \end{cases}$$

This implies that the B-coordinates are given by

$$[\vec{u}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \quad [\vec{v}]_B = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$$

Then the sum of the vectors is

$$\vec{u} + \vec{v} = (c_1 + d_1)\vec{b}_1 + (c_2 + d_2)\vec{b}_2 + \cdots + (c_n + d_n)\vec{b}_n \tag{6}$$

which means that

$$[\vec{u} + \vec{v}]_B = \begin{bmatrix} c_1 + d_1 \\ c_2 + d_2 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = [\vec{u}]_B + [\vec{v}]_B \tag{7}$$

So the first property of a linear transformation is met.

2. Let $\vec{u} \in V$, and $k \in \mathbb{R}$. We want $[k\vec{u}]_B = k\,[\vec{u}]_B$. Let $\vec{u} = c_1\vec{b}_1 + \cdots + c_n\vec{b}_n$. Then

$$k\vec{u} = k\left(c_1\vec{b}_1 + \cdots + c_n\vec{b}_n\right) = (kc_1)\vec{b}_1 + (kc_2)\vec{b}_2 + \cdots + (kc_n)\vec{b}_n$$

which implies that

$$[k\vec{u}]_B = \begin{bmatrix} kc_1 \\ kc_2 \\ \vdots \\ kc_n \end{bmatrix} = k\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = k\,[\vec{u}]_B$$

So the second property is also met.

Example problem: Verify that the vectors

$$\begin{aligned} \vec{p}_1(t) &= 1 + 2t \\ \vec{p}_2(t) &= 4 + t + 5t^2 \\ \vec{p}_3(t) &= 3 + 2t \end{aligned}$$

are linearly dependent in $\mathbb{P}_2$. Because the coordinate mapping is an isomorphism, the set of vectors $\{\vec{p}_1, \vec{p}_2, \vec{p}_3\}$ is linearly dependent in $\mathbb{P}_2$ if and only if $\{\,[\vec{p}_1]_B\,, [\vec{p}_2]_B\,, [\vec{p}_3]_B\,\}$ is linearly dependent in $\mathbb{R}^3$. The basis of $\mathbb{P}_2$ is $B = \{1, t, t^2\}$, so we can take the B-coordinates of each vector to be the coefficients on the terms.

$$\{\,[\vec{p}_1]_B\,, [\vec{p}_2]_B\,, [\vec{p}_3]_B\,\} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} \right\}$$

Now we just need to check if these vectors are dependent. Put them in the columns of a matrix.

$$\begin{bmatrix} 1 & 4 & 3 \\ 2 & 1 & 2 \\ 0 & 5 & 0 \end{bmatrix} \rightarrow \text{row reduce} \rightarrow \begin{bmatrix} \mathbf{1} & 4 & 3 \\ 0 & \mathbf{1} & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

There is not a pivot in every column, so these vectors are linearly dependent, which implies that the original vectors $\{\vec{p}_1, \vec{p}_2, \vec{p}_3\}$ are also linearly dependent in $\mathbb{P}_2$.

Another example:

$$\vec{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}, \ \vec{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \ \vec{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}, \ B = \{\vec{v}_1, \vec{v}_2\}$$

a.) Check that B is a basis for

$$H = \text{Span}\,\{\vec{v}_1, \vec{v}_2\}$$

By definition, the vectors in B span H, but we also need to check that the vectors are independent. The vector $\vec{v}_2 \neq k\vec{v}_1$ for any scalar k, so $\vec{v}_1$ and $\vec{v}_2$ are linearly independent, which implies that B is a basis for H.

b.) Determine if $\vec{x} \in H$. In order for $\vec{x}$ to be in H, it must be a linear combination of $\vec{v}_1$ and $\vec{v}_2$, such that

$$\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2$$

which we can rewrite as

$$\begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \vec{x}$$

$$\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & & \vec{x} \end{bmatrix}$$

$$\begin{bmatrix} 3 & -1 & & 3 \\ 6 & 0 & & 12 \\ 2 & 1 & & 7 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} \mathbf{1} & 0 & & 2 \\ 0 & \mathbf{1} & & 3 \\ 0 & 0 & & 0 \end{bmatrix}$$

This system is consistent, so there does exist a linear combination of vectors $\vec{v}_1$ and $\vec{v}_2$ that produces $\vec{x}$, so yes $\vec{x} \in H$.

c.) Find $[\vec{x}]_B$. Recall that $B = \{\vec{v}_1, \vec{v}_2\}$. Then, from the previous part of this problem, we found that $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ such that

$$\begin{aligned} \vec{x} &= c_1\vec{v}_1 + c_2\vec{v}_2 \\ &= 2\vec{v}_1 + 3\vec{v}_2 \end{aligned}$$

which means that the coordinate vector of $\vec{x}$ relative to B is $[\vec{x}]_B = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

Change of Basis

If B and C are bases for a vector space V and $\vec{x} \in V$, how are $[\vec{x}]_B$ and $[\vec{x}]_C$ related? Consider the following example. Given the bases $B = \left\{\vec{b}_1, \vec{b}_2\right\}$ and $C = \{\vec{c}_1, \vec{c}_2\}$, let

$$\begin{aligned} \vec{b}_1 &= 4\vec{c}_1 + \vec{c}_2 \\ \vec{b}_2 &= -6\vec{c}_1 + \vec{c}_2 \\ [\vec{x}]_B &= \begin{bmatrix} 3 \\ 1 \end{bmatrix} \end{aligned}$$

Find $[\vec{x}]_C$. Firstly, $[\vec{x}]_B = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ implies that $\vec{x} = 3\vec{b}_1 + \vec{b}_2$. Then since coordinate mapping is an isomorphism,

$$[\vec{x}]_C = \left[3\vec{b}_1 + \vec{b}_2\right]_C = 3\left[\vec{b}_1\right]_C + \left[\vec{b}_2\right]_C = \left[\left[\vec{b}_1\right]_C \quad \left[\vec{b}_2\right]_C\right] \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

where the thing on the left is a matrix whose columns are $\left[\vec{b}_1\right]_C$ and $\left[\vec{b}_2\right]_C$. This just represents taking three times $\vec{b}_1$ in terms of C and adding one times $\vec{b}_2$. Then, since $\vec{b}_1 = 4\vec{c}_1 + \vec{c}_2$ and $\vec{b}_2 = -6\vec{c}_1 + \vec{c}_2$, we can write

$$\left[\vec{b}_1\right]_C = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \quad \text{and} \quad \left[\vec{b}_2\right]_C = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$$

So the equation above becomes

$$[\vec{x}]_C = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

which means $\vec{x} = 6\vec{c}_1 + 4\vec{c}_2$. Note that the vector $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ was $[\vec{x}]_B$. Then the matrix we multiplied by $[\vec{x}]_B$, which effectively translated the B-coordinates into C-coordinates, is written as $\underset{C \leftarrow B}{P}$. This example motivates the following theorem.

Theorem:
If $B = \left\{\vec{b}_1, \ldots, \vec{b}_n\right\}$ and $C = \{\vec{c}_1, \ldots, \vec{c}_n\}$ are bases of a vector space V, then there exists an $n \times n$ matrix $\underset{C \leftarrow B}{P}$ such that for any $\vec{x} \in V$,

$$[\vec{x}]_C = \underset{C \leftarrow B}{P} [\vec{x}]_B \tag{8}$$

where $\underset{C \leftarrow B}{P}$ is called the **change of coordinates matrix**, and

$$\underset{C \leftarrow B}{P} = \left[\left[\vec{b}_1\right]_C \quad \left[\vec{b}_2\right]_C \quad \cdots \quad \left[\vec{b}_n\right]_C \right] \tag{9}$$

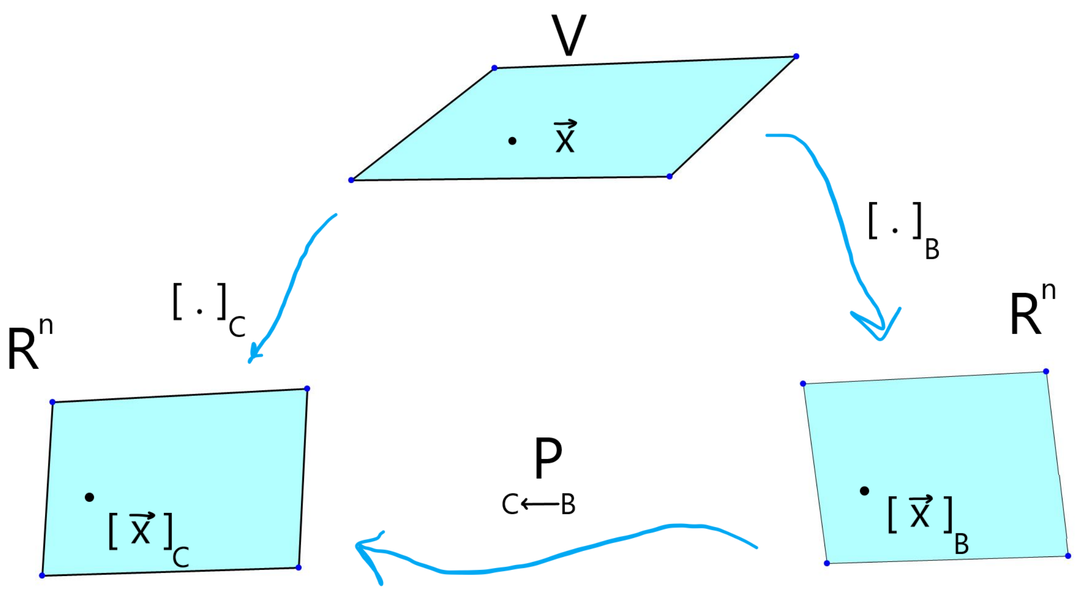


Figure 2: We can change between bases for a vector space V

To go the other way, from C to B,

$$\underset{B \leftarrow C}{P} = \underset{C \leftarrow B}{P}^{-1} \tag{10}$$

which follows from Equation (8).

Lecture 18

March 30, 2022

More on changing coordinates

Recall that the change of coordinates matrix is the matrix $\underset{C \leftarrow B}{P}$ such that

$$[\,\vec{x}\,]_C = \underset{C \leftarrow B}{P}\,[\,\vec{x}\,]_B \tag{1}$$

and its columns are the vectors in B relative to C.

$$\underset{C \leftarrow B}{P} = \begin{bmatrix} \begin{bmatrix} \vec{b}_1 \end{bmatrix}_C & \begin{bmatrix} \vec{b}_2 \end{bmatrix}_C & \dots & \begin{bmatrix} \vec{b}_n \end{bmatrix}_C \end{bmatrix} \tag{2}$$

Remarks:

1. The columns of $\underset{C \leftarrow B}{P}$ are linearly independent (because the columns of B are linearly independent and $[\ ]_C$ is a linear transformation. So after they get mapped to C-coordinates, they are still independent).

2. The matrix $\underset{C \leftarrow B}{P}$ is invertible and the inverse is

$$\underset{C \leftarrow B}{P}^{-1} = \underset{B \leftarrow C}{P} \tag{3}$$

3. If $V = \mathbb{R}^n$ and $\mathcal{E}$ is the standard basis, (so that $\mathcal{E} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$), then $\underset{\mathcal{E} \leftarrow B}{P} = P_B$, which is the change of coordinates matrix from Lecture 16. Recall that

$$P_B = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_n \end{bmatrix} \quad \text{and} \quad \vec{b}_i = \begin{bmatrix} \vec{b}_i \end{bmatrix}_{\mathcal{E}} \tag{4}$$

then if you multiply this matrix by any vector in B-coordinates, the result will give you that vector in standard coordinates.

Example: Let

$$\vec{b}_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix},\ \vec{b}_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix},\ \vec{c}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix},\ \vec{c}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$$

and

$$B = \left\{\vec{b}_1, \vec{b}_2\right\},\ C = \{\vec{c}_1, \vec{c}_2\}$$

Find the change of coordinates matrix from B to C, $\underset{C\leftarrow B}{P}$. To find this matrix, we need to write the vectors in B in terms of those in C. That is, we need to find values of $\alpha_1, \alpha_2, \beta_1, \beta_2$ such that

$$\alpha_1\vec{c}_1 + \alpha_2\vec{c}_2 = \vec{b}_1$$
$$\beta_1\vec{c}_1 + \beta_2\vec{c}_2 = \vec{b}_2$$

We can rewrite each equation as

$$\begin{bmatrix} \vec{c}_1 & \vec{c}_2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \vec{b}_1$$
$$\begin{bmatrix} \vec{c}_1 & \vec{c}_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \vec{b}_2$$

and the goal is to solve both equations simultaneously. By matrix multiplication,

$$\begin{bmatrix} \vec{c}_1 & \vec{c}_2 \end{bmatrix} \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 \end{bmatrix} \tag{5}$$

To solve for this matrix, we could just as well solve the augmented matrices for $\vec{b}_1$ and $\vec{b}_2$ separately, $\left[\begin{array}{cc:c} \vec{c}_1 & \vec{c}_2 & \vec{b}_1 \end{array}\right]$ and $\left[\begin{array}{cc:c} \vec{c}_1 & \vec{c}_2 & \vec{b}_2 \end{array}\right]$. This will give you $\left[\begin{array}{cc:c} 1 & 0 & \alpha_1 \\ 0 & 1 & \alpha_2 \end{array}\right]$, and $\left[\begin{array}{cc:c} 1 & 0 & \beta_1 \\ 0 & 1 & \beta_2 \end{array}\right]$. Then, you will have taken the same steps to row reduce both matrices, so if you had instead applied those steps to $\left[\begin{array}{cc:cc} \vec{c}_1 & \vec{c}_2 & \vec{b}_1 & \vec{b}_2 \end{array}\right]$, then the result after row reducing will be $\left[\begin{array}{cc:cc} 1 & 0 & \alpha_1 & \beta_1 \\ 0 & 1 & \alpha_2 & \beta_2 \end{array}\right]$. The following is a proof that this is true. Based on Equation 5,

$$\begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} = \begin{bmatrix} \vec{c}_1 & \vec{c}_2 \end{bmatrix}^{-1} \begin{bmatrix} \vec{b}_1 & \vec{b}_2 \end{bmatrix} \tag{6}$$

Apply the algorithm for finding an inverse matrix.

$$\left[\vec{c}_1 \quad \vec{c}_2 \;\middle|\; I_2\right] \to \left[I_2 \;\middle|\; \begin{bmatrix}\vec{c}_1 & \vec{c}_2\end{bmatrix}^{-1}\right]$$

Then if we multiply the right part of the augmented matrix by $\begin{bmatrix}\vec{b}_1 & \vec{b}_2\end{bmatrix}$ and carry that through, we get

$$\left[\vec{c}_1 \quad \vec{c}_2 \;\middle|\; \vec{b}_1 \quad \vec{b}_2\right] \to \left[I_2 \;\middle|\; \begin{bmatrix}\vec{c}_1 & \vec{c}_2\end{bmatrix}^{-1}\begin{bmatrix}\vec{b}_1 & \vec{b}_2\end{bmatrix}\right]$$

Based on Equation 6, we can rewrite this as

$$\left[\vec{c}_1 \quad \vec{c}_2 \;\middle|\; \vec{b}_1 \quad \vec{b}_2\right] \to \left[\begin{array}{c|cc} I_2 & \alpha_1 & \beta_1 \\ & \alpha_2 & \beta_2 \end{array}\right] \tag{7}$$

So we have finished the proof: by row-reducing the augmented matrix, we can solve for the values that we're after. In the problem given,

$$\left[\vec{c}_1 \quad \vec{c}_2 \;\middle|\; \vec{b}_1 \quad \vec{b}_2\right] = \left[\begin{array}{cc|cc} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right]$$

$$\text{Row reduce } \to \left[\begin{array}{cc|cc} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right]$$

So

$$\begin{aligned} \alpha_1 &= 6, \ \beta_1 = 4 \\ \alpha_2 &= -5, \ \beta_2 = -3 \end{aligned}$$

which implies that

$$\begin{aligned} \vec{b}_1 &= 6\vec{c}_1 - 5\vec{c}_2 \\ \vec{b}_2 &= 4\vec{c}_1 - 3\vec{c}_2 \end{aligned}$$

so that the C-coordinates of $\vec{b}_1$ and $\vec{b}_2$ are given by the coefficients on these vectors.

$$\left[\vec{b}_1\right]_C = \begin{bmatrix} 6 \\ -5 \end{bmatrix} \quad \text{and} \quad \left[\vec{b}_2\right]_C = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

This means the change of coordinates matrix is

$$\underset{C \leftarrow B}{P} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}$$

which, notice, is exactly what was on the right side of the augmented matrix. This example motivates the following method.

General Method:

Row reducing $\left[\begin{array}{cccc:cccc} \vec{c}_1 & \vec{c}_2 & \dots & \vec{c}_n & \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_n \end{array}\right]$ yields

$$\left[\begin{array}{c:c} I_n & \underset{C \leftarrow B}{P} \end{array}\right] \tag{8}$$

Dimension of a vector space

Theorem: If $B = \left\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\right\}$ is a basis for a vector space V, then any set of vectors in V of *size more than* n must be linearly dependent. The idea of the proof is this: the coordinate mapping $[\ \ ]_B$ is an isomorphism: $V \to \mathbb{R}^n$. So the set of vectors $\{\vec{v}_1, \dots, \vec{v}_p\}$ in V, with $p > n$, gets mapped to $\left\{[\ \vec{v}_1\ ]_B, \dots, [\ \vec{v}_p\ ]_B\right\}$, which is a set of vectors in $\mathbb{R}^n$, and it is linearly independent if and only if the original set is. So to show the theorem, we just need to show that $\left\{[\ \vec{v}_1\ ]_B, \dots, [\ \vec{v}_p\ ]_B\right\}$ is dependent. For example, if $p = 3$ and $n = 2$, then we can put the B-coordinate vectors in the columns of a matrix to test their independence. Assuming the maximum number of pivots possible, the matrix will look like something of the form

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

The third column does not have a pivot, so it can be written as a linear combination of the others, and hence the columns are linearly dependent. Another way to see this is that, in general, more vectors than entries in each vector implies that the set of columns is linearly dependent in $\mathbb{R}^n$. Then, since $[\ \ ]_B$ is an isomorphism, the set $\{\vec{v}_1, \dots, \vec{v}_p\}$ is also linearly dependent.

Fact: If $T \colon V \to W$ is an isomorphism, then $\{\vec{v}_1, \dots, \vec{v}_k\}$ is linearly independent in V if and only if $\{T(\vec{v}_1), \dots, T(\vec{v}_k)\}$ is linearly independent in W.

Theorem: If V has a basis with n elements, then every basis of V has n elements.

Definition: If V is spanned by a finite set of elements, then V is finite-dimensional, and the **dimension of** V, denoted by

$$\dim V \tag{9}$$

is the number of vectors in a basis for V.

Remarks:

1. If $V = \left\{\vec{0}\right\}$, then $\dim V$ is defined to be zero.
2. If V is not spanned by a finite set, then V is infinite-dimensional.

Examples:

1. $\dim \mathbb{R}^n = n$, because one basis of $\mathbb{R}^n$ is $\mathcal{E} = \{\vec{e}_1, \cdots, \vec{e}_n\}$, which has n elements.
2. $\dim \mathbb{P}_2 = 3$, where $\mathbb{P}_2$ is the set of all polynomials of degree less than or equal to two, and the standard basis of that is $\{1, t, t^2\}$.
3. $\dim \mathbb{P}_n = n + 1$
4. The symbol $\mathbb{P}$ represents the set of all polynomials, and it has a basis $\{1, t, t^2, t^3, \cdots\}$. Since there are endless elements, $\mathbb{P}$ is infinite-dimensional.
5. The set $\{(x_1, x_2, \cdots) \colon x_i \in \mathbb{R}\}$ is a vector space which is infinite-dimensional.

Example problem
Find $\dim H$, where

$$H = \left\{ \begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} : a, b, c, d \in \mathbb{R} \right\}$$

H is a subspace of $\mathbb{R}^4$. We can rewrite the vectors by separating each of the free variables.

$$\begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} = a \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

So $H = \text{Span}\left\{ \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix} \right\}$. Call these vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$. Notice that $\vec{v}_3 = -2\vec{v}_2$, so we throw out $\vec{v}_3$. The remaining set $\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$ is a linearly independent set, (which one can check by putting them in the columns of a matrix and seeing if it is invertible). So $\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$ is a basis for H, and

$$\dim H = 3$$

Dimension of a subspace:
We look at examples, and classify subspaces of $\mathbb{R}^3$ by their dimension.
dimension 0: Origin $\left\{\vec{0}\right\}$
dimension 1: Lines through the origin (which are essentially copies of $\mathbb{R}$ because we can assign each vector in one of these lines to a real number). $\text{Span}\{\vec{v}\} = \{k\vec{v} \colon k \in \mathbb{R}\}$

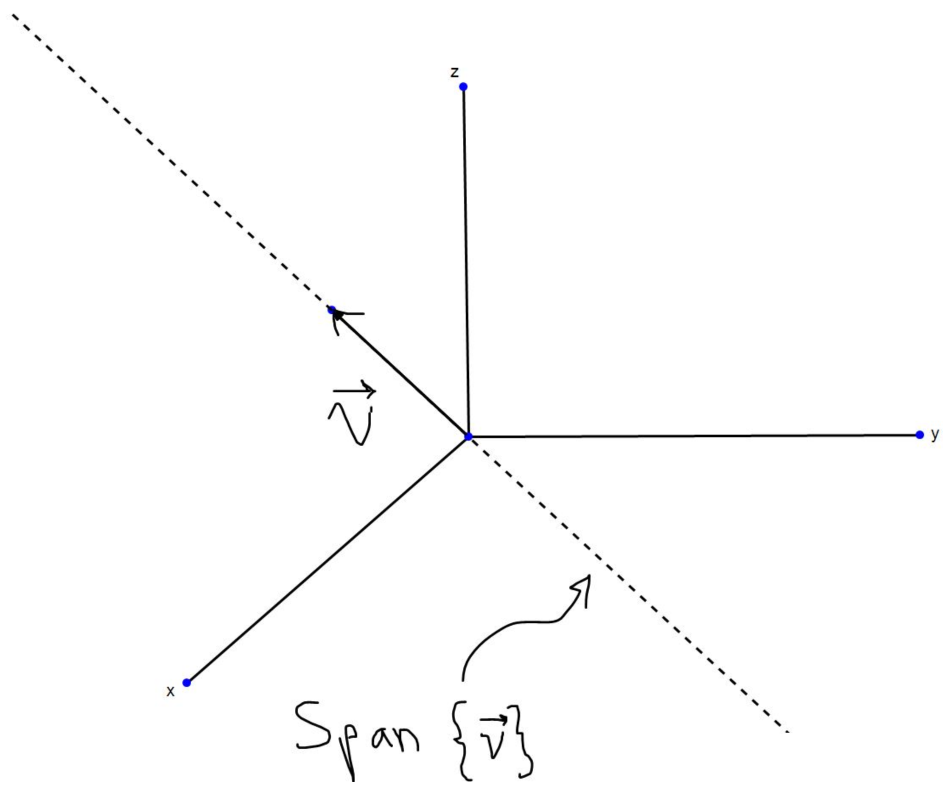


Figure 1: Span of a vector

dimension 2: Planes through the origin (which are essentially copies of $\mathbb{R}^2$). $\text{Span}\{\vec{u}, \vec{v}\} = \{c_1\vec{u} + c_2\vec{v} \colon c_1, c_2 \in \mathbb{R}\}$.

dimension 3: $\mathbb{R}^3$.

Lecture 19

April 4, 2022

Theorem: If H is a subspace of a finite-dimensional vector space V, then

1. Any linearly independent set in H can be expanded to a basis for H.
2. If H is finite-dimensional, then

$$\dim H \leq \dim V \tag{1}$$

Theorem: If V is a p-dimensional vector space with $p \geq 1$, then any linearly independent set in V with exactly p elements is a basis for V. For example,

$$\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \end{bmatrix} \right\}$$

is a linearly independent set in $\mathbb{R}^2$ and $\dim \mathbb{R}^2 = 2$. So by the theroem, $\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^2$.

Nul A and Col A:

1. dim Nul A is the number of free variables in the equation $A\vec{x} = \vec{0}$. Equivalently, it is the number of non-pivot columns in A.
2. dim Col A is the number of pivot columns in A.

Definitions:
The **rank** of a matrix A is dim Col A, and we write it as **rank** A. The **nullity** of a matrix A is dim Nul A.

Rank-Nullity Theorem:
If A is an $n \times n$ matrix, then

$$\text{rank } A + \dim \text{ Nul } A = n \tag{2}$$

because rank A is the number of pivot columns in A and dim Nul A is the number of non-pivot columns in A, so their sum must be the total number of columns in A. Remember that rank A = dim Col A.

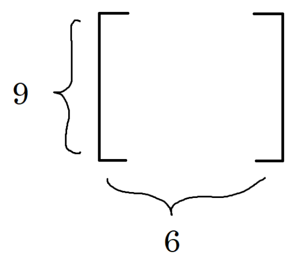
Example: Can a 6×9 matrix A have nullity 2? If the nullity is 2, then since $n = 9$, by the Rank-Nullity Theorem: rank $A + 2 = 9$, so

$$\text{rank } A = 9 - 2 = 7$$

But Col$A$ is a subspace of $\mathbb{R}^6$ which means dim Col$A \leq 6$. So rank$A \leq 6$, and rank$A$ cannot be 7. So the nullity cannot be 2.

Another example: What is the rank of a 7×9 matrix with Nullity 2? Use the Rank-Nullity Theorem: rank + Nullity = 9, and then since Nullity=2, rank= $9 - 2 = 7$.

Another example: Let A be a 9×6 matrix. The maximum number of pivots is 6, but it can be fewer. If there are fewer, then dim Nul$A > 0$.



Since there are 9 rows, Col$A$ is a subspace of $\mathbb{R}^9$, but then also since we know that the max number of pivot columns is 6,

$$0 \leq \dim \text{ Col}A \leq 6 \tag{3}$$

Theorem: If A is an $n \times n$ matrix. The following are equivalent to A begin invertible

1. Columns of A form a basis for $\mathbb{R}^n$.

2. Col$A = \mathbb{R}^n$

3. dim Col$A$ = n

4. rank $A = n$

5. Nul$A = \left\{\vec{0}\right\}$

6. dim Nul$A = 0$

7. Nullity of $A = 0$

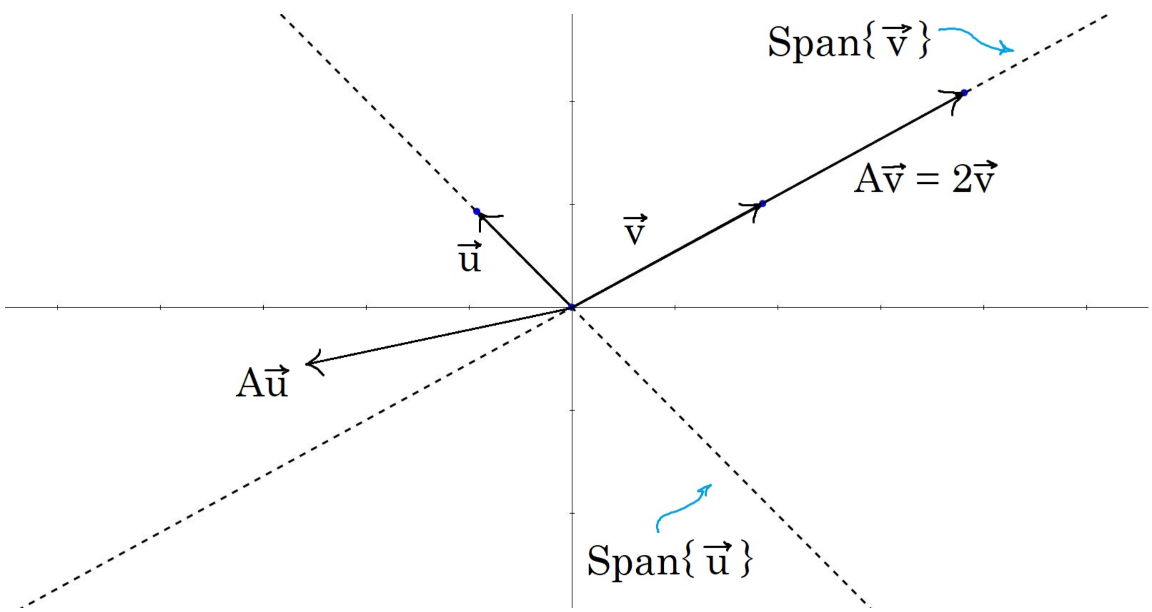
Example

If $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $\vec{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, then find $A\vec{u}$ and $A\vec{v}$.

$$A\vec{u} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$$

$$A\vec{v} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

Notice for the two vectors that $A\vec{v} = 2\vec{v}$, but there exists no k for which $A\vec{u} = k\vec{u}$. That is, $A\vec{u} \neq k\vec{u}$ for all k.



This graph shows that $A\vec{v}$ lies along the span of $\vec{v}$, whereas $A\vec{u}$ does not fall on the span of $\vec{u}$, so you could never get it as a multiple of $\vec{u}$.

Definition: An **eigenvector** of an $n \times n$ matrix A is a non-zero vector $\vec{v}$ such that

$$A\vec{v} = \lambda\vec{v} \tag{4}$$

for some scalar λ. A scalar λ is an **eigenvalue** of A if there is a non-trivial solution to

$$A\vec{x} = \lambda\vec{x} \tag{5}$$

Note that eigenvectors are never zero vectors, but eigenvalues can be zero. We will regularly use the notation **e-vectors** for eigenvectors and **e-values** for eigenvalues.

Example:
Let $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\vec{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$, and $\vec{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. Are $\vec{u}$ and $\vec{v}$ e-vectors of A? For this, we want to see if the matrix-vector product gives a multiple of each vector.

$$A\vec{u} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4\vec{u}$$

So yes, $\vec{u}$ is an e-vector of A.

$$A\vec{v} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda\vec{v}$$

for any λ, so no $\vec{v}$ is not an e-vector of A.

Another example: Show that 7 is an e-value of A from the example above. For this, we need to find a solution to

$$A\vec{x} = 7\vec{x}$$

We can rewrite this equation as

$$A\vec{x} - 7\vec{x} = \vec{0}$$

$$(A - 7I)\vec{x} = \vec{0}$$

where we use I because the 7 term is begin subtracted from another matrix, and $\vec{x}I = \vec{x}$ so it is the same. Then,

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix}$$

$$= \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}$$

This means that the columns of $A - 7I$ are linearly dependent, so $(A - 7I)\vec{x} = \vec{0}$ has a non-trivial solution, which implies that 7 is an e-value for A. To find a corresponding e-vector, augment the matrix with the zero vector.

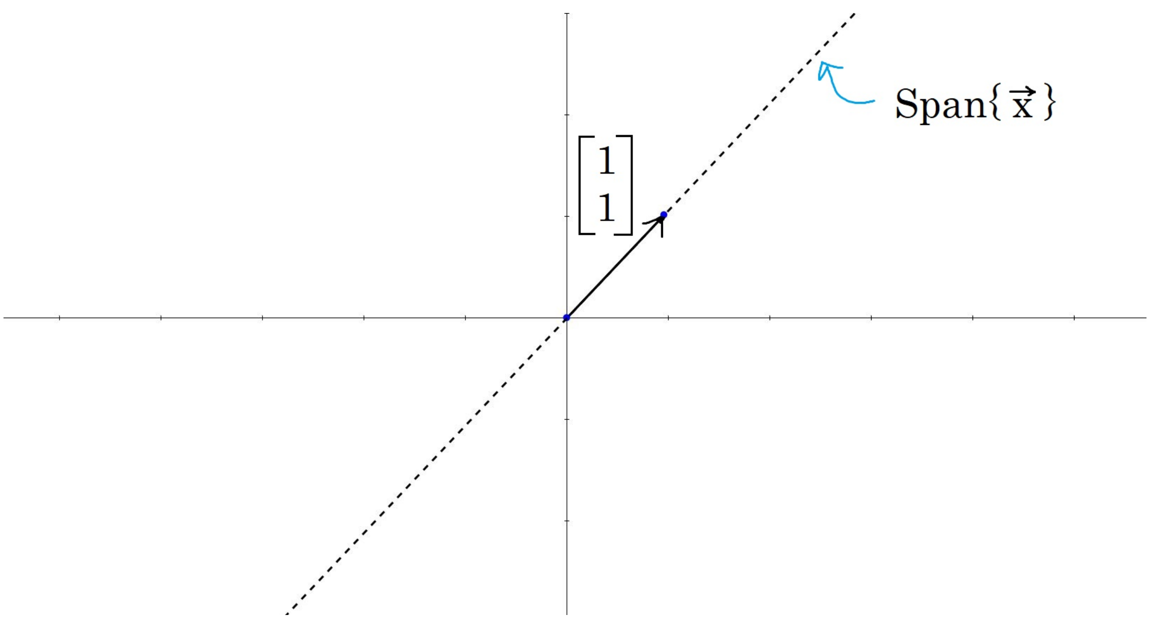
$$\left[\begin{array}{cc:c} -6 & 6 & 0 \\ 5 & -5 & 0 \end{array}\right] \to \text{ row reduce } \to \left[\begin{array}{cc:c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

So x_2 is free, and

$$x_1 - x_2 = 0$$

$$x_1 = x_2$$

which implies that $\vec{x} = \begin{bmatrix} x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ for $x_2 \in \mathbb{R}$. So any vector of this form would be an eigenvector for an e-value of 7.



Lecture 20

April 6, 2022

More on Eigenvalues and Eigenvectors

Example from last time: If $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$ then 7 is an eigenvalue of A, i.e. $A\vec{x} = 7\vec{x}$ has a non-trivial solution. To show this, move $7\vec{x}$ to the left side, and factor out $\vec{x}$:

$$A\vec{x} - 7\vec{x} = \vec{0}$$

$$(A - 7I)\vec{x} = \vec{0}$$

This is essentially the same as solving something like $A\vec{x} = \vec{0}$, except that now we're just working with a different matrix.

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix}$$

$$A - 7I = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}$$

since the columns of this matrix are linearly dependent, there will be non-trivial solutions to

$$(A - 7I)\vec{x} = \vec{0}$$

So 7 is an eigenvalue of A. We want to find all the vectors that correspond this eigenvalue. To find an eigenvector, augment the matrix $A - 7I$ with the zero vector:

$$\left[\begin{array}{cc:c} -6 & 6 & 0 \\ 5 & -5 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{cc:c} -1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

The first row implies $x_1 = x_2$, where x_2 is free. So $\vec{x} = \begin{bmatrix} x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ for any $x_2 \in \mathbb{R}$. This means the span of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is the solution set of our equation and any vector of the form $\vec{v} = \begin{bmatrix} k \\ k \end{bmatrix}$, where $k \neq 0$, is an e-vector associated with the e-value 7 of A.

To find other e-values of A, replace 7 with λ in $(A - 7I)\vec{x} = \vec{0}$, and apply the same method:

1. A scalar λ is an e-value of A if and only if $(A - \lambda I)\vec{x} = \vec{0}$ has a non-trivial solution.

2. The set of all solutions to $(A - \lambda I)\vec{x} = \vec{0}$ is the null space of $A - \lambda I$. This is called the **eigenspace**, (the set of all eigenvectors) of A corresponding to λ. (This is a subspace of $\mathbb{R}^n$ if A is $n \times n$).

Notation: We call an eigenspace e-space. If A is fixed, write E_λ for the e-space of A corresponding to λ, i.e.

$$E_\lambda = \left\{ \vec{x} \in \mathbb{R}^n \colon A\vec{x} = \lambda \vec{x} \right\} \tag{1}$$

where A is $n \times n$.

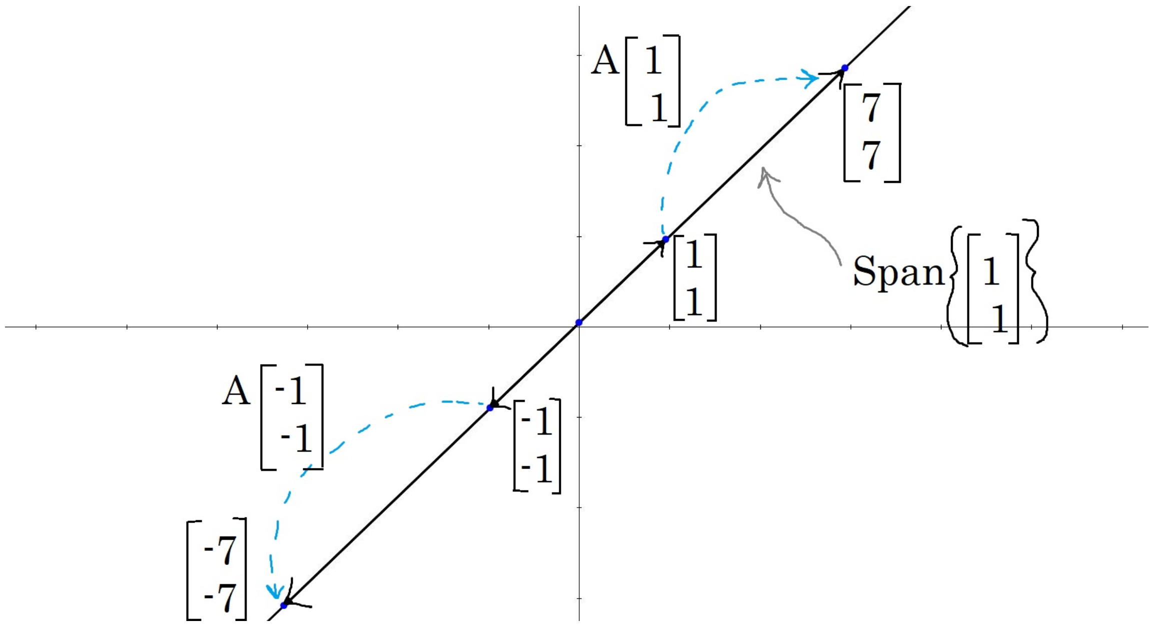
Example: In the previous example,

$$E_7 = \left\{ \begin{bmatrix} k \\ k \end{bmatrix} : k \in \mathbb{R} \right\}$$

$$= \left\{ x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} : x_2 \in \mathbb{R} \right\}$$

$$= \text{Span} \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

Graphically, the e-space corresponding to A is a line, and multiplying A by any vector in that line will give 7 times that vector.



Example:

$$A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & 1 & 8 \end{bmatrix}$$

An e-value of A is $\lambda = 2$. Find a basis for E_2.

$$E_2 = \mathrm{Nul}(A - 2I)$$

$$A - 2I = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}$$

$$\to \text{row reduce} \to \begin{bmatrix} \mathbf{2} & -1 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

x_2 and x_3 are free, and

$$2x_1 - x_2 + 6x_3 = 0$$

So

$$x_1 = \frac{1}{2}x_2 - 3x_3$$

which means that

$$E_2 = \mathrm{Nul}(A - 2I) = \left\{ \begin{bmatrix} \frac{1}{2}x_2 - 3x_3 \\ x_2 \\ x_3 \end{bmatrix} : x_2, x_3 \in \mathbb{R} \right\}$$

We want to find a basis for E_2, so we need to find a spanning set for this. Factor out the free variables:

$$E_2 = \left\{ x_2 \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} : x_2, x_3 \in \mathbb{R} \right\}$$

$$= \mathrm{Span} \left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

These two vectors are linearly independent, so they form a basis for E_2.

$$\text{Basis for } E_2 = \left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Since there are two vecors in this basis,

$$\dim E_2 = 2$$

Theorem: If $\vec{v}_1, \dots, \vec{v}_r$ are e-vectors of a matrix A corresponding to distinct e-values $\lambda_1, \lambda_2, \dots, \lambda_r$, respectively, then $\{\vec{v}_1, \dots, \vec{v}_r\}$ is linearly independent.

How to find e-values of a matrix:

Example: Find all e-values of $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$. By definition, λ is an e-value of A if and only if $(A - \lambda I)\vec{x} = \vec{0}$ has a non-zero solution, which is true if and only if $A - \lambda I$ is not invertible. The matrix $A - \lambda I$ is not invertible if and only if $\det(A - \lambda I) = 0$. So we just need to find all values of λ for which this is true.

$$A - \lambda I = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

$$= \begin{bmatrix} 2 - \lambda & 3 \\ 3 & -6 - \lambda \end{bmatrix}$$

So

$$\begin{aligned}\det(A - \lambda I) &= (2-\lambda)\cdot(-6-\lambda) - 3\cdot 3 \\ &= -12 - 2\lambda + 6\lambda + \lambda^2 - 9 \\ &= \lambda^2 + 4\lambda - 21\end{aligned}$$

Now we just need to find the roots of this polynomial to solve $\det(A - \lambda I) = \lambda^2 + 4\lambda - 21 = 0$. We can re-write it as

$$\lambda^2 + 4\lambda - 21 = (\lambda + 7)(\lambda - 3)$$

So, the eignvalues of A are $\lambda_1 = -7$ and $\lambda_2 = 3$. These are all the e-values of A.

Summary of steps:
We started with the equation $(A - \lambda I)\vec{x} = \vec{0}$ and then used the fact that this has non-trivial solutions to say that $A - \lambda I$ is not invertible, which is equivalent to saying $\det(A - \lambda I) = 0$. So we took the determinant of that matrix, and expressed it as a polynomial in terms of λ. Then we set the polynomial equal to zero to find its roots and see for which values of λ the determinant is zero. The set of roots is the set of all e-values.

Lecture 21

April 25, 2022

Finding Eigenvalues

For the equation $A\vec{x} = \lambda\vec{x}$, λ is an e-value and $\vec{x}$ is an e-vector. We have looked at the set of all eigenvectors associated with the e-value λ: $E_\lambda = \{\vec{x}\colon A\vec{x} = \lambda\vec{x}\}$. We can re-write this by taking the following steps

$$A\vec{x} = \lambda\vec{x} \iff A\vec{x} - \lambda\vec{x} = \vec{0}$$

$$\iff A\vec{x} - (\lambda I)\vec{x} = \vec{0}$$

$$\iff (A - \lambda I)\vec{x} = \vec{0}$$

$$\iff \vec{x} \in \mathrm{Nul}(A - \lambda I)$$

So an equivalent way of writing the eigenspace is as

$$E_\lambda = \mathrm{Nul}(A - \lambda I) \tag{1}$$

If you want to find E_λ, you just need to find the null space of $A - \lambda I$.

If λ is an e-value for A, then $E_\lambda \neq \left\{\vec{0}\right\}$, which is the same thing as saying there is a non-trivial solution to $(A - \lambda I)\vec{x} = \vec{0}$. This is equivalent to saying the columns of $A - \lambda I$ are linearly dependent, which is true if and only if $\det(A - \lambda I) = 0$. So, λ is an e-value of A if and only if $\det(A - \lambda I) = 0$. And in order to find e-values of A, it suffices to find the solutions to the equation

$$\det(A - \lambda I) = 0 \tag{2}$$

where the variable is λ. Taking the determinant of this matrix will eventually give you an n^{th} degree polynomial in terms of λ. This is called the characteristic polynomial.

Definition: The equation $\det(A - \lambda I) = 0$ is the **characteristic polynomial** (or characteristic equation) of A. This is an n^{th} degree polynomial in λ if A is $n \times n$.

Example: Find the characteristic equation of

$$A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We need to find the determinant of $A - \lambda I$.

$$A - \lambda I = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix}$$

This matrix is upper-triangular, so the determinant is the product of the diagonal entries.

$$\det(A - \lambda I) = (5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda)$$

Set this equal to zero.

$$0 = (5-\lambda)^2(3-\lambda)(1-\lambda)$$

So $\lambda = 1, 3, 5$ where 5 has a multiplicity $(5-\lambda)$ appears twice in the characteristic polynomial.

Definition: The **algebraic multiplicity** of an e-value of λ is the multiplicity of the root of the characteristic polynomial.

Example: The characteristic polynomial of a 6×6 matrix is $\lambda^6 - 4\lambda^5 - 12\lambda^4$. Find the e-values with algebraic multilplicities. For this, set the polynomial equal to zero:

$$\lambda^4(\lambda^2 - 4\lambda - 12) = 0$$

So either $\lambda = 0$, or

$$\lambda^2 - 4\lambda - 12 = 0$$

$$(\lambda - 6)(\lambda + 2) = 0$$

$$\lambda = 6, -2$$

So the e-values are

$$\begin{cases} \lambda = 0 & \text{multiplicity of } 4 \\ \lambda = 6 & \text{multiplicity of } 1 \\ \lambda = -2 & \text{multiplicity of } 1 \end{cases}$$

Note that sometimes these roots are complex. If you count multiplicity, for any polynomial of degree n you will get n roots, though some may not be real. For example, $\lambda^2 + 1 = 0$ has the roots $\lambda = \pm i$.

Similarity: If A and B are $n \times n$ matrices, then A is **similar** to B if there is an invertible matrix P such that

$$A = PBP^{-1} \tag{3}$$

Or equivalently,

$$B = P^{-1}AP$$

Theorem: If A and B are similar, then they have the same characteristic polynomial and so the same e-values, (with the same multiplicity because the polynomials are the same). Note:

1. Not every pair of matrices with the same e-values are similar, (they also need to have the same e-vectors).
2. Similarity is different from row equivalency, because row operations change the e-values. (If $A \sim B$, then $A = EB$ for some invertible matrix E, and E will change the e-value).

Why is the theorem true?

Proof:

If $B = P^{-1}AP$, we want to show that this implies $\det(B - \lambda I) = \det(A - \lambda I)$, so that A and B have the same characteristic polynomial.

$$B = P^{-1}AP$$

$$\begin{aligned} B - \lambda I &= P^{-1}AP - \lambda I \\ &= P^{-1}AP - \lambda(P^{-1}P) \\ &= P^{-1}AP - P^{-1}\lambda P \\ &= P^{-1}(A - \lambda I)P \end{aligned}$$

So $\det(B-\lambda I) = \det(P^{-1}(A-\lambda I)P)$ Then we use the determinant property $\det(AB) = \det(A)\cdot \det(B)$, so that

$$\det(B-\lambda I) = \det(P^{-1}) \cdot \det(A-\lambda I) \cdot \det(P)$$

$$= \frac{1}{\det(P)} \det(P) \cdot \det(A-\lambda I)$$

$$\det(B-\lambda I) = \det(A-\lambda I)$$

So if it is given that $B = P^{-1}AP$, then A and B have the same characteristic polynomials. This motivates the idea that some matrices are diagonalizable—or able to be be made similar to a diagonal matrix—which is useful because diagonal matrices have easy determinants. For example, if $D = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$ then

$$\det(D-\lambda I) = \det\left(\begin{bmatrix} 3-\lambda & 0 \\ 0 & 5-\lambda \end{bmatrix}\right)$$

$$= (3-\lambda)(5-\lambda)$$

So $\lambda = 3, 5$ are the e-values. The goal of diagonalization is to see if we can do this.

Diagonalization:

Let $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$. Find D^k.

$$k = 2: \quad D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}$$

So in general,

$$D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \quad \text{for any } k.$$

Example: Find A^k if $A = PDP^{-1}$ where $P = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$ and $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$, (i.e. A is similar to D).

$$k = 2: \quad A^2 = (PDP^{-1})(PDP^{-1})$$

$$= (PD)(PP^{-1})(DP^{-1})$$
$$= PD^2P^{-1}$$

So in general,

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & -1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{bmatrix}$$

we were able to use the fact that diagonal matrices are simpler, to find A^k. (more on diagonalization next lecture)