Linear Algebra: Supplementary Notes

Ethan Nadler

These notes were created as supplementary material for an introductory linear algebra course at The University of California, Santa Barbara. They rely heavily on David C. Lay's *Linear Algebra and its Application*; many of the exercises and definitions come directly from this text. However, they also include several original exercises, and some of the derivations differ from Lay's presentation in order to maintain a coherent narrative within each lesson.

I aimed to create a set of notes that provides a challenge for students and emphasizes some of the more "important" concepts in the subject without spending too much time developing a theoretical framework. Many of the lessons assume some background knowledge, hopefully provided by the course lecture. The material is not meant to provide a comprehensive overview of linear algebra or even of Lay's text; while the sequence of topics usually follows the text, the order was determined by the material presented in lecture.

Week 1

A **system of linear equations** is either **consistent** (if it has one solution or infinitely many solutions) or **inconsistent** (if it has no solutions). Note that if a system is consistent, there's no "in-between" case; a consistent system has either one solution or infinitely many! Try to convince yourself of this by picturing the solution to a system of two equations as the intersection of two lines in the x_1, x_2 plane. (What does this look like in 3D? Does the argument still work?)

The **coefficient matrix** of a system of n linear equations,

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = c_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = c_2 \\ . \\ . \\ a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n = c_n \end{gathered} \tag{1}$$

is the $n \times n$ matrix

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \tag{2}$$

and the corresponding **augmented matrix** includes the constant terms $c_1, \ldots, c_n$ on the right hand side of the matrix.

To solve a system of equations, use the following **elementary row operations** on the augmented matrix:

1. Replace a row by the sum of itself and a multiple of any other row.
2. Switch any two rows.
3. Multiply any row by a nonzero constant.

Below is one example of a system to solve using the row operations. **Important:** practice solving systems using row operations until you're completely comfortable with the process.

Question: What does the row reduced augmented matrix look like when a system has one solution, infinitely many solutions, or no solution?

Ex. Solve the system using elementary row operations.

$$\begin{aligned} x_1 - 3x_3 &= 8 \\ 2x_1 + 2x_2 + 9x_3 &= 7 \\ x_2 + 5x_3 &= -2 \end{aligned} \tag{3}$$

Answer: $x_1 = 5$, $x_2 = 3$, $x_3 = -1$.

Ex. For the following augmented matrices, determine the value(s) of h that make the corresponding systems consistent. Justify your answers!

$$\begin{bmatrix} 1 & h & 4 \\ 3 & 6 & 8 \end{bmatrix} \tag{4}$$

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

Answer: 1. All $h \neq 2$. *2. All* h.

Ex. Suppose the following system is consistent for all values of f and g. What can you conclude about the coefficients c and d? (Hint: think about f and g as completely unrelated in general.)

$$\begin{aligned} 2x_1 + 4x_2 &= f \\ cx_1 + dx_2 &= g \end{aligned} \tag{6}$$

Answer: $d \neq 2c$. *(Strictly speaking, we must have* $d \neq 2c$ *whenever* $g \neq cf$*; what happens if* $g = cf$*?)*

Week 2

A **linear combination** of a set of vectors $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is any vector $\mathbf{y} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n}$, where $c_1, \ldots, c_n$ are constants. Note that the vectors $\mathbf{v_j}$ should have the same size and that some of the c_i may be 0. The idea of representing an object (such as a vector) as a linear combination of other objects of the same kind is an extremely important concept! The set of *all* linear combinations of a set of vectors $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is called the **span** of the vectors, which we denote by $\mathrm{Span}\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$.

Let's think about span geometrically: in two dimensions, it's clear that the vectors $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ span the entire x, y plane since we can write any vector $\mathbf{v}$ by decomposing it into $\mathbf{x}$ and $\mathbf{y}$ components in the usual way:

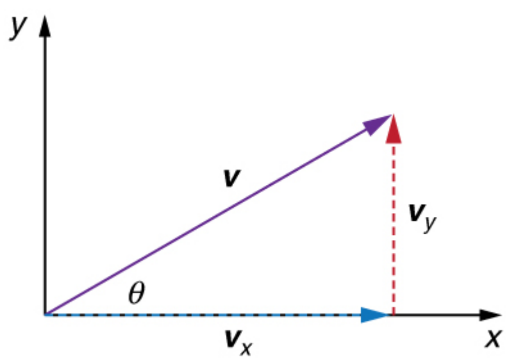

Fig. 1.— Components of a vector in Cartesian coordinates: $\mathbf{v} = \mathbf{v_x} + \mathbf{v_y} = v_x\mathbf{x} + v_y\mathbf{y}$, where $v_x = \cos\theta\ v$ and $v_y = \sin\theta\ v$.

Question: What if we use different vectors $\mathbf{x}'$ *and* $\mathbf{y}'$ *rather than* $\mathbf{x}$ *and* $\mathbf{y}$ *to describe vectors in the plane? Do any two vectors span the plane? Give an example of two vectors that do* ***not*** *span the plane.*

Ex. For what value(s) of h is $\mathbf{b}$ in the plane spanned by $\mathbf{a_1}$ and $\mathbf{a_2}$, where

$$\mathbf{a_1} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}, \mathbf{a_2} = \begin{bmatrix} -5 \\ -8 \\ 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ -5 \\ h \end{bmatrix}? \tag{7}$$

Answer: $h = 3$.

Ex. Construct a 3×3 matrix $\mathbf{A}$ (with $\mathbf{A} \neq \mathbf{0}$) and a vector b in $\mathbb{R}^3$ such that b is *not* in the set spanned by the columns of $\mathbf{A}$. Then repeat the exercise, but construct $\mathbf{A}$ such that *none* of its entries are 0.

Answer: There are an infinite number of possibilities!

Consider matrix equations of the form $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is an $n \times n$ matrix and $\mathbf{x}$ and $\mathbf{b}$ are $n \times 1$ column vectors. Try to convince yourself of the following statement: the equation $\mathbf{Ax} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combination of the columns of $\mathbf{A}$.

Here are two very important properties of any $m \times n$ matrix $\mathbf{A}$. For any vectors $\mathbf{u}$ and $\mathbf{v}$ (of the appropriate size) and any constant c,

1. $\mathbf{A}(\mathbf{u} + \mathbf{v}) = \mathbf{Au} + \mathbf{Av}$.
2. $\mathbf{A}(c\mathbf{u}) = c(\mathbf{Au})$.

In other words, if $\mathbf{A}$ multiplies a linear combination of vectors $c_1\mathbf{v_1} + \cdots + c_n\mathbf{v_n}$, we can multiply each vector by $\mathbf{A}$, multiply by the corresponding coefficient, and add the results together: $\mathbf{A}(c_1\mathbf{v_1} + \cdots + c_n\mathbf{v_n}) = c_1(\mathbf{Av_1}) + \cdots + c_n(\mathbf{Av_n})$. This property is called **linearity**.

Ex. Let

$$\mathbf{A} = \begin{bmatrix} 3 & -1 \\ -9 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{8}$$

Show that the equation $\mathbf{Ax} = \mathbf{b}$ does not have a solution for all possible $\mathbf{b}$. For what set of $\mathbf{b}$ is there a solution?

Answer: A solution exists for all $\mathbf{b}$ *with* $b_2 = -3b_1$.

Ex. Let $\mathbf{A}$ be a 3×2 matrix. Explain why the equation $\mathbf{Ax} = \mathbf{b}$ cannot be consistent for all $\mathbf{b}$ in $\mathbb{R}^3$. Can you generalize your argument for an arbitrary $\mathbf{A}$ with n columns and $m > n$ rows?

Hint: Consider $\mathbf{Ax} = \mathbf{b}$ *explicitly for the* 3×2 *case:*

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{9}$$

How many equations does this represent, in how many unknowns?

Week 3

Ex. Determine if the following set of vectors spans $\mathbb{R}^3$:

$$\left\{ \begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}2\\2\\2\end{bmatrix}, \begin{bmatrix}2\\4\\6\end{bmatrix} \right\}. \tag{10}$$

Answer: The set does ***not*** *span $\mathbb{R}^3$. Things to think about: What is the minimum number of vectors needed to span $\mathbb{R}^3$? Why? Is there a "quick" way to tell whether a given set of vectors spans $\mathbb{R}^n$? (Revisit this question after we've covered linear independence).*

A **homogeneous** system of linear equations has the form $\mathbf{Ax} = \mathbf{0}$. Convince yourself of the following statement: $\mathbf{Ax} = \mathbf{0}$ has a nontrivial solution (i.e., $\mathbf{x} \neq \mathbf{0}$) if and only if the equation has at least one free variable. Homogeneous systems are very important because *any linear combination of solutions to a homogeneous system is still a solution*! That is, if we have solutions $\mathbf{u}$ and $\mathbf{v}$ of $\mathbf{Ax} = \mathbf{0}$, meaning $\mathbf{Au} = \mathbf{0}$ and $\mathbf{Av} = \mathbf{0}$, then

$$\mathbf{A}(c\mathbf{u} + d\mathbf{v}) = \mathbf{A}(c\mathbf{u}) + \mathbf{A}(d\mathbf{v}) = c(\mathbf{Au}) + d(\mathbf{Av}) = \mathbf{0} + \mathbf{0} = \mathbf{0}, \tag{11}$$

and the linear combination $c\mathbf{u} + d\mathbf{v}$ is still a solution of $\mathbf{Ax} = \mathbf{0}$.

A **nonhomogeneous** system of linear equations has the form $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{b} \neq \mathbf{0}$. We will now see the importance of homogeneous solutions. Suppose that $\mathbf{p}$ is a solution of $\mathbf{Ax} = \mathbf{b}$, and suppose that $\mathbf{v_h}$ is a solution of the corresponding homogeneous system $\mathbf{Ax} = \mathbf{0}$. Then the *general solution* to $\mathbf{Ax} = \mathbf{b}$ is $\mathbf{p} + \mathbf{v_h}$, since

$$\mathbf{A}(\mathbf{p} + \mathbf{v_h}) = \mathbf{A}(\mathbf{p}) + \mathbf{A}(\mathbf{v_h}) = \mathbf{b} + \mathbf{0} = \mathbf{b}. \tag{12}$$

In practice, this form of a solution will automatically result from row reduction. For example, consider the system $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix}3 & 5 & -4\\-3 & -2 & 4\\6 & 1 & -8\end{bmatrix}, \mathbf{b} = \begin{bmatrix}7\\-1\\-4\end{bmatrix}. \tag{13}$$

The augmented matrix row reduces to

$$\left(\begin{array}{ccc|c} 1 & 0 & -4/3 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right), \tag{14}$$

giving $x_2 = 2$ and $x_1 = -1 + \frac{4}{3}x_3$, with x_3 free. A solution to the system is then

$$\mathbf{x} = \begin{bmatrix} -1 + \frac{4}{3}x_3 \\ 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} \equiv \mathbf{p} + x_3\mathbf{v_h}. \tag{15}$$

The "particular" and homogeneous solutions both come out of the row reduction process. The homogeneous solution is always multiplied by an arbitrary **parameter** (here called x_3).

Question: What does the row of zeros in Eq. 5 tell us about the solutions of the corresponding system – is there a homogeneous solution? What if there were no row of zeros? (Always look out for rows of zeros!)

Ex. True or False:

(a) A homogeneous equation is always consistent.

(b) If $\mathbf{x}$ is a nontrivial solution of $\mathbf{Ax} = \mathbf{0}$, then every entry in $\mathbf{x}$ is nonzero.

(c) The equation $\mathbf{Ax} = \mathbf{b}$ is homogeneous if the zero vector is a solution.

(d) If $\mathbf{Ax} = \mathbf{b}$ is consistent, then the solution set is obtained by translating the solution set of $\mathbf{Ax} = \mathbf{0}$.

Answer: T, F, T, T.

A set of vectors $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is **linearly independent** if the equation

$$x_1\mathbf{v_1} + x_2\mathbf{v_2} + \cdots + x_n\mathbf{v_n} = \mathbf{0} \tag{16}$$

only has the trivial solution $\mathbf{x} = \mathbf{0}$. On the other hand, the set is **linearly dependent** if there exist $x_1, \ldots, x_n$, not all zero, such that $x_1\mathbf{v_1} + \cdots + x_n\mathbf{v_n} = \mathbf{0}$. This definition immediately implies that *the columns of a matrix* $\mathbf{A}$ *are linearly independent if and only if the equation* $\mathbf{Ax} = \mathbf{0}$ *has only the trivial solution.*

Question: For a set consisting of one vector, $\mathbf{v}$*, under what conditions is the set linearly independent and under what conditions is it linearly dependent? What if the set has two vectors,* $\mathbf{v_1}$ *and* $\mathbf{v_2}$*?*

Here's an important theorem: **A set of two or more vectors is linearly dependent if and only if at least one of the vectors is a linear combination of the others**. Let's get some intuition for this statement; consider the set $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$, and suppose that one of the vectors – say, $\mathbf{v_2}$ – is a linear combination of the others. Then there exist c_1, c_3 such that $\mathbf{v_2} = c_1\mathbf{v_1} + c_3\mathbf{v_3}$, and it's easy to find a nontrivial solution of equation (7); just take $x_1 = -c_1$, $x_2 = 1$, and $x_3 = -c_3$, for example:

$$x_1\mathbf{v_1} + x_2\mathbf{v_2} + x_3\mathbf{v_3} = -c_1\mathbf{v_1} + \mathbf{v_2} - c_3\mathbf{v_3} = \mathbf{v_2} - \mathbf{v_2} = \mathbf{0}. \tag{17}$$

Ex. Determine if each set of vectors is linearly independent or linearly dependent:

(a) $\left\{ \begin{bmatrix}1\\7\\6\end{bmatrix}, \begin{bmatrix}2\\0\\9\end{bmatrix}, \begin{bmatrix}3\\1\\5\end{bmatrix}, \begin{bmatrix}4\\1\\8\end{bmatrix} \right\}$

(b) $\left\{ \begin{bmatrix}2\\3\\5\end{bmatrix}, \begin{bmatrix}0\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\8\end{bmatrix} \right\}$

(c) $\left\{ \begin{bmatrix}-2\\4\\6\\10\end{bmatrix}, \begin{bmatrix}3\\-6\\-9\\15\end{bmatrix} \right\}.$

Answer: Dependent, dependent, independent.

Ex. Find the value(s) of h for which the vectors are linearly dependent:

$$\begin{bmatrix}1\\5\\-3\end{bmatrix}, \begin{bmatrix}-2\\-9\\6\end{bmatrix}, \begin{bmatrix}3\\h\\-9\end{bmatrix}. \tag{18}$$

Answer: All h. (You could row reduce, but there's a shortcut!)

Linear Transformations: A **transformation** $T : \mathbb{R}^n \to \mathbb{R}^m$ is a mapping that assigns each $\mathbf{x}$ in $\mathbb{R}^n$ a vector $T(\mathbf{x})$ in $\mathbb{R}^m$. $\mathbb{R}^n$ is called the **domain** of T, $\mathbb{R}^m$ is called the **codomain** of T, and the set of all **images** $T(\mathbf{x})$ is called the **range** of T. For a **linear** transformation,

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

2. $\mathbf{T}(c\mathbf{u}) = c(\mathbf{T}(\mathbf{u}))$

for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ and any constant c. Again, as we saw with matrices, if a linear transformation acts on a linear combination of vectors, the result is the sum of the linear transformation acting on each vector multiplied by the corresponding coefficient: $\mathbf{T}(c_1\mathbf{v_1} + \cdots + c_n\mathbf{v_n}) = c_1(\mathbf{T}(\mathbf{v_1})) + \cdots + c_n(\mathbf{T}(\mathbf{v_n}))$.

Any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ can be expressed as an $m \times n$ matrix! There are several nice examples with illustrations in the textbook, including a projection transformation and a shear transformation.

Question: Let $\mathbf{A}$ be a $j \times k$ matrix. What must a and b be in order to define $T : \mathbb{R}^a \to \mathbb{R}^b$ by $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$?

Ex. Let

$$\mathbf{A} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}, \tag{19}$$

and define a transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$.

(a) Find the image of $\mathbf{u}$ under the transformation T.

(b) Find an $\mathbf{x}$ in $\mathbb{R}^2$ whose image under T is $\mathbf{b}$.

(c) Is there more than one $\mathbf{x}$ whose image under T is $\mathbf{b}$?

(d) Is $\mathbf{c}$ in the range of the transformation T?

Answer: (a) $\begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$, *(b)* $\begin{bmatrix} 3/2 \\ 1/2 \end{bmatrix}$, *(c) no, (d) no.*

How do we actually find the matrix that corresponds to a given linear transformation? The answer is actually very simple: if $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then the columns of the matrix $\mathbf{A}$ that represents this transformation are given by the action of T on the columns of the $n \times n$ identity matrix I_n:

$$\mathbf{A} = \big[T(\mathbf{e_1}) \ldots T(\mathbf{e_n})\big], \tag{20}$$

where $\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e_n} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$ are **standard basis vectors** in $\mathbb{R}^n$. This works because any vector $\mathbf{x}$ can be decomposed into a piece along each basis vector. For example, in $\mathbb{R}^2$,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1 \mathbf{e_1} + x_2 \mathbf{e_2}, \tag{21}$$

so $T(\mathbf{x}) = x_1 T(\mathbf{e_1}) + x_2 T(\mathbf{e_2})$.

Ex. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation that maps $\mathbf{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ into $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ and maps $\mathbf{v} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ into $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$.

(a) Find $T(2\mathbf{u})$ and $T(2\mathbf{u} + 3\mathbf{v})$.

(b) Find the matrix $\mathbf{A}$ such that $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$.

Answer: (a) $\begin{bmatrix} 8 \\ 2 \end{bmatrix}, \begin{bmatrix} 11 \\ -7 \end{bmatrix}$, *(b)* $\begin{bmatrix} -8/3 & 3 \\ -5 & 4 \end{bmatrix}$.

Week 4

Matrix Operations: Let $\mathbf{A}$ and $\mathbf{B}$ be $m \times n$ matrices:

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

There are several matrix operations we can perform with these matrices:

1. Addition:

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix}. \tag{23}$$

   Matrix addition only works if $\mathbf{A}$ *and* $\mathbf{B}$ *have the same size!*

2. Scalar Multiplication:

$$c\mathbf{A} = \begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1n} \\ ca_{21} & ca_{22} & \dots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \dots & ca_{mn} \end{bmatrix}. \tag{24}$$

Now let $\mathbf{A}$ be an $m \times n$ matrix and $\mathbf{B}$ be an $n \times p$ matrix. We can then perform **matrix multiplication** to form the $m \times p$ matrix $\mathbf{C} = \mathbf{AB}$. For example, if $\mathbf{A}$ and $\mathbf{B}$ are 3×3 matrices, we have

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}. \tag{25}$$

This is a good exercise to work out on your own! The answer is

$$\mathbf{C} = \begin{bmatrix} a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31} & a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32} & a_{11}b_{13}+a_{12}b_{23}+a_{13}b_{33} \\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31} & a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32} & a_{21}b_{13}+a_{22}b_{23}+a_{23}b_{33} \\ a_{31}b_{11}+a_{32}b_{21}+a_{33}b_{31} & a_{31}b_{12}+a_{32}b_{22}+a_{33}b_{32} & a_{31}b_{13}+a_{32}b_{23}+a_{33}b_{33} \end{bmatrix}. \tag{26}$$

Warning – unlike addition or scalar multiplication, *the order matters for matrix multiplication*! In general, $\mathbf{AB} \neq \mathbf{BA}$.

There are two more useful matrix operations: we can raise a matrix to the kth power,

$$\mathbf{A}^k = \mathbf{AA}\ldots\mathbf{A} \ (k \text{ times}), \tag{27}$$

and we can take the *transpose* of a matrix by interchanging its rows and columns. For example, the transpose of the 3×3 matrix $\mathbf{A}$ from equation (4) is

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}. \tag{28}$$

Ex. Let

$$\mathbf{A} = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} -1 & 1 \\ 3 & 4 \end{bmatrix}, \ \mathbf{C} = \begin{bmatrix} -3 & -5 \\ 2 & 1 \end{bmatrix}. \tag{29}$$

Compute $\mathbf{AB}$ and $\mathbf{AC}$.

Answer: $\mathbf{AB} = \mathbf{AC} = \begin{bmatrix} -21 & -21 \\ 7 & 7 \end{bmatrix}$. *Notice that* $\mathbf{AB} = \mathbf{AC}$*, even though* $\mathbf{B} \neq \mathbf{C}$*! There is no "cancellation" allowed in equations involving matrix multiplication.*

Ex. Let

$$\mathbf{A} = \begin{bmatrix} 3 & -6 \\ -2 & 4 \end{bmatrix}. \tag{30}$$

Construct a 2×2 matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{0}$. Use two different nonzero columns for $\mathbf{B}$.

Answer: $\mathbf{B} = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$ *is one possible choice.*

An $n \times n$ matrix $\mathbf{A}$ is **invertible** if there is an $n \times n$ matrix $\mathbf{C}$ such that $\mathbf{AC} = \mathbf{CA} = \mathbf{I}_n$, where $\mathbf{I}_n$ is the $n \times n$ identity matrix. $\mathbf{C}$ is then the **inverse** of $\mathbf{A}$, denoted by $\mathbf{C} = \mathbf{A}^{-1}$. Note that not all matrices are invertible!

There's a simple formula for the inverse of a 2×2 matrix, and a simple procedure for finding the inverse of a larger matrix. For a 2×2 matrix,

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \implies \mathbf{A}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \tag{31}$$

Clearly, $\mathbf{A}^{-1}$ only exists if the **determinant** $\det \mathbf{A} = ad - bc \neq 0$. To find the inverse of a larger $n \times n$ matrix $\mathbf{B}$, row reduce the augmented matrix $\begin{bmatrix}\mathbf{B} \mid \mathbf{I}_n\end{bmatrix}$ until $\mathbf{B}$ has been reduced to $\mathbf{I}_n$; then whatever is left on the augmented side of the matrix is $\mathbf{B}^{-1}$.

If $\mathbf{A}$ is an invertible $n \times n$ matrix, then the equation $\mathbf{A}\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ for each $\mathbf{b}$ in $\mathbb{R}^n$.

Ex. Use matrix inversion to solve the system

$$\begin{aligned} 8x_1 + 6x_2 &= 2 \\ 5x_1 + 4x_2 &= -1. \end{aligned} \tag{32}$$

Answer: $\mathbf{x} = \begin{bmatrix} 7 \\ -9 \end{bmatrix}$.

Ex. Find the inverse of the matrix $\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$, if it exists.

Answer: $\mathbf{A}^{-1} = \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix}$.

Ex. Use matrix inversion to solve the system $\mathbf{A}\mathbf{x} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{33}$$

Answer: $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

Notice that if $\mathbf{A}^{-1}$ exists, then the solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$ is unique – there is only one solution! This means that any matrix with linearly dependent rows (i.e., any matrix that can be row reduced to something with a row of zeros) is *not* invertible, since the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ would then contain a free variable and have more than one solution.

Properties of Linear Transformations: A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is **onto** if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at least one $\mathbf{x}$ in $\mathbb{R}^n$. In words, T maps vectors in its domain onto *every* vector in its codomain. Or, if you like, the *range* of T is equal to its codomain.

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is **one-to-one** if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of at most one $\mathbf{x}$ in $\mathbb{R}^n$. That is, each vector in $\mathbb{R}^n$ gets mapped to a different vector in $\mathbb{R}^m$. Or, equivalently, no two vectors have the same image under T.

Be careful – a general linear transformation can have any combination of these properties. There are transformations that are one-to-one but not onto, onto but not one-to-one, both onto and one-to-one, and ones that are neither of the two!

How to check whether a linear transformation is one-to-one or onto: Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation that is represented by a matrix $\mathbf{A}$. Then

1. T is onto if and only if the columns of $\mathbf{A}$ span $\mathbb{R}^m$. *Ask yourself, "does* $\mathbf{A}$ *span* $\mathbb{R}^m$*?"*
2. T is one-to-one if and only if the equation $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution. *Ask yourself, "are the columns of* $\mathbf{A}$ *linearly independent?"*

Ex. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation that is represented by the matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix}. \tag{34}$$

Is T one-to-one? Is T onto?

Answer: T is one-to-one, but not onto. Question: Can a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$ ever be onto?

Ex. Let T be a linear transformation whose action is given by

$$T(x_1, x_2, x_3, x_4) = 3x_1 + 4x_3 - 2x_4. \tag{35}$$

(a) T takes vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$. What are n and m?

(b) Find the standard matrix of T. Is T onto? Is T one-to-one?

Answer: $n = 4$, $m = 1$, $\mathbf{A} = \begin{bmatrix} 3 & 0 & 4 & -2 \end{bmatrix}$. T is onto, but not one-to-one.

Week 5

The invertible matrix theorem is a list of conditions that tells us whether an $n \times n$ matrix $\mathbf{A}$ is invertible. It's not important to memorize all of the various conditions; however, it's very important to understand why these conditions are all equivalent.

Invertible matrix theorem: Let $\mathbf{A}$ be an $n \times n$ matrix. Then the following statements are equivalent (the statements are either all true, if $\mathbf{A}$ is invertible, or all false, if not):

(a) $\mathbf{A}$ is an invertible matrix.

(b) $\mathbf{A}$ is row equivalent to the $n \times n$ identity matrix. (*This should make sense, considering the method for finding* $\mathbf{A}^{-1}$ *– we row reduce* $\begin{bmatrix}\mathbf{A} \mid \mathbf{I}_n\end{bmatrix}$ *until* $\mathbf{A}$ *looks like* $\mathbf{I}_n$*!*)

(c) $\mathbf{A}$ has n pivot positions. (*This is just a restatement of (b); if* $\mathbf{A}$ *has less than* n *pivots, it can't be row equivalent to* $\mathbf{I}_n$*.*)

(d) The equation $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.

(e) The columns of $\mathbf{A}$ form a linearly independent set. (*This just follows from the definition of linear independence.*)

(f) The linear transformation $\mathbf{x} \to \mathbf{Ax}$ is one-to-one. (*Recall how we check whether a linear transformation is one-to-one.*)

(g) The equation $\mathbf{Ax} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$. (*This should make sense because* $\mathbf{A}$ *is row equivalent to* $\mathbf{I}_n$*.*)

(h) The columns of $\mathbf{A}$ span $\mathbb{R}^n$. (*A restatement of (c) and/or (g).*)

(i) The linear transformation $\mathbf{x} \to \mathbf{Ax}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.

(j) There is an $n \times n$ matrix $\mathbf{C}$ such that $\mathbf{CA} = \mathbf{I}_n$.

(k) There is an $n \times n$ matrix $\mathbf{D}$ such that $\mathbf{AD} = \mathbf{I}_n$.

(l) $\mathbf{A}^{\mathrm{T}}$ is an invertible matrix.

Don't let this list intimidate you. One of the main takeaways is that solving either $\mathbf{CA} = \mathbf{I}_n$ or $\mathbf{AD} = \mathbf{I}_n$ is sufficient for finding $\mathbf{A}^{-1}$.

Invertible linear transformations: The question of whether a linear transformation T is invertible is very closely related to the question of matrix inversion. In particular, if

$T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation with standard matrix $\mathbf{A}$, then T is invertible if and only if $\mathbf{A}$ is invertible. If $\mathbf{A}$ is invertible, we can form the inverse transformation $S : \mathbb{R}^n \to \mathbb{R}^n$, whose standard matrix is $\mathbf{A}^{-1}$. S and T then satisfy $S(T(\mathbf{x})) = \mathbf{x}$ and $T(S(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$.

Ex. With as few calculations as possible, determine if the following matrices are invertible:

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

Answer: Not invertible, invertible, invertible.

Ex. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation whose action is given by

$$T(x_1, x_2) = (-5x_1 + 9x_2, 4x_1 - 7x_2). \tag{36}$$

Show that T is invertible and find a formula for $T^{-1}(x_1, x_2)$.

Answer: T is invertible since $\begin{bmatrix} -5 & 9 \\ 4 & -7 \end{bmatrix}$ is invertible. $T^{-1}(x_1, x_2) = (7x_1 + 9x_2, 4x_1 + 5x_2)$.

A **subspace** of $\mathbb{R}^n$ is any set H in $\mathbb{R}^n$ such that

1. The zero vector is in H.
2. For each $\mathbf{u}$ and $\mathbf{v}$ in H, the sum $\mathbf{u} + \mathbf{v}$ is in H.
3. For each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H.

For $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in $\mathbb{R}^n$, the set of all linear combinations of these vectors is a subspace of $\mathbb{R}^n$. Try to justify this for yourself!

The **column space** of a matrix $\mathbf{A}$ is the set of all linear combinations of the columns of $\mathbf{A}$. The **null space** of a matrix $\mathbf{A}$ is the set of all solutions of the equation $\mathbf{Ax} = \mathbf{0}$.

Ex. Let

$$\mathbf{A} = \begin{bmatrix} 2 & -3 & -4 \\ -8 & 8 & 6 \\ 6 & -7 & -7 \end{bmatrix}, \quad \mathbf{p} = \begin{bmatrix} 6 \\ -10 \\ 11 \end{bmatrix}. \tag{37}$$

(a) How many vectors are in the column space of $\mathbf{A}$?

(b) Is $\mathbf{p}$ in the column space of $\mathbf{A}$?

Answer: Infinitely many; yes.

Part (a) is potentially confusing. There are infinitely many vectors in the column space of $\mathbf{A}$ *simply because there are infinitely many ways to take a linear combination of the columns. That is, there are infinitely many choices of* c_1, c_2, c_3 *in the linear combination* $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$, *where* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ *are the columns of* $\mathbf{A}$.

Ex. Consider the matrix

$$\begin{bmatrix} 3 & 2 & 1 & -5 \\ -9 & -4 & 1 & 7 \\ 9 & 2 & -5 & 1 \end{bmatrix}. \tag{38}$$

(a) The column space of $\mathbf{A}$ is a subspace of $\mathbb{R}^p$. What is p?

(b) The null space of $\mathbf{A}$ is a subspace of $\mathbb{R}^q$. What is q?

(c) Find a nonzero vector in the column space of $\mathbf{A}$ and find a nonzero vector in the null space of $\mathbf{A}$.

Answer: $p = 3$, $q = 4$. *For (c), possible answers are* $\begin{bmatrix} 1 \\ -3 \\ 3 \end{bmatrix}$ *for the column space and* $\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}$ *for the null space.*

Much of the material we've covered so far leads up to the following definition: A **basis** for a subspace H of $\mathbb{R}^n$ is a linearly independent set in H that spans H. The concept of a basis is extremely important! We've already seen the **standard basis** for $\mathbb{R}^n$:

$$\left\{ \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ \dots, \ \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\}. \tag{39}$$

To find the basis for the null space of a matrix $\mathbf{A}$, we simply solve the system $\mathbf{Ax} = \mathbf{0}$ and write the solution in parametric form. The linearly independent vectors that multiply the free parameters are then the elements of the basis for Nul $\mathbf{A}$. On the other hand, the basis for the column space of a matrix is simply the set containing the pivot columns of $\mathbf{A}$.

Question: If $\mathbf{A}$ *is an* $n \times n$ *invertible matrix, why is* Col $\mathbf{A}$ *a basis for* $\mathbb{R}^n$*? What is the null space of an invertible matrix?*

Ex. Determine if the following vectors form a basis for $\mathbb{R}^3$:

$$\begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}, \ \begin{bmatrix} 5 \\ 0 \\ 4 \end{bmatrix}, \ \begin{bmatrix} 6 \\ 3 \\ 2 \end{bmatrix}. \tag{40}$$

Answer: The vectors form a basis for $\mathbb{R}^3$*. You could check if the vectors span* $\mathbb{R}^3$ *and then check if they are linearly independent, but try to do it in one step!*

Ex. Consider the matrix $\mathbf{A}$, which row reduces as follows:

$$\mathbf{A} = \begin{bmatrix} 1 & 4 & 8 & -3 & -7 \\ -1 & 2 & 7 & 3 & 4 \\ -2 & 2 & 9 & 5 & 5 \\ 3 & 6 & 9 & -5 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 8 & 0 & 5 \\ 0 & 2 & 5 & 0 & -1 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{41}$$

Find a basis for Col $\mathbf{A}$ and a basis for Nul $\mathbf{A}$.

Answer: Col $\mathbf{A}$: $\left\{ \begin{bmatrix} 1 \\ -1 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 2 \\ 6 \end{bmatrix}, \begin{bmatrix} -3 \\ 3 \\ 5 \\ -5 \end{bmatrix} \right\}$, Nul $\mathbf{A}$: $\left\{ \begin{bmatrix} 2 \\ -2.5 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ 0.5 \\ 0 \\ -4 \\ 1 \end{bmatrix} \right\}$.

An interesting pattern is starting to show up – invertible matrices, which have n pivots, have n vectors in the basis for their column space (we already knew this had to be true, since the standard basis vectors in $\mathbb{R}^n$ form a basis for any n linearly independent vectors, such as the columns of an invertible matrix). In the previous example, $\mathbf{A}$ had 3 vectors in the basis for its column space since it had less than $n = 5$ pivots. The remaining $5 - 3 = 2$ degrees of freedom (the number of free variables!) showed up as vectors in the basis for the null space of $\mathbf{A}$! The **rank theorem** encapsulates this relationship.

First, we need a definition: the **dimension** $\dim H$ of a subspace H is the number of vectors in any basis for H. Then, for a matrix $\mathbf{A}$ with n columns, the rank theorem says

$$\dim(\text{Col } \mathbf{A}) + \dim(\text{Nul } \mathbf{A}) = n. \tag{42}$$

Question: Can $\mathbb{R}^3$ contain a 4-dimensional subspace? Explain.

Ex. True or False:

(a) Each line in $\mathbb{R}^n$ is a one-dimensional subspace of $\mathbb{R}^n$.

(b) If a set of p vectors spans a p-dimensional subspace H of $\mathbb{R}^n$, then these vectors form a basis for H.

(c) The dimension of Nul $\mathbf{A}$ is the number of free variables in the equation $\mathbf{Ax} = \mathbf{0}$.

Answer: F, T, T.

Week 6

The **determinant** is a quantity associated with any $n \times n$ matrix. Determinants have many useful properties and an interesting meaning in terms of linear transformations, but for now we'll focus on how to calculate them. We talked about the basic **cofactor expansion** method for calculating determinants in class; this is the most reliable way to find a determinant. Note that you can perform the cofactor expansion along any row or column!

There are sometimes easier methods than the cofactor expansion for finding determinants. For example, consider the cofactor expansion for a 3×3 matrix:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}. \tag{43}$$

If the matrix is *upper triangular*, this becomes

$$\begin{vmatrix} a & b & c \\ 0 & e & f \\ 0 & 0 & i \end{vmatrix} = a \begin{vmatrix} e & f \\ 0 & i \end{vmatrix} - b \begin{vmatrix} 0 & f \\ 0 & i \end{vmatrix} + c \begin{vmatrix} 0 & e \\ 0 & 0 \end{vmatrix} = aei - b(0) + c(0) = aei. \tag{44}$$

Thus, the determinant of an upper triangular matrix is the product of the diagonal entries. This has some interesting consequences; for example, if the matrix is non-invertible then there will be a row of zeros and the determinant will vanish! This proves that a matrix is invertible if and only if its determinant is nonzero.

Ex. Find the determinant of each matrix using a cofactor expansion.

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

(b) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 3 & 6 & 8 & 0 \\ 4 & 7 & 9 & 10 \end{bmatrix}$

Answer: (a) -23, (b) 400.

Since taking the determinant of a triangular matrix is so simple, you might wonder whether you can row reduce a matrix to bring it into echelon form before taking its determinant. Unfortunately it's not quite this easy, since performing row operations changes the determinant in general. In particular, for a square matrix $\mathbf{A}$,

1. If a multiple of one row is added to another row, the determinant does not change.
2. If two rows of $\mathbf{A}$ are switched, the determinant is multiplied by -1.
3. If a row of $\mathbf{A}$ is scaled by a number k, then the determinant is multiplied by k.

Finally, here are two properties of determinants for $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$:

1. $\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})$
2. $\det(\mathbf{A}^{\mathrm{T}}) = \det \mathbf{A}$

Note that $\det(\mathbf{A}+\mathbf{B})$ is **not** equal to $\det \mathbf{A} + \det \mathbf{B}$ in general!

Ex. Calculate the determinant by row reduction methods.

$$\begin{bmatrix} 1 & -1 & 3 & 0 \\ 0 & 1 & 5 & 4 \\ -1 & 2 & 8 & 5 \\ 3 & -1 & -2 & 3 \end{bmatrix} \tag{45}$$

Answer: 3.

Ex. Let $\mathbf{A}$ and $\mathbf{B}$ be 3×3 matrices, with $\det \mathbf{A} = 4$ and $\det \mathbf{B} = -3$. Use properties of determinants to compute:

(a) $\det \mathbf{AB}$

(b) $\det 5\mathbf{A}$

(c) $\det \mathbf{B}^{\mathrm{T}}$

(d) $\det \mathbf{A}^{-1}$

(e) $\det \mathbf{A}^3$

Answer: (a) -12, (b) 500, (c) -3, (d) 1/4, (e) 64.

Ex. Use determinants to decide whether the following vectors are linearly independent:

$$\begin{bmatrix} 7 \\ -4 \\ -6 \end{bmatrix}, \begin{bmatrix} -8 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \\ 5 \end{bmatrix}. \tag{46}$$

Answer: Linearly independent.

Ex. Use row reduction to show that

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

(b) $\begin{vmatrix} a & b & c \\ d & e & f \\ 0 & 0 & i \end{vmatrix} = aei - dbi.$

Vector Spaces: A vector space V is a set of objects, called vectors, on which we define addition and multiplication operations that satisfy several axioms. Many of the axioms might seem obvious, but it will pay off to learn them! For all vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in V and for all scalars c and d, we have:

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
3. There is a vector $\mathbf{0}$ in V such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
4. For each $\mathbf{u}$ in V, there is a vector $-\mathbf{u}$ in V such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
5. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
6. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
7. $c(d\mathbf{u}) = (cd)\mathbf{u}$.
8. $1\mathbf{u} = \mathbf{u}$.

Additionally, we require that vector spaces are *closed* under addition and scalar multiplication. This should remind you of subspaces, and indeed, a subspace is just a subset of a vector space that contains $\mathbf{0}$ and that is closed under addition and scalar multiplication.

A useful example of a vector space is the set of polynomials of degree n, called $\mathbb{P}_n$. This vector space consists of all polynomials of the form

$$\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n. \tag{47}$$

Try to convince yourself that $\mathbb{P}_n$ satisfies all of the addition, multiplication, and closure axioms.

When studying subspaces, we saw that the span of a set of vectors in $\mathbb{R}^n$ always forms a subspace of $\mathbb{R}^n$. A similar statement holds in the more general vector space context: If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are in a vector space V, then $\text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a subspace of V. Try to prove this yourself!

Ex. Determine if the given set is a subspace of $\mathbb{P}_n$, for an appropriate value of n.

(a) All polynomials of degree 3 with integer coefficients.

(b) All polynomials of degree n such that $\mathbf{p}(0) = \mathbf{0}$.

Answer: The first set is not a subspace; the second set is a subspace.

Ex. Let W be the set of all vectors of the form

$$\begin{bmatrix} 2b + 3c \\ -b \\ 2c \end{bmatrix}, \tag{48}$$

where b and c are arbitrary constants. Find vectors $\mathbf{u}$ and $\mathbf{v}$ such that $W = \text{Span}\{\mathbf{u}, \mathbf{v}\}$. Why does this show that W is a subspace of $\mathbb{R}^3$?

Answer: $W = \text{Span}\left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} \right\}$.

Ex. Let W be the set of all vectors of the form

(a) $\begin{bmatrix} 4a+3b \\ 0 \\ a+3b+c \end{bmatrix}$

(b) $\begin{bmatrix} 4a+3b \\ 1 \\ a+3b+c \end{bmatrix}$

where a, b, and c are arbitrary constants. Is W a vector space?

Answer: (a) Yes, (b) no.

Ex. Let W be the following set of vectors:

(a) $\left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a+b+c=2 \right\}$

(b) $\left\{ \begin{bmatrix} p \\ q \\ r \\ s \end{bmatrix} : p-3q=4s,\ 2p=s+5r \right\}$

Is W a vector space?

Answer: (a) No, (b) yes.

Week 7

Before we continue with the vector space material, let's review column and null spaces. If $\mathbf{A}$ is an $m \times n$ matrix, then the **null space** of $\mathbf{A}$, which is the set of all solutions to $\mathbf{Ax} = \mathbf{0}$, is a subspace of $\mathbb{R}^n$. The **column space** of $\mathbf{A}$, which is the span of the columns of $\mathbf{A}$, is a subspace of $\mathbb{R}^m$.

Prove these statements about the null space and the column space – what three properties do you need to check for each to show that it is a subspace?

We now generalize our previous results about transformations from $\mathbb{R}^n \to \mathbb{R}^m$ to results about transformations between vector spaces. A **linear transformation** $T : V \to W$ from a vector space V to a vector space W is a rule that assigns each vector $\mathbf{x}$ in V a unique vector $T(\mathbf{x})$ in W such that

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

for all $\mathbf{u}$, $\mathbf{v}$ in V and any constant c. The **kernel** (or null space – the terms are interchangeable) of a linear transformation T is the set of all $\mathbf{u}$ in V such that $T(\mathbf{u}) = \mathbf{0}$. *Note that this* $\mathbf{0}$ *is the zero vector in* W*!* The **range** of T is the set of all possible images $T(\mathbf{x})$ in W. It should be clear that if T is a transformation such that $T(\mathbf{x}) = \mathbf{Ax}$ then the kernel of T is the null space of $\mathbf{A}$ and the range of T is the column space of $\mathbf{A}$.

Ex. If $T : \mathbb{P}_2 \to \mathbb{R}^2$ is given by $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix}$,

(a) Show that T is a linear transformation.

(b) Find a polynomial $\mathbf{p}(t)$ in $\mathbb{P}_2$ that spans the kernel of T.

(c) Find a set of vectors in $\mathbb{R}^2$ that spans the range of T.

Answer: (a) Check the linearity properties, (b) $\mathbf{p}(t) = t - t^2$*, (c)* $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$.

The concepts of linear independence and basis extend naturally to vector spaces. In particular, a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in a vector space V is **linearly independent** if the equation

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0} \tag{49}$$

only has the trivial solution $c_1 = \cdots = c_n = 0$. A set of vectors $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ in a vector space V is a **basis** for V if the vectors are both linearly independent and they span V.

For example, the **standard basis** for $\mathbb{P}_n$ is the set $\{1, t, t^2, \ldots, t^n\}$. It's clear that this set spans $\mathbb{P}_n$, since a linear combination of these vectors is the most general polynomial of degree n: $a_0(1) + a_1(t) + a_2(t^2) + \cdots + a_n(t^n)$ can describe any polynomial in $\mathbb{P}_n$. The set is also linearly independent, since the only solution of $a_0 + a_1t + a_2t^2 + \cdots + a_nt^n = 0$ is the trivial solution $a_1 = \cdots = a_n = 0$. (proving this is slightly tricky).

Question: Consider the polynomials $\mathbf{p}(t) = 1+t^2$ *and* $\mathbf{q}(t) = 1-t^2$ *in* $\mathbb{P}_2$. *Is* $\{\mathbf{p}, \mathbf{q}\}$ *a linearly independent set? Why or why not?*

Finally, let's look at the idea of a spanning set through an example. Consider the polynomials $\mathbf{p}(t) = 1 + t^2$, $\mathbf{q}(t) = 1 - t^2$ and $\mathbf{r}(t) = 2$ in $\mathbb{P}_2$. It's clear that $\mathbf{p} + \mathbf{q} = \mathbf{r}$, so these vectors are not linearly independent. The **spanning set theorem** says that if a vector $\mathbf{v}$ in a set S is a linear combination of the other vectors, then the set S' formed by removing $\mathbf{v}$ from S has the same span as S: $\text{Span}(S) = \text{Span}(S')$. In our example, this means that $\text{Span}\{\mathbf{p}, \mathbf{q}, \mathbf{r}\} = \text{Span}\{\mathbf{p}, \mathbf{q}\}$.

Since basis vectors must be linearly independent, a valid basis for a subspace corresponds to the smallest possible spanning set for that subspace. Continuing with the example, $\{\mathbf{p}, \mathbf{q}\}$ (or $\{\mathbf{p}, \mathbf{r}\}$, or $\{\mathbf{q}, \mathbf{r}\}$) is a valid basis for $\text{Span}\{\mathbf{p}, \mathbf{q}, \mathbf{r}\}$, but $\{\mathbf{p}, \mathbf{q}, \mathbf{r}\}$ is not.

Ex. True or false:

(a) A set consisting of a single vector is linearly dependent.

(b) If a subspace $H = \text{Span}\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$, then $\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ is a basis for H.

(c) A linearly independent set in a subspace H is a basis for H.

(d) If the matrix $\mathbf{A}$ is row equivalent to $\mathbf{B}$, then the pivot columns of $\mathbf{B}$ form a basis for Col $\mathbf{A}$.

Answer: (a) F (unless it's the zero vector), (b) F, (c) F, (d) F (remember that we want the pivot columns of the original matrix).

Week 8

Change of Basis: Consider a basis $\mathbb{B}$ for $\mathbb{R}^2$ consisting of the vectors $\mathbf{b}_1$ and $\mathbf{b}_2$, where we know the components of $\mathbf{b}_1$ and $\mathbf{b}_2$ with respect to the standard basis. If we know the components of a vector $\mathbf{x}$ with respect to the standard basis, recall that we can write it as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2. \tag{50}$$

We now ask what the components of $\mathbf{x}$ are in the new basis $\mathbb{B}$. That is, if we write

$$\mathbf{x} = x_1'\mathbf{b}_1 + x_2'\mathbf{b}_2, \tag{51}$$

then what are the **coordinates** x_1' and x_2' of $\mathbf{x}$ with respect to $\mathbb{B}$? Remember that we know the coordinates of $\mathbf{b}_1$ and $\mathbf{b}_2$ with respect to the standard basis, so we can write $\mathbf{b}_1 = b_{11}\mathbf{e}_1 + b_{12}\mathbf{e}_2$ and $\mathbf{b}_2 = b_{21}\mathbf{e}_1 + b_{22}\mathbf{e}_2$. Inserting these into equation (2) gives

$$\mathbf{x} = (x_1'b_{11} + x_2'b_{21})\mathbf{e}_1 + (x_1'b_{12} + x_2'b_{22})\mathbf{e}_2. \tag{52}$$

Comparing equations (1) and (3), we see that $x_1 = x_1'b_{11} + x_2'b_{21}$ and $x_2 = x_1'b_{12} + x_2'b_{22}$, or

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \end{bmatrix}. \tag{53}$$

Letting $[\mathbf{x}]_{\mathbb{B}}$ denote the coordinates of $\mathbf{x}$ with respect to the basis $\mathbb{B}$, we therefore have

$$[\mathbf{x}] = \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix} [\mathbf{x}]_{\mathbb{B}} = \mathbf{P}[\mathbf{x}]_{\mathbb{B}}. \tag{54}$$

$\mathbf{P} = [\mathbf{b}_1\ \mathbf{b}_2]$ is called the change of coordinates matrix. Note that $[\mathbf{x}]$ with respect to the standard basis is the same thing as $\mathbf{x}$, but I've used the brackets in equation (5) to be consistent with the lecture.

Ex. Let $\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$. Find the coordinates of $\mathbf{x}$ with respect to the $\mathbb{B}$ basis.

Answer: $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2$, *or* $[\mathbf{x}]_{\mathbb{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

Ex. Let

$$\mathbf{b}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}, \ \mathbf{b}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}. \tag{55}$$

$\mathbb{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ is a basis for $H = \text{Span}\{\mathbf{b}_1, \mathbf{b}_2\}$. Determine if $\mathbf{x}$ is in H. If it is, find the coordinates of $\mathbf{x}$ relative to $\mathbb{B}$.

Answer: $\mathbf{x}$ *is in* H*, and* $[\mathbf{x}]_{\mathbb{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

Ex. In $\mathbb{P}_2$, find the change of coordinates matrix from the basis $\mathbb{B} = \{1 - 2t + t^2, 3 - 5t + 4t^2, 2t + 3t^2\}$ to the standard basis $\{1, t, t^2\}$. Then find the coordinates of the vector $\mathbf{p} = -1 + 2t$ in the $\mathbb{B}$ basis.

Answer: $\mathbf{P} = \begin{bmatrix} 1 & 3 & 0 \\ -2 & -5 & 2 \\ 1 & 4 & 3 \end{bmatrix}$, $[\mathbf{p}]_{\mathbb{B}} = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$.

Ex. Consider the polynomials 1, $1 - t$, $2 - 4t + t^2$, and $6 - 18t + 9t^2 - t^3$.

(a) Show that these polynomials form a basis $\mathbb{B}$ for $\mathbb{P}_3$.

(b) Find the coordinates of $\mathbf{p} = 5 + 5t - 2t^2$ relative to $\mathbb{B}$.

Answer: $[\mathbf{p}]_{\mathbb{B}} = \begin{bmatrix} 6 \\ 3 \\ -2 \\ 0 \end{bmatrix}$.

Dimension: Continuing the theme from last week, we now generalize some of the results for subspaces of $\mathbb{R}^n$ to results for general vector spaces:

1. The **dimension** of a vector space V is the number of vectors in any basis for V.
2. If H is a subspace of a vector space V, then $\dim H \leq \dim V$.
3. Let $T : V \to W$ be a linear transformation from a vector space V to a vector space W. Then $\dim(\text{Range}(T)) + \dim(\text{Kernel}(T)) = \dim V$. (This should remind you of the rank theorem!)

Ex. Find the dimension of the subspace H of all vectors in $\mathbb{R}^3$ whose first and third entries are equal, and find an explicit basis for this subspace.

Answer: $\dim H = 2$, $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$.

Ex. True or False:

(a) The dimension of the vector space $\mathbb{P}_n$ is n.

(b) $\mathbb{R}^2$ is a two-dimensional subspace of $\mathbb{R}^3$.

(c) If there is a linearly independent set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in a vector space V, then $\dim V \geq p$.

(d) If there is a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ that spans a vector space V, then $\dim V \leq p$.

Answer: F, F, T, T. Careful, part (b) is tricky!

We've seen many properties of column spaces. Now let's look at the **row space**, which is the set of all linear combinations of the row vectors of a matrix $\mathbf{A}$ (or the span of the rows of $\mathbf{A}$). For an $m \times n$ matrix, the row vectors are elements of $\mathbb{R}^n$, but they are written horizontally rather than vertically (for example, $\mathbf{r} = [r_1,\ r_2,\ \ldots,\ r_n]$). The row space is clearly a vector space since it's the span of a collection of vectors.

A key fact about row spaces is that *row reduction does not change the row space of a matrix.* So, to find the basis for the row space of a matrix, we can row reduce as much as possible (i.e., to reduced echelon form) and just read off a basis!

For example, if the matrix $\mathbf{A}$ row reduces as follows,

$$\mathbf{A} = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{56}$$

then a basis for Row $\mathbf{A}$ is just $\{[1, 3, -5, 1, 5],\ [0, 1, -2, 2, -7],\ [0, 0, 0, -4, 20]\}$.

Recall the rank theorem: $\dim(\text{Col } \mathbf{A}) + \dim(\text{Nul } \mathbf{A}) = n$, where n is the number of columns in $\mathbf{A}$. It turns out that $\dim(\text{Col } \mathbf{A}) = \dim(\text{Row } \mathbf{A})$, so we call this number the **rank** of $\mathbf{A}$.

Question: If $\mathbf{A}$ *is a* 7×9 *matrix with a two-dimensional null space, what is rank* $\mathbf{A}$*? Can a* 6×9 *matrix have a two-dimensional null space?*

Ex. If the matrix $\mathbf{A}$ row reduces as follows,

$$\mathbf{A} = \begin{bmatrix} 2 & 6 & -6 & 6 & 3 & 6 \\ -2 & -3 & 6 & -3 & 0 & -6 \\ 4 & 9 & -12 & 9 & 3 & 12 \\ -2 & 3 & 6 & 3 & 3 & -6 \end{bmatrix} \sim \begin{bmatrix} 2 & 6 & -6 & 6 & 3 & 6 \\ 0 & 3 & 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{57}$$

(a) Find rank $\mathbf{A}$ and dim(Nul $\mathbf{A}$) (no long calculations allowed!).

(b) Find bases for Col $\mathbf{A}$, Row $\mathbf{A}$, and Nul $\mathbf{A}$.

Answer: rank $\mathbf{A} = 3$, $\dim(\text{Nul } \mathbf{A}) = 3$,

Col $\mathbf{A}$: $\left\{ \begin{bmatrix} 2 \\ -2 \\ 4 \\ -2 \end{bmatrix}, \begin{bmatrix} 6 \\ -3 \\ 9 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 3 \\ 3 \end{bmatrix} \right\}$,

Row $\mathbf{A}$: $\{[2, 6, -6, 6, 3, 6], [0, 3, 0, 3, 3, 0], [0, 0, 0, 0, 3, 0]\}$,

Nul $\mathbf{A}$: $\left\{ \begin{bmatrix} 3 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Week 9

An **eigenvector** of an $n \times n$ matrix $\mathbf{A}$ is a nonzero vector $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ. Here, $\mathbf{x}$ is the eigenvector corresponding to the **eigenvalue** λ. Note that there may be more than one eigenvector corresponding to a single eigenvalue (but not the other way around!).

The eigenvalue equation is closely related to the equation $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}$. For a given λ, the set of all nontrivial solutions to this equation (i.e., the set of all eigenvectors corresponding to λ) is called the **eigenspace** of $\mathbf{A}$ corresponding to λ. This means that the eigenspace of $\mathbf{A}$ corresponding to λ is just the null space of $(\mathbf{A} - \lambda\mathbf{I}_n)$!

Ex. Let $\mathbf{A} = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. Find a basis for the eigenspace corresponding to the eigenvalue $\lambda = 2$.

Answer: One possible basis is $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 6 \\ 1 \end{bmatrix} \right\}$, *but the particular basis you find will depend on the free variables you choose.*

Let's try to get some intuition for what's going on here by finding the eigenvalues and eigenvectors of the following matrix:

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}. \tag{58}$$

It should be clear that the equation $\mathbf{A}\mathbf{x} = \mathbf{0}$ has two free variables. But $\mathbf{A}\mathbf{x} = \mathbf{0}$ is just the eigenvalue equation with $\lambda = 0$! So, if we can find vectors in the null space of $\mathbf{A}$, these will be eigenvectors corresponding to $\lambda = 0$.

What do these vectors look like? Trial and error (or a short calculation) leads to

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \tag{59}$$

(or a linear combination of these). Here we have two linearly independent eigenvectors corresponding to the *same* eigenvalue. However, the reverse of this can't happen:

If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent.

Question: For a 2×2 matrix $\mathbf{A}$, can there be more than two eigenvectors corresponding to a single eigenvalue? What is the maximum number of distinct eigenvalues $\mathbf{A}$ can have?

It's very useful to think of eigenvectors and eigenvalues geometrically in terms of linear transformations. For a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, eigenvectors of $\mathbf{A}$ correspond to vectors that T scales by some number λ, without changing their direction.

Ex. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation that

(a) Reflects vectors about the x-axis;

(b) Rotates vectors by 180 degrees counterclockwise.

Using pictures, find the eigenvectors and corresponding eigenvalues for each transformation.

Answer: (a) $\lambda = 1 \to \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\lambda = -1 \to \begin{bmatrix} 0 \\ -1 \end{bmatrix}$; *(b)* $\lambda = -1 \to \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Part (b) is a very good example to learn from. The two eigenvectors corresponding to $\lambda = -1$ actually form a basis for $\mathbb{R}^2$, so the eigenspace corresponding to $\lambda = -1$ is all of $\mathbb{R}^2$. This means that the null space of the matrix $(\mathbf{A} - (-1)\mathbf{I}_2)$ is two-dimensional. *When can this happen? Think about the rank theorem!*

The procedure for finding the eigenvalues and eigenvectors of any matrix is straightforward. Remember that we need to find nontrivial solutions to $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}$. This equation only has a nontrivial solution if there is at least one free variable, meaning that $(\mathbf{A} - \lambda\mathbf{I}_n)$ is not invertible. But this means that $\det(\mathbf{A} - \lambda\mathbf{I}_n) = 0$!

Note: $\det(\mathbf{A} - \lambda\mathbf{I}_n) = 0$ is called the **characteristic equation**, and the **characteristic polynomial** is just this equation written as a polynomial in λ. For an $n \times n$ matrix, the polynomial will be of order λ^n.

So, here's the step-by-step procedure for finding the eigenvalues and eigenvectors of any $n \times n$ matrix $\mathbf{A}$:

1. Solve $\det(\mathbf{A} - \lambda\mathbf{I}_n) = 0$ to find the eigenvalues.

2. For each eigenvalue λ_i, solve $(\mathbf{A} - \lambda_i\mathbf{I}_n)\mathbf{x} = \mathbf{0}$ to find the corresponding eigenvectors.

Note: $\lambda = 0$ is an allowed eigenvalue, but $\mathbf{x} = \mathbf{0}$ is never an allowed eigenvector.

Ex. Find the eigenvalues and corresponding eigenvectors of the following matrices.

(a) $\begin{bmatrix} 8 & 2 \\ 3 & 3 \end{bmatrix}$

(b) $\begin{bmatrix} 8 & 4 \\ 4 & 8 \end{bmatrix}$

(c) $\begin{bmatrix} 3 & 0 & 0 \\ 2 & 1 & 4 \\ 1 & 0 & 4 \end{bmatrix}$

Answer: (a) $\lambda = 9 \to \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\lambda = 2 \to \begin{bmatrix} 1 \\ -3 \end{bmatrix}$; *(b)* $\lambda = 12 \to \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\lambda = 4 \to \begin{bmatrix} 1 \\ -1 \end{bmatrix}$;
(c) $\lambda = 4 \to \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix}$, $\lambda = 3 \to \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, $\lambda = 1 \to \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

The **multiplicity** of an eigenvalue λ is the number of times it appears as a root of the characteristic equation. For example, if the characteristic equation of a 6×6 matrix is $\lambda^6 - 4\lambda^5 - 12\lambda^4 = \lambda^4(\lambda - 6)(\lambda + 2)$, then the eigenvalues are 0 (multiplicity 4), 6 (multiplicity 1), and -2 (multiplicity 1); or, we could list them as $0, 0, 0, 0, 6, -2$.

Ex. For the following matrix, find the eigenvalues (including their multiplicities) and find a basis for the eigenspace corresponding to each eigenvalue:

$$\begin{bmatrix} 5 & 5 & 0 & 2 \\ 0 & 2 & -3 & 6 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 5 \end{bmatrix}. \tag{60}$$

Answer: $\lambda = 5$ *(multiplicity 2),* $\lambda = 3$ *(multiplicity 1),* $\lambda = 2$ *(multiplicity 1);* $\lambda = 5 \to$

$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}$, $\lambda = 3 \to \left\{ \begin{bmatrix} -15 \\ 6 \\ -2 \\ 0 \end{bmatrix} \right\}$, $\lambda = 2 \to \left\{ \begin{bmatrix} -5 \\ 3 \\ 0 \\ 0 \end{bmatrix} \right\}$.

Two $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ are **similar** if there is an invertible matrix P such that $\mathbf{A} = \mathbf{PBP}^{-1}$ (or $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$). If $\mathbf{A}$ and $\mathbf{B}$ are similar matrices, then they have the same characteristic equation and the same eigenvalues. However, the eigenvectors corresponding to each eigenvalue will *not* generally be the same for each matrix. In addition, if two matrices have the same eigenvalues then they are *not* necessarily similar.

If an $n \times n$ matrix $\mathbf{A}$ has n linearly independent eigenvectors, then $\mathbf{A}$ is similar to a diagonal matrix $\mathbf{D}$. The entries of this diagonal matrix are the eigenvalues corresponding to each eigenvector, and the columns of the matrix $\mathbf{P}$ that brings $\mathbf{A}$ into a diagonal form are the eigenvectors of $\mathbf{A}$. Let's sum this up using equations:

If $\mathbf{A}$ has n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ (*note: some of these eigenvalues might be repeated!*), then $\mathbf{A} = \mathbf{PDP}^{-1}$, where the matrix $\mathbf{P} = [\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_n]$ and

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}. \tag{61}$$

Bringing a matrix into a diagonal form using its eigenvectors is called **diagonalization**. Here are the steps to diagonalize any $n \times n$ matrix $\mathbf{A}$:

1. Find the eigenvalues and the corresponding eigenvectors of $\mathbf{A}$.
2. Check whether there are n linearly independent eigenvectors.
3. If there are n linearly independent eigenvectors, then $\mathbf{A}$ is diagonalizable, meaning $\mathbf{A} = \mathbf{PDP}^{-1}$. The columns of $\mathbf{P}$ are the eigenvectors and the diagonal entries of $\mathbf{D}$ are the corresponding eigenvalues.

Let's look at an example: suppose we want to diagonalize the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}. \tag{62}$$

This means that we need to find an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = \mathbf{PDP}^{-1}$. Step 1 gives us the eigenvalues $\lambda_1 = -2$, $\lambda_2 = -2$, $\lambda_3 = 1$, and the corresponding eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}. \tag{63}$$

These vectors are clearly linearly independent, but it's good to do step 2 carefully by row reducing the matrix formed out of these vectors. We have three linearly independent eigenvectors, so $\mathbf{A}$ is diagonalizable: $\mathbf{A} = \mathbf{PDP}^{-1}$, with

$$\mathbf{P} = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}, \ \mathbf{D} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{64}$$

You should check that this works by taking the inverse of $\mathbf{P}$ and finding $\mathbf{PDP}^{-1}$. (There's also a way to check whether $\mathbf{A} = \mathbf{PDP}^{-1}$ works without having to find $\mathbf{P}^{-1}$!)

Ex. Diagonalize the following matrices, if possible:

(a) $\begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 \\ 6 & -1 \end{bmatrix}$

Answer: (a) Not diagonalizable; (b) $\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

Diagonalization is useful for finding the powers of a matrix. If $\mathbf{A}$ is a diagonalizable matrix with $\mathbf{A} = \mathbf{PDP}^{-1}$, then $\mathbf{A}^2 = \mathbf{PDP}^{-1}\mathbf{PDP}^{-1} = \mathbf{PDDP}^{-1} = \mathbf{PD}^2\mathbf{P}^{-1}$. This works the same way for higher powers, so $\mathbf{A}^k = \mathbf{PD}^k\mathbf{P}^{-1}$. This is helpful because $\mathbf{D}^k$ is very easy to find (remember, $\mathbf{D}$ is diagonal!).

Ex. For the matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 6 & -1 \end{bmatrix}$ from the previous exercise, find $\mathbf{A}^5$.

Answer: $\mathbf{A}^5 = \mathbf{A}$.

Week 10

Ex. Diagonalize the following matrix, if possible:

$$\begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 1 & 4 & -3 & 0 \\ -1 & -2 & 0 & -3 \end{bmatrix}. \tag{65}$$

Answer: The matrix is diagonalizable, with

$$\mathbf{P} = \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ -1 & -1 & 0 & 1 \end{bmatrix}, \ \mathbf{D} = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}. \tag{66}$$

Recall that diagonalization allows us to find the powers of a matrix very quickly. In particular, if $\mathbf{A} = \mathbf{PDP}^{-1}$, then $\mathbf{A}^k = \mathbf{PD}^k\mathbf{P}^{-1}$. This is "easy" because $\mathbf{D}^k$ is simply the matrix with each diagonal entry of $\mathbf{D}$ raised to the kth power. However, don't forget to do the final matrix multiplication step with $\mathbf{P}$ and $\mathbf{P}^{-1}$!

Ex. Compute the following powers of each matrix using diagonalization:

(a) $\mathbf{A}^4$, where $\mathbf{A} = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$.

(b) $\mathbf{B}^2$, where $\mathbf{B} = \begin{bmatrix} a & 0 \\ 2(a-b) & b \end{bmatrix}$ $(a \neq b)$.

Answer: (a) $\begin{bmatrix} 46 & -45 \\ 30 & -29 \end{bmatrix}$, *(b)* $\begin{bmatrix} a^2 & 0 \\ 2(a^2-b^2) & b^2 \end{bmatrix}$.

Ex. True or False:

(a) If an $n \times n$ matrix $\mathbf{A}$ is diagonalizable, then $\mathbf{A}$ has n distinct eigenvalues.

(b) If a matrix $\mathbf{A}$ is invertible, then $\mathbf{A}$ is diagonalizable.

(c) If the eigenvectors of an $n \times n$ matrix $\mathbf{A}$ form a basis for $\mathbb{R}^n$, then $\mathbf{A}$ is diagonalizable.

(d) $\mathbf{A}$ is a 5×5 matrix with two distinct eigenvalues; one eigenspace is three-dimensional, and the other eigenspace is two-dimensional. $\mathbf{A}$ is diagonalizable.

Answer: F, F, T, T.

Warning – so far, all of the matrices that we've encountered have had real eigenvalues, but this isn't always the case. Consider the matrix

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{67}$$

The characteristic equation for this matrix is $\lambda^2 + 1 = 0$, so there are *no real eigenvalues*! (The complex eigenvalues are $\pm i$.)

Here are two useful facts about similar matrices that we mentioned briefly last week:

1. Similar matrices have the same determinant. If $\mathbf{A}$ is similar to $\mathbf{B}$, then $\mathbf{A} = \mathbf{PBP}^{-1}$ for some invertible matrix $\mathbf{P}$. Then

$$\det(\mathbf{A}) = \det(\mathbf{PBP}^{-1}) = \det(\mathbf{P})\,\det(\mathbf{B})\,\det(\mathbf{P}^{-1}) = \det(\mathbf{P})\,\det(\mathbf{B})\,\frac{1}{\det(\mathbf{P})} = \det(\mathbf{B}). \tag{68}$$

2. Similar matrices have the same eigenvalues. If $\mathbf{A} = \mathbf{PBP}^{-1}$, then

$$\mathbf{A} - \lambda\mathbf{I}_n = \mathbf{PBP}^{-1} - \lambda\mathbf{PP}^{-1} = \mathbf{P}(\mathbf{BP}^{-1} - \lambda\mathbf{P}^{-1}) = \mathbf{P}(\mathbf{B} - \lambda\mathbf{I}_n)\mathbf{P}^{-1}. \tag{69}$$

   This means that $\det(\mathbf{A} - \lambda\mathbf{I}_n) = \det(\mathbf{B} - \lambda\mathbf{I}_n)$, so $\mathbf{A}$ and $\mathbf{B}$ have the same characteristic equation and the same eigenvalues.

We can combine many of the concepts we've seen so far, including linear transformations, abstract vector spaces, and eigenvectors/eigenvalues, by looking at the eigenvectors of a general linear transformation $T : V \to V$. Remember that we can write the action of T on vectors in V as $T(\mathbf{x}) = \mathbf{Ax}$, where the standard matrix $\mathbf{A} = [T(\mathbf{e}_1) \dots T(\mathbf{e}_n)]$ ($\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is the standard basis for V).

We can convert vectors in V to column vectors by using their coordinates with respect to the standard basis. For example, the polynomial $\mathbf{p}(t) = 1 + t - 5t^2$ in $\mathbb{P}_2$ corresponds to the coordinate vector

$$[\mathbf{p}] = \begin{bmatrix} 1 \\ 1 \\ -5 \end{bmatrix}. \tag{70}$$

Make sure to pay attention to the vector space that you're working in; if we had been thinking of $\mathbf{p}(t) = 1 + t - 5t^2$ as an element of $\mathbb{P}_3$, then its coordinate vector would have been

$$[\mathbf{p}] = \begin{bmatrix} 1 \\ 1 \\ -5 \\ 0 \end{bmatrix}. \tag{71}$$

Ex. Let $T : \mathbb{P}_2 \to \mathbb{P}_2$ be given by $T(\mathbf{p}) = \mathbf{p}(0) + \mathbf{p}(0)t + \mathbf{p}(-1)t^2$.

(a) Find the standard matrix $\mathbf{A}$ for this transformation.

(b) What are the eigenvalues and eigenvectors of $\mathbf{A}$? Write the eigenvectors as polynomials.

(c) Is $\mathbf{A}$ diagonalizable?

Answer: (a) $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix}$, *(b)* $\lambda = 1 \to 1 + t,\ t^2,\ \lambda = 0 \to t + t^2$, *(c) yes.*

Ex. Let $T : \mathbb{P}_1 \to \mathbb{P}_1$ be given by $T(\mathbf{p}) = \mathbf{p}'(1) + 2\mathbf{p}'(1)t$.

(a) Find the standard matrix $\mathbf{A}$ for this transformation.

(b) What are the eigenvalues and eigenvectors of $\mathbf{A}$? Write the eigenvectors as polynomials.

(c) Diagonalize $\mathbf{A}$, if possible.

(d) Find $T^4(\mathbf{q})$, where $\mathbf{q} = 5 + 5t$. (Hint: use part (c)!)

Answer: (a) $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}$, *(b)* $\lambda = 2 \to 1 + 2t,\ \lambda = 0 \to 1 + 0t$, *(c)* $\mathbf{P} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$, $\mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$,

(d) $T^4(\mathbf{q}) = 40 + 80t$.

One question that we've never really addressed is why diagonalization works: Why should an $n \times n$ matrix with n linearly independent eigenvectors be similar to a diagonal matrix, and why do the eigenvectors and eigenvalues show up in the formula $\mathbf{A} = \mathbf{PDP}^{-1}$? The answer is very elegant, and it's related to the concept of change of basis.

Recall that if we have a basis $\mathbb{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ for $\mathbb{R}^n$, then the coordinates of a vector $\mathbf{x}$ with respect to the standard basis are related to the coordinates of $\mathbf{x}$ with respect to the $\mathbb{B}$ basis by

$$[\mathbf{x}] = \mathbf{P}[\mathbf{x}]_{\mathbb{B}}, \tag{72}$$

where $\mathbf{P} = [\mathbf{b}_1 \ldots \mathbf{b}_n]$.

How does a *matrix* transform under a change of basis? So far, we've always implicitly used matrices defined with respect to the standard basis; when we write $\mathbf{A}$, we mean $[\mathbf{A}(\mathbf{e}_1) \ldots \mathbf{A}(\mathbf{e}_n)]$. To find $[\mathbf{A}]_{\mathbb{B}}$, we need to account for the fact that both the basis vectors *and* the coordinate system that we're expressing the columns of $\mathbf{A}$ in change. In particular, $[\mathbf{A}]_{\mathbb{B}} = [[\mathbf{A}(\mathbf{b}_1)]_{\mathbb{B}} \ldots [\mathbf{A}(\mathbf{b}_n)]_{\mathbb{B}}]$. But $\mathbf{A}(\mathbf{b})$ is just a vector, so we can use equation 6 to write $[\mathbf{A}(\mathbf{b})]_{\mathbb{B}} = \mathbf{P}^{-1}\mathbf{A}(\mathbf{b})$. Then

$$[\mathbf{A}]_{\mathbb{B}} = [\mathbf{P}^{-1}\mathbf{A}(\mathbf{b}_1) \ldots \mathbf{P}^{-1}\mathbf{A}(\mathbf{b}_n)] = \mathbf{P}^{-1}\mathbf{A}[\mathbf{b}_1 \ldots \mathbf{b}_n] = \mathbf{P}^{-1}\mathbf{AP}, \tag{73}$$

or equivalently $\mathbf{A} = \mathbf{P}[\mathbf{A}]_{\mathbb{B}}\mathbf{P}^{-1}$.

This should look very familiar! What we've found is that applying a similarity transformation to $\mathbf{A}$ is related to changing the basis we're working in. In particular, if we write $\mathbf{A} = \mathbf{PDP}^{-1}$, the matrix $\mathbf{D}$ is just the old matrix $\mathbf{A}$ expressed in the basis formed by the columns of $\mathbf{P}$.

Finally, we can relate this to eigenvectors and eigenvalues. If an $n \times n$ matrix $\mathbf{A}$ has n linearly independent eigenvectors, then $\mathbf{A} = \mathbf{PDP}^{-1}$ means that the matrix $\mathbf{A}$, expressed in a basis formed by its eigenvectors, is diagonal. This should make sense: if we align our coordinate system with the eigenvectors of $\mathbf{A}$, then $\mathbf{A}$ simply scales each basis vector and doesn't "mix" them together. The amount that each vector gets scaled by is the corresponding eigenvalue, which is why these are the entries of the matrix $\mathbf{D}$!