MATH031 Spring 2024 Worksheet Solutions & Handouts & Theorems & Knuckles

TAylor

Compiled at 22:13, Friday 25th October, 2024

Contents

Week #1 Worksheet Solutions

1.1 Row Operations and Systems of Equations

1. State in words the next two elementary row operations you think should be performed in the process of solving the system associated with the following augmented matrix.

$$\left[\begin{array}{cccc|c} 1 & -6 & 4 & 0 & 1 \\ 0 & -2 & -7 & 0 & 4 \\ 0 & 0 & 1 & 2 & -3 \\ 0 & 0 & 3 & 1 & 6 \end{array}\right]$$

There is no one correct answer I'm looking for. Just describe two (productive) operations *you* would do.

Solution. Some examples would be

$$\begin{aligned} R_3 : &\quad -3R_3 + R_4 \\ R_2 : &\quad R_2 + 7R_3 \\ R_1 : &\quad R_1 - 4R_3 \\ R_1 : &\quad R_1 - 3R_2 \end{aligned}$$

In an attempt to eliminate the nonpivot entries in the second and third columns. ■

2. The augmented matrix of some linear systems have been reduced to the following forms.

(i) Why does this system require no further solving to describe the solution set? Describe the solution set (if it exists, or explain why a solution does not exist).

$$\left[\begin{array}{ccc|c} 1 & 7 & 3 & -4 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array}\right]$$

Solution. Row 3 says $0 = 1$, so no solution. ■

(ii) Explain why $x_1 = 0$, $x_2 = 0$, $x_3 = 0$ (the trivial solution) is the only solution to this system without performing any more row operations.

$$\left[\begin{array}{ccc|c} 1 & -4 & 9 & 0 \\ 0 & 1 & 7 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right]$$

Solution. Equation 3 tells us $x_3 = 0$. Plugging that into equation 2 tells us $x_2 = 0$. Plugging $x_2 = x_3 = 0$ into equation 1 tells us $x_1 = 0$. ■

1.2 Linear Transformation Preview

3. Consider the system of equations

$$\left[\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right] \iff \begin{array}{ccccccc} x & & - & z & & = & 1 \\ & y & - & z & & = & 2 \\ & & & & w & = & 3 \end{array}$$

(a) Show $\begin{bmatrix} x = 1 \\ y = 2 \\ z = 0 \\ w = 3 \end{bmatrix}$ is a solution by plugging it in. We'll call this solution $\mathbf{x}_1$.

Solution.

$$\begin{aligned} 1 \quad - 0 &= 1 \\ 2 - 0 &= 2 \quad \checkmark \\ 3 &= 3 \end{aligned}$$

■

(b) We define the function

$$T(x,y,z,w) = \begin{bmatrix} x-z \\ y-z \\ w \end{bmatrix}$$

$$= x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix} + w \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

(Notice how the vectors on each variable are the columns of the original matrix.)

Compute $T(1,2,0,3)$ (evaluating T at $\mathbf{x}_1$). Also compute $T(1,1,1,0)$.

(Just use the $T(x,y,z,w) = \begin{bmatrix} x-z \\ y-z \\ w \end{bmatrix}$ form.)

Solution.

$$T(1,2,0,3) = \begin{bmatrix} 1-0 \\ 2-0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad T(1,1,1,0) = \begin{bmatrix} 1-1 \\ 1-1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

■

(c) Explain why the function problem

$$T(x, y, z, w) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\text{That is, } \begin{bmatrix} x - z \\ y - z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

is equivalent to the original system of equations, and explain how the previous part shows $\mathbf{x}_1$ is a solution.

Solution. We have just put each equation into the entry of a vector. Two vectors are equal if and only if each entry is equal, so saying that these two vectors are equal requires all three equations to be satisfied.

In short: we just put the equations in the entries of a vector. ■

(d) Show that for any $t \in \mathbb{R}$,

$$\mathbf{x}_g = \begin{bmatrix} x = 1 + t \\ y = 2 + t \\ z = t \\ w = 3 \end{bmatrix}$$

is also a solution (just plug it in).

Solution.

$$T(1 + t, 2 + t, t, 3) = \begin{bmatrix} (1 + t) - (t) \\ (2 + t) - (t) \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \checkmark$$

■

(e) Show that

$$T(kx, ky, kz, kw) = k\,T(x, y, z, w)$$

Solution.

$$T(kx, ky, kz, kw) = \begin{bmatrix} kx - kz \\ ky - kz \\ kw \end{bmatrix} = k \begin{bmatrix} x - z \\ y - z \\ w \end{bmatrix} = kT(x, y, z, w)$$

■

(f) Define

$$\mathbf{x}_h = \mathbf{x}_g - \mathbf{x}_1 = \begin{bmatrix} x = t \\ y = t \\ z = t \\ w = 0 \end{bmatrix}$$

What is $T(\mathbf{x}_h) = T(t, t, t, 0)$? Explain how the previous part implies this by knowing $T(1, 1, 1, 0)$.

Solution.

$$T(t, t, t, 0) = \begin{bmatrix} t - t \\ t - t \\ 0 \end{bmatrix} = \vec{0}$$

Since $T(1, 1, 1, 0) = \vec{0}$, and $T(t, t, t, 0) = tT(1, 1, 1, 0)$ from the previous part, it must be zero for all t. That is, because $t\vec{0} = \vec{0}$ for all t. ■

Remark 1. It is a linear algebra fact that if $\mathbf{x}_1$ and $\mathbf{x}_2$ are both particular solutions to a system (or both preimages of a vector), then their difference $\mathbf{x}_2 - \mathbf{x}_1$ is in the kernel. This is straightforward to prove:

$$T(\mathbf{x}_2 - \mathbf{x}_1) = T(\mathbf{x}_2) - T(\mathbf{x}_1) = \mathbf{b} - \mathbf{b} = \vec{0}$$

4. Find the general solutions of the systems with given augmented matrices (if a solution exists). Identify the pivot columns.

(a) $\left[\begin{array}{ccc|c} 1 & 4 & 0 & 7 \\ 2 & 9 & 0 & 10 \end{array}\right]$ (b) $\left[\begin{array}{ccc|c} 1 & -2 & -1 & 3 \\ 3 & -6 & -2 & 11 \end{array}\right]$

(c) $\left[\begin{array}{ccc|c} 1 & -4 & 2 & 0 \\ -3 & 12 & -6 & 0 \\ -2 & 8 & -4 & 0 \end{array}\right]$

(d) $\left[\begin{array}{ccccc|c} 1 & 2 & -5 & -15 & 1252534\ln(2) & -5 \\ 0 & 0 & 1 & 3 & -\pi^{23}\cos(1) & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$

Solution. The REFs of the matrices are as follows:

(a) $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 23 \\ 0 & 1 & 0 & -4 \end{array}\right] \implies \begin{cases} x_1 = 23 \\ x_2 = -4 \\ x_3 = t \end{cases}$ Pivots: x_1, x_2

(b) $\left[\begin{array}{ccc|c} 1 & -2 & 0 & 5 \\ 0 & 0 & 1 & 2 \end{array}\right] \implies \begin{cases} x_1 = 5 + 2t \\ x_2 = t \\ x_3 = 2 \end{cases}$ Pivots: x_1, x_3

(c) $\left[\begin{array}{ccc|c} 1 & -4 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \implies \begin{cases} x_1 = 4t - 2s \\ x_2 = t \\ x_3 = s \end{cases}$ Pivots: x_1

(d) $\left[\begin{array}{ccccc|c} 1 & 2 & 0 & 0 & 0 & 5 \\ 0 & 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \implies \begin{cases} x_1 = 5 - 2t \\ x_2 = t \\ x_3 = 2 - 3s \\ x_4 = s \\ x_5 = 0 \end{cases}$ Pivots: x_1, x_3, x_5

$\blacksquare$

5. Consider the matrix and its reduced echelon form

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 & 1 \\ 1 & 2 & 3 & 2 & 3 \\ 1 & 2 & 1 & 0 & 2 \end{pmatrix}, \qquad R = \begin{pmatrix} 1 & 2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The pivot columns are columns 1, 3, and 5.

(a) Verify the dependence relations of the nonpivot columns of R:

- Col 2 = 2(Col 1)
- Col 4 = Col 3 − Col 1

Solution.

$$\begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \checkmark \qquad \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \checkmark$$

■

(b) Verify these relations hold for A as well. This is true in general: the column dependencies are unchanged by row operations.

Solution.

$$\begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \checkmark \qquad \begin{bmatrix} 0 \\ 0 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \checkmark$$

■

(c) The dependence relations can be rewritten as

- $-2(\text{Col } 1) + \text{Col } 2 = 0$

- Col 1 − Col 3 + Col 4 = 0

Verify that

$$\begin{bmatrix} x_1 = -2t + s \\ x_2 = t \\ x_3 = -s \\ x_4 = s \\ x_5 = 0 \end{bmatrix}$$

is a solution to

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

for all $s, t \in \mathbb{R}$. Can you see how the dependence relations give this solution?

Solution.

$$\begin{aligned} 1(-2t+s) + 2(t) + 1(-s) + 0(s) + 1(0) &= 0 \quad \checkmark \\ 1(-2t+s) + 2(t) + 1(-s) + 0(s) + 1(0) &= 0 \quad \checkmark \\ 1(-2t+s) + 2(t) + 3(-s) + 2(s) + 3(0) &= 0 \quad \checkmark \\ 1(-2t+s) + 2(t) + 1(-s) + 0(s) + 2(0) &= 0 \quad \checkmark \end{aligned}$$

All I expect from your solutions is to notice that the numbers from the column dependence relations sort of match the numbers in the solution. But here's a way to look at it.

If we just consider the t terms (i.e., we can let $s = 0$), we get

$$\begin{aligned} 1(-2t) + 2(t) + 1(0) + 0(0) + 1(0) &= 0 \\ 1(-2t) + 2(t) + 1(0) + 0(0) + 1(0) &= 0 \\ 1(-2t) + 2(t) + 3(0) + 2(0) + 3(0) &= 0 \\ 1(-2t) + 2(t) + 1(0) + 0(0) + 2(0) &= 0 \end{aligned}$$

Which we can see cancel out exactly because the second column is twice the first. That is, if we take −2 of the first column and add the second column, we get zero. So we can multiply it by t and also get zero.

$-2(\text{Col } 1) + \text{Col } 2 = 0 \implies -2t(\text{Col } 1) + t(\text{Col } 2) = 0$

Similarly, removing the t terms:

$$\begin{aligned}
1(s) + 0(s) + 1(-s) + 0(s) + 1(0) = 0 \\
1(s) + 0(s) + 1(-s) + 0(s) + 1(0) = 0 \\
1(s) + 0(s) + 3(-s) + 2(s) + 3(0) = 0 \\
1(s) + 0(s) + 1(-s) + 0(s) + 2(0) = 0
\end{aligned}$$

Here we are taking 1 of the first column, -1 of the third column, and 1 of the fourth column and then multiplying by s.

$\text{Col } 1 - \text{Col } 3 + \text{Col } 4 = 0$

$\implies s(\text{Col } 1) + -s(\text{Col } 3) + s(\text{Col } 4) = 0$

Later in the quarter, I can give a better explanation for why this works. ■

Week #2 Worksheet Solutions

2.1 Vectors, Vector / Matrix Equations

1. Compute $\mathbf{u}+\mathbf{v}$ and $\mathbf{u}-2\mathbf{v}$

$$\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$$

Solution. $\mathbf{u}+\mathbf{v} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$ and $\mathbf{u}-2\mathbf{v} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$ ■

2. Write a system of equations that is equivalent to the given vector equation and solve it (if a solution exists).

$$x_1 \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ -6 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

Solution.

$$\begin{array}{rcrcrcr} x_1 & & & + & 5x_3 & = & 1 \\ -2x_1 & + & x_2 & - & 6x_3 & = & -1 \\ & & 2x_2 & + & 8x_3 & = & 2 \end{array} \iff \left[\begin{array}{ccc|c} 1 & 0 & 5 & 1 \\ -2 & 1 & -6 & -1 \\ 0 & 2 & 8 & 2 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|c} 1 & 0 & 5 & 1 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \implies \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1-5t \\ 1-4t \\ t \end{bmatrix}$$

One particular solution is $\begin{bmatrix} x_1 = 1 \\ x_2 = 1 \\ x_3 = 0 \end{bmatrix}$, which corresponds to the fact that

$$1\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + 0\begin{bmatrix} 5 \\ -6 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

■

3. Determine if $\mathbf{b}$ is a linear combination of $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_3$.

(a)

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 5 \\ -6 \\ 8 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

Hint: Look at problem 2.

Solution. This is directly implied by problem two, since we got a solution. Remember, **a vector equation (and by extension, a system of equations) is consistent if and only if the RHS vector is a linear combination of the variable vectors.** (i.e. $\mathbf{b}$ is in the column space / span of the variable vectors).

We showed that

$$1\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + 0\begin{bmatrix} 5 \\ -6 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

so we've written $\mathbf{b}$ as a linear combination of the $\mathbf{a}$ vectors.

Note: this means $\mathbf{b}$ is in the span of $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$. ■

(b)

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 2 \\ 0 \\ 8 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -5 \\ 1 \\ -7 \end{bmatrix}$$

Solution. Writing

$$x_1 \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 0 \\ 8 \end{bmatrix} = \begin{bmatrix} -5 \\ 1 \\ -7 \end{bmatrix}$$

as an augmented matrix,

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & -5 \\ -2 & 4 & 0 & 1 \\ 2 & 4 & 8 & -7 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 2 & -5 \\ 0 & 4 & 4 & -9 \\ 0 & 4 & 4 & 3 \end{array}\right]$$

From here, we can see that equations two and three tell us $4x_2 + 4x_3$ need to be equal to both -9 and 3.

You may check with a calculator that $-9 \neq 3$, so this system is inconsistent and has no solution. That is, we *cannot* write $\begin{bmatrix} -5 \\ 1 \\ -7 \end{bmatrix}$ as a linear combination of the **a** vectors.

Note: this means **b** is **not** in the span of $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$. ■

4. Compute the products using both (i) the column perspective definition ($A\mathbf{x} = x_1\mathbf{a}_1 + \ldots + x_n\mathbf{a}_n$. See page 35 in the textbook), and (ii) the row-vector rule for computing $A\mathbf{x}$ (see page 38 in the textbook). If a product is undefined, explain why.

   Note: These methods are equivalent, but the point of this problem is to use both so you can *see* that they're equivalent.

   (a) $\begin{bmatrix} 2 \\ 6 \\ -1 \end{bmatrix} \begin{bmatrix} 5 \\ -1 \end{bmatrix}$ (b) $\begin{bmatrix} 8 & 3 & -4 \\ 5 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

   Solution.

(a) is not defined, because we have a $(3\times\underline{1})\cdot(\underline{2}\times 1)$. $1 \neq 2$. Recall that matrix multiplication is defined only when the inner dimensions match (the number of columns of the left matrix match the rows of the right matrix)

$$(m \times \underline{\mathbf{n}}) \cdot (\underline{\mathbf{n}} \times p) \mapsto m \times p$$

(b) This multiplication is defined, because the left matrix has three columns and the right matrix has three rows. Both have three, so we can multiply them.

(i)

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

(ii)

$$\begin{bmatrix} 8 & 3 & -4 \\ 5 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \cdot 1 + 3 \cdot 1 + (-4) \cdot 1 \\ 5 \cdot 1 + 1 \cdot 1 + 2 \cdot 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$$

These methods *are* nearly identical, but there are some *minor* nuances.

■

5. Write the system first as a vector equation and then as a matrix equation. (Do not solve it)

$$\begin{array}{rcrcrcr} 3x_1 & + & x_2 & - & 5x_3 & = & 9 \\ & & x_2 & + & 4x_3 & = & 0 \end{array}$$

Solution.

$$x_1 \begin{bmatrix} 3 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} -5 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ 0 \end{bmatrix}$$

$$\iff \begin{bmatrix} 3 & 1 & -5 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 0 \end{bmatrix}$$

$$\iff \left[\begin{array}{ccc|c} 3 & 1 & -5 & 9 \\ 0 & 1 & 4 & 0 \end{array}\right]$$

Technically, these are three different forms, but they are all equivalent.

■

6. List five vectors in span $\{\mathbf{v}_1, \mathbf{v}_2\}$. For each vector, show the weights on $\mathbf{v}_1$ and $\mathbf{v}_2$ used to generate the vector and list three entries of the vector. Don't make a sketch.

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -5 \\ 3 \\ 0 \end{bmatrix}$$

Hint: This is easier than it sounds. Just pick some constants!

Solution. Remember the span of a set of vectors is just all the linear combinations. So

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2\} = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 : c_1, c_2 \in \mathbb{R}\}$$

Which is the mathematician's way of saying all the vectors of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2$$

where we just pick some constants c_1, c_2. So let's pick some nice constants!

$$1\mathbf{v}_1 + 0\mathbf{v}_2 = \mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix} \qquad 0\mathbf{v}_1 + 1\mathbf{v}_2 = \mathbf{v}_2 = \begin{bmatrix} -5 \\ 3 \\ 0 \end{bmatrix}$$

$$0\mathbf{v}_1 + 0\mathbf{v}_2 = 0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad -1\mathbf{v}_1 + 0\mathbf{v}_2 = -\mathbf{v}_1 = \begin{bmatrix} -3 \\ -1 \\ 2 \end{bmatrix}$$

$$0\mathbf{v}_1 + (-1)\mathbf{v}_2 = -\mathbf{v}_2 = \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix} \qquad 1\mathbf{v}_1 + 1\mathbf{v}_2 = \begin{bmatrix} -2 \\ 4 \\ -2 \end{bmatrix}$$

■

Remark 2. An astute student made the insightful observation that this process is equivalent to finding random vectors on the plane generated (spanned) by $\mathbf{v}_1$ and $\mathbf{v}_2$. If you are so inclined to see that this is true, you can go to https://www.geogebra.org/3d and plot some of these points and the plane

$$3x + 5y + 7z = 0$$

which is the equation for the plane generated by $\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$ and $\begin{bmatrix} -5 \\ 3 \\ 0 \end{bmatrix}$.

You can also verify that

$$\begin{bmatrix} x = 3 \\ y = 1 \\ z = -2 \end{bmatrix}, \quad \begin{bmatrix} x = -5 \\ y = 3 \\ z = 0 \end{bmatrix}$$

both satisfy the equation $3x + 5y + 7z = 0$.

You don't need to know this, but if you are interested in how to get this equation, here are two methods:

(a) The cross product:

$$(3,1,-2)\times(-5,3,0)=(6,10,14)$$

$$\implies 6x+10y+14z=0 \implies 3x+5y+7z=0$$

(b) Attempting to solve for a,b,c such that

$$ax+by+cz=0$$

for both

$$\begin{bmatrix} x=3 \\ y=1 \\ z=-2 \end{bmatrix}, \quad \begin{bmatrix} x=-5 \\ y=3 \\ z=0 \end{bmatrix}$$

This gives the equations

$$\begin{matrix} 3a+b-2c=0 \\ -5a+3b+0c=0 \end{matrix} \iff \left[\begin{array}{ccc|c} 3 & 1 & -2 & 0 \\ -5 & 3 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|c} 1 & 0 & -\frac{3}{7} & 0 \\ 0 & 1 & -\frac{5}{7} & 0 \end{array}\right] \implies \begin{bmatrix} a=3t \\ b=5t \\ c=7t \end{bmatrix}$$

Pick $t=1$ to get $a=3, b=5, c=7$.

The benefit to this method is that it works for any sized vectors (under the conditions where the span can be described by a single equation), where the cross product only works for this purpose in $\mathbb{R}^3$.

2.2 Eigenvector introduction

7. In linear algebra, one of the *most important* types of equations are of the form

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

For example,

$$\begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x \begin{bmatrix} 2 \\ -3 \end{bmatrix} + y \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}$$

for some number λ. Most linear algebra applications are actually related to equations of this form.

(a) Let $\lambda = 1$, so that the equation is

$$x \begin{bmatrix} 2 \\ -3 \end{bmatrix} + y \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

Rearrange this equation so that it's in the regular vector form

$$x\mathbf{a}_1 + y\mathbf{a}_2 = \mathbf{0}$$

and then solve it. (You should get infinitely many solutions).
Hint: $\begin{bmatrix} x \\ y \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Solution.

$$\begin{aligned} x \begin{bmatrix} 2 \\ -3 \end{bmatrix} + y \begin{bmatrix} -1 \\ 4 \end{bmatrix} - \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \iff x \begin{bmatrix} 2 \\ -3 \end{bmatrix} + y \begin{bmatrix} -1 \\ 4 \end{bmatrix} - x \begin{bmatrix} 1 \\ 0 \end{bmatrix} - y \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \iff x \begin{bmatrix} 1 \\ -3 \end{bmatrix} + y \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

From here, you *can* make this an augmented matrix and solve it, but we can also see that if we let $x = y$, that would definitely give us a solution. Another way to look at it is

$$x\begin{bmatrix}1\\-3\end{bmatrix}+y\begin{bmatrix}-1\\3\end{bmatrix}$$

$$=x\begin{bmatrix}1\\-3\end{bmatrix}-y\begin{bmatrix}1\\-3\end{bmatrix}$$

$$=(x-y)\begin{bmatrix}1\\-3\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

Again, the only way this is true is if $x = y$. That tells us our solutions should be

$$\begin{bmatrix}x=t\\y=t\end{bmatrix}$$

with one particular solution being $\begin{bmatrix}x=1\\y=1\end{bmatrix}$.

Let's also verify our work:

$$t\begin{bmatrix}2\\-3\end{bmatrix}+t\begin{bmatrix}-1\\4\end{bmatrix}=1\begin{bmatrix}t\\t\end{bmatrix}\quad\checkmark$$

$$1\begin{bmatrix}2\\-3\end{bmatrix}+1\begin{bmatrix}-1\\4\end{bmatrix}=1\begin{bmatrix}1\\1\end{bmatrix}\quad\checkmark$$

So this is actually a valid solution! ■

(b) Pick some nonzero solution $\mathbf{v}$ (set $t = 1$ or something) and verify that

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

(If multiplying $A\mathbf{v}$ does not give back $\mathbf{v}$, you might want to check your work for the previous part)

Explain why

$$A^{99999999999999}\mathbf{v}=\mathbf{v}$$

(Do *not* actually calculate $A^{99999999999999}$).

Solution. Let's pick our nice particular solution $\mathbf{v} = \begin{bmatrix}1\\1\end{bmatrix}$.

$$A\mathbf{v} = \begin{bmatrix}2 & -1\\-3 & 4\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = 1\begin{bmatrix}2\\-3\end{bmatrix} + 1\begin{bmatrix}-1\\4\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix}$$

Exactly as we saw in the previous part. We can also check with the row-vector rule:

$$\begin{bmatrix}2 & -1\\-3 & 4\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}2(1) + (-1)(1)\\-3(1) + 4(1)\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix} \quad \checkmark$$

So, why is $A^{99999999999999}\mathbf{v} = \mathbf{v}$? Well, if A doesn't change $\mathbf{v}$, then doing it multiple times won't change it still. Another way to look at it is

$$\mathbf{v} = A\mathbf{v} = A^2\mathbf{v} = A^3\mathbf{v} = \dots$$

■

(c) Do what you did in part 1 for $\lambda = 2$.

$$x\begin{bmatrix}2\\-3\end{bmatrix} + y\begin{bmatrix}-1\\4\end{bmatrix} = \begin{bmatrix}2x\\2y\end{bmatrix}$$

How many solutions do you get this time?

Solution.

$$\begin{aligned}
x\begin{bmatrix}2\\-3\end{bmatrix} + y\begin{bmatrix}-1\\4\end{bmatrix} - \begin{bmatrix}2x\\2y\end{bmatrix} &= \begin{bmatrix}0\\0\end{bmatrix}\\
\iff x\begin{bmatrix}2\\-3\end{bmatrix} + y\begin{bmatrix}-1\\4\end{bmatrix} - x\begin{bmatrix}2\\0\end{bmatrix} - y\begin{bmatrix}0\\2\end{bmatrix} &= \begin{bmatrix}0\\0\end{bmatrix}\\
\iff x\begin{bmatrix}0\\-3\end{bmatrix} + y\begin{bmatrix}-1\\2\end{bmatrix} &= \begin{bmatrix}0\\0\end{bmatrix}
\end{aligned}$$

In equation form, this says

$$\begin{array}{rcrcl} & & -y & = & 0 \\ -3x & + & 2y & = & 0 \end{array}$$

The first equation tells us $y = 0$, and then plugging that into the second equation tells us $x = 0$. Therefore, we only get the trivial solution, and this is very boring. That is, there is only <u>one</u> solution: and it's the zero solution. ■

Remark 3. One observation to make is that we have essentially shown that we *can* find a nonzero vector $\mathbf{v}$ such that

$$A\mathbf{v} = 1\mathbf{v}$$

(such as $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$) but there is *not* a nonzero vector such that

$$A\mathbf{v} = 2\mathbf{v}$$

This means that 1 is a very special number associated with A, but 2 is not. You will talk more extensively about what's going on here, and how to find these special numbers, later in the course.

Remark 4. An astute student also noticed the similarities of the equation

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

to the Lagrange multiplier equation

$$\nabla f = \lambda \nabla g$$

And though the Lagrange multiplier method is more general (since it doesn't have to be linear), there are actually cases where a Lagrange multiplier problem turns into the former.

One example that comes up very often (and is actually the basis of data/image compression and principle component analysis) is trying to maximize the magnitude of the output of multiplying by A. That is, maximizing $\|A\mathbf{x}\|^2$, while restricting ourselves to unit vectors $\|\mathbf{x}\| = 1$ (without this restriction, we wouldn't be able to find an answer since we could just make the vectors super big). This results in the equation

$$A^T A\mathbf{x} = \lambda\mathbf{x} \tag{1}$$

You haven't yet learned about what A^T means (and this class doesn't cover magnitudes or inner products, unfortunately), but later on you can research principle component analysis (PCA), singular value decomposition (SVD), and how linear algebra is used in image compression, if you are interested.

2.3 Polynomial Interpolation / Least Squares

8. In statistics and many other fields, we often want a polynomial that fits data points. Let's consider a very simple case where we just have two data points.

$$(1, 1), \quad (3, 5)$$

(a) Using your basic algebra knowledge, find the equation of a line that passes through these two points (use the point-slope formula $y = y_1 + m(x - x_1)$).

Solution. $m = \frac{5-1}{3-1} = 2$

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

Writing it as $-1 + 2x$ will actually help us keep things consistent for other parts. ■

(b) In linear algebra, we can view this from a different perspective. We suppose that our polynomial is of the form

$$p(x) = a_0 + a_1 x$$

with one variable for each data point. Show that these data points give us a system of equations

$$\begin{array}{ccccc} a_0 & + & 1a_1 & = & 1 \\ a_0 & + & 3a_1 & = & 5 \end{array}$$

and write it as a vector equation.
Hint: What are $p(1)$ and $p(3)$?

Solution. To find $p(1)$ or $p(3)$ we just plug in 1 or 3 for x.

$$p(1) = a_0 + 1a_1$$
$$p(3) = a_0 + 3a_1$$

If we want this to pass through the point $(1, 1)$, that precisely means $p(1) = 1$. So that gives us

$$a_0 + 1a_1 = 1$$

Similarly, to pass through $(3, 5)$, we need $p(3) = 5$, which gives us the second equation. ■

(c) Solve the system and show it gives you the same line you computed in the first part.

Solution. The augmented matrix is

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

Thus, $a_0 = -1$ and $a_1 = 2$. So

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

Exactly as we calculated in the previous part. ■

(d) If we add another data point $(4,4)$ (which adds a new equation $a_0+4a_1=4$) show that there is no solution to the system (implying there is no $p(x)$ of the form a_0+a_1x that passes through those three points).

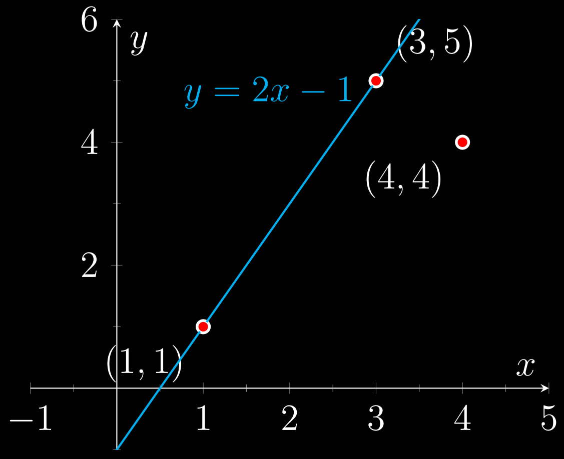
Solution. Adding the third equation

$$\begin{array}{ccccc} a_0 & + & 1a_1 & = & 1 \\ a_0 & + & 3a_1 & = & 5 \\ a_0 & + & 4a_1 & = & 4 \end{array}$$

Row reducing the augmented matrix results in

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

Therefore, we get no solution. This is hopefully intuitive, because we can't draw a straight line between these three points.



■

(e) If we add an x^2 term

$$p(x)=a_0+a_1x+a_2x^2$$

Explain why

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 3 & 9 & 5 \\ 1 & 4 & 16 & 4 \end{array}\right]$$

is the system of equations that correspond to the constraints:

$$p(1) = 1, \quad p(3) = 5, \quad p(4) = 4$$

Also, explain why these constraints are equivalent to passing through the data points $(1,1), (3,5), (4,4)$

Solution. $p(x) = a_0 + a_1 x + a_2 x^2$ means that $p(1) = 1$ becomes

$$p(1) = a_0 + 1a_1 + 1^2 a_2 = 1$$

which gives us the first equation of the given augmented matrix. The exact same logic gives us the rows

$$\left[\begin{array}{ccc|c} 1 & 3 & 3^2 & 5 \\ 1 & 4 & 4^2 & 4 \end{array}\right]$$

■

(f) The REF of the matrix is

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

write the polynomial $p(x)$. Compute $p(1), p(3), p(4)$ to verify that it passes through the points

$$(1,1), \quad (3,5), \quad (4,4)$$

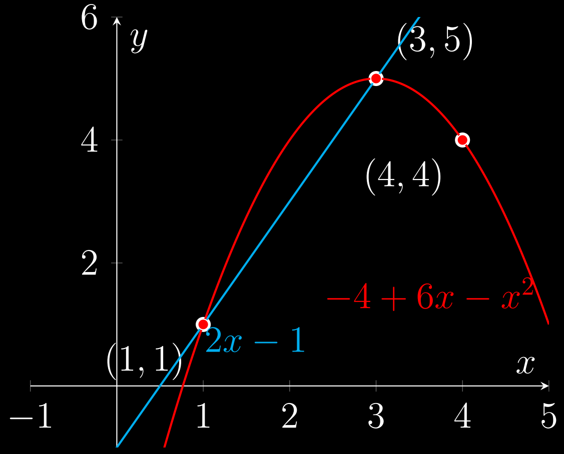
I recommend you plot the quadratic in Desmos and verify (just by looking) that it passes through the points as well.

Solution. This REF matrix tells us $a_0 = -4$, $a_1 = 6$, and $a_2 = -1$. Thus,

$$p(x) = -4 + 6x - x^2$$

We can also see

$$\begin{aligned} p(1) &= -4 + 6(1) - (1)^2 = 1 \quad \checkmark \\ p(3) &= -4 + 6(3) - (3)^2 = 5 \quad \checkmark \\ p(4) &= -4 + 6(4) - (4)^2 = 4 \quad \checkmark \end{aligned}$$



∎

2.3.1 Least Squares

(g) It's often incredibly important and useful in application to find the line of best fit (that is, the linear polynomial that gets "closest" to the data points). You can actually use linear algebra to compute it. We won't get into that in this class, but I'll show you how it's done (without getting into *why* it works).

The system you computed in a previous part was

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 3 & 5 \\ 1 & 4 & 4 \end{array}\right]$$

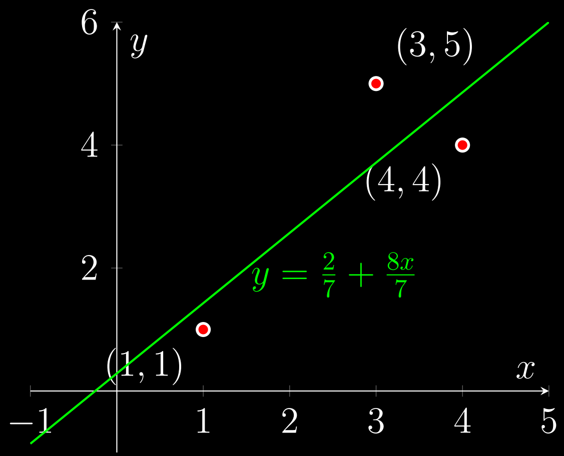
This system didn't have a solution, but we can find the "closest" solution.

What we do is multiply by the coefficient matrix turned on its side (what is called the transpose).

Remark 5. This "transpose" is the A^T found in equation (1).

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 3 & 5 \\ 1 & 4 & 4 \end{array}\right] = \left[\begin{array}{cc|c} 3 & 8 & 10 \\ 8 & 26 & 32 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 2/7 \\ 0 & 1 & 8/7 \end{array}\right]$$

All you need to do for this part is graph the line $p(x) = \frac{2}{7} + \frac{8}{7}x$ alongside the data points and just verify that it gets pretty close. Just plot it on Desmos and give a very quick rough sketch.



Pretty close, huh?

Note: Notice that we "happened" to get a consistent system. It turns out multiplying by the matrix turned on its side actually always gives you a consistent system (which is actually the "closest" solution possible!). The reasons for why this work are unfortunately quite advanced, but if you are interested, you can research "Least squares".

(h) Use your calculus knowledge to explain why

$$\left[\begin{array}{cccc|c} 1 & x_1 & x_1^2 & x_1^3 & y_1 \\ 0 & 1 & 2x_1 & 3x_1^2 & 0 \\ 1 & x_2 & x_2^2 & x_2^3 & y_2 \\ 0 & 1 & 2x_2 & 3x_2^2 & 0 \end{array}\right]$$

is the system that corresponds to finding a cubic polynomial that satisfies the constrains

$$p(x_1) = y_1, \quad p'(x_1) = 0, \quad p(x_2) = y_2, \quad p'(x_2) = 0$$

*Do **not** solve it. Just explain how the system corresponds to the constraints.*
Hint: Suppose

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$$

and consider $p'(x_1)$ and $p'(x_2)$.

Solution. The first and third row are just the degree three polynomial version of what we did in part (e).

$$p(x_1) = a_0 + a_1 x_1 + a_2 x_1^2 + a_3 x_1^3 = y_1 \implies \left[\begin{array}{cccc|c} 1 & x_1 & x_1^2 & x_1^3 & y_1 \end{array}\right]$$

and similar for x_2. The second and fourth row are the weird ones that look different from what we've had so far. But let's consider the constraints on $p'(x)$. (Note: it's often a good strategy in mathematics to look at your givens and simply write what they *mean*)

$$\begin{aligned} p'(x) &= a_1 + 2a_2 x + 3a_3 x^2 \\ &\implies \begin{cases} p'(x_1) = a_1 + 2a_2 x_1 + 3a_3 x_1^2 = 0 \\ p'(x_2) = a_1 + 2a_2 x_2 + 3a_3 x_2^2 = 0 \end{cases} \\ &\implies \left[\begin{array}{cccc|c} 0 & 1 & 2x_1 & 3x_1^2 & 0 \\ 0 & 1 & 2x_2 & 3x_2^2 & 0 \end{array}\right] \end{aligned}$$

Therefore, combining the rows/equations together gives us the augmented matrix we were given. ■

Optional Problems

9. Describe all solutions of $Ax = 0$ in parametric vector form, where A is row equivalent to the given matrix.

$$(a)\quad A = \begin{bmatrix} 1 & -2 & -9 & 5 \\ 0 & 1 & 2 & -6 \end{bmatrix}$$

$$(b)\quad A = \begin{bmatrix} 3 & -9 & 6 \\ -1 & 3 & -2 \end{bmatrix}$$

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

Solution. (a)

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

$$\begin{bmatrix} x_1 = 5t + 7s \\ x_2 = -2t + 6s \\ x_3 = t \\ x_4 = s \end{bmatrix} \implies \mathbf{x} = t \begin{bmatrix} 5 \\ -2 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 7 \\ 6 \\ 0 \\ 1 \end{bmatrix}$$

Remark 6. From the REF, we can see

- 5(Col 1)-2(Col 2)+(Col 3)=0 $\implies \begin{bmatrix} 5 \\ -2 \\ 1 \\ 0 \end{bmatrix}$
- 7(Col 1)+6(Col 2)+(Col 4)=0 $\implies \begin{bmatrix} 7 \\ 6 \\ 0 \\ 1 \end{bmatrix}$

This method requires less writing so, I'm just going to use it since these are optional problems.

(b)

$$\begin{bmatrix} 3 & -9 & 6 \\ -1 & 3 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

We can do 2(Col 1)+3(Col 2)=0 and -2(Col 1)+(Col 3)=0 so that tells us

$$\mathbf{x} = t \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

(a)

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

- -5(Col 1)+(Col 2)=0 $\Longrightarrow \begin{bmatrix} -5 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$
- -8(Col 1)+7(Col 3)+(Col 4)=0 $\Longrightarrow \begin{bmatrix} -8 \\ 0 \\ 7 \\ 1 \\ 0 \\ 0 \end{bmatrix}$

- -1(Col 1)-4(Col 3)+(Col 5)=0 $\implies \begin{bmatrix} -1 \\ 0 \\ -4 \\ 0 \\ 1 \\ 0 \end{bmatrix}$

$$\mathbf{x} = t_1 \begin{bmatrix} -5 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -8 \\ 0 \\ 7 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_3 \begin{bmatrix} -1 \\ 0 \\ -4 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

■

10. Given A and $\mathbf{b}$, write the augmented matrix for the linear system that corresponds to the matrix equation $A\mathbf{x} = \mathbf{b}$. Then solve the system and write the solution as a vector.

$$A = \begin{bmatrix} 1 & 2 & 1 \\ -3 & -1 & 2 \\ 0 & 5 & 3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

Solution.

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ -3 & -1 & 2 & 1 \\ 0 & 5 & 3 & -1 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{3}{5} \\ 0 & 1 & 0 & -\frac{4}{5} \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\implies \begin{bmatrix} x_1 = \frac{3}{5} \\ x_2 = -\frac{4}{5} \\ x_3 = 1 \end{bmatrix} \implies \mathbf{x} = \begin{bmatrix} \frac{3}{5} \\ -\frac{4}{5} \\ 1 \end{bmatrix}$$

■

Week #3 Worksheet Solutions

3.1 Linear Independence and Span

1. Find four vectors in $\mathbb{R}^2$, $\mathbf{a}_1, \mathbf{a}_2, \mathbf{b}_1, \mathbf{b}_2$ such that

   - $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}_1$ is inconsistent
   - $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}_2$ is consistent.

   And then (without doing any more work) find one 2×2 matrix A such that

   - $A\mathbf{x} = \mathbf{b}_1$ is inconsistent
   - $A\mathbf{x} = \mathbf{b}_2$ is consistent.

   Solution. The tricky part about this problem is that in order for the system to possibly be inconsistent, the $\mathbf{a}$ vectors have to be linearly dependent. Otherwise, the $\mathbf{a}$ vectors will span $\mathbb{R}^2$ and every system will be consistent.

   I'll pick the absolute easiest solution I can think of (that isn't completely trivial).

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\mathbf{b}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \mathbf{b}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

   Then we have

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The first is clearly inconsistent and a homogeneous system is always consistent.

To satisfy the other part, we just use the **a** vectors as the columns of A.

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Another totally valid solution: $\mathbf{a}_1 = \mathbf{a}_2 = \mathbf{b}_2 = \vec{0}$ and $\mathbf{b}_1 \neq \vec{0}$ would totally work. ■

2. Determine if the vectors are linearly independent. Justify each answer.

$$(a) \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \begin{bmatrix} -3 \\ 9 \end{bmatrix} \qquad (b) \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ -8 \end{bmatrix}, \begin{bmatrix} -3 \\ 4 \\ 1 \end{bmatrix}$$

$$(c) \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 7 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ -8 \end{bmatrix} \qquad (d) \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \begin{bmatrix} -2 \\ -8 \end{bmatrix}$$

Solution. Note that there are *many* ways to determine linear (in)dependence. I won't write out every single way. Your method could totally be valid.

(a) We can see that $\mathbf{v}_2 = -3\mathbf{v}_1$, so they're dependent.

(b)

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

We get a sort of triangular form here. If we were to have some linear combination to get zero, the first row of the system we'd get would imply that $c_3 = 0$. Substituting that into the second equation would give $c_2 = 0$ etc. Therefore, they're independent. This is similar to a problem on the first worksheet.

(c)

$$\begin{bmatrix}5\\0\\0\end{bmatrix}, \begin{bmatrix}7\\2\\-4\end{bmatrix}, \begin{bmatrix}9\\4\\-8\end{bmatrix}$$

If we were to try to get a linear combination of these vectors to get zero, then we can always choose the constant c_1 on the first vector to cancel out whatever is in the first entry. Therefore, we just need to focus on the second and third rows of the second and third vector.

$$\begin{bmatrix}2\\-4\end{bmatrix}, \begin{bmatrix}4\\-8\end{bmatrix}$$

We can see that $c_2 = -2c_3$ will cancel out the second and third entries.

$$c_1\begin{bmatrix}5\\0\\0\end{bmatrix} - 2c_3\begin{bmatrix}7\\2\\-4\end{bmatrix} + c_3\begin{bmatrix}9\\4\\-8\end{bmatrix} = c_1\begin{bmatrix}5\\0\\0\end{bmatrix} + c_3\begin{bmatrix}-5\\0\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

Picking $c_1 = 1$ and $c_3 = 1$ will work, so there's a nontrivial linear combination. Therefore, they are dependent, since there are only two vectors.

(d)

$$\begin{bmatrix}-1\\4\end{bmatrix}, \begin{bmatrix}-2\\-8\end{bmatrix}$$

Here, we can see that neither vector is a scalar multiple of the other. That is enough to say they're independent.

■

3. Determine if the columns of the matrix form a linearly independent set. Then determine if the function $T(\mathbf{x}) = A\mathbf{x}$ is injective. Justify each answer.

$$(a)\begin{bmatrix} -4 & -3 & 0 \\ 0 & -1 & 4 \\ 1 & 0 & 3 \\ 5 & 4 & 6 \end{bmatrix} \qquad (b)\begin{bmatrix} 1 & -3 & 3 & -2 \\ -3 & 7 & -1 & 2 \\ 0 & 1 & -4 & 3 \end{bmatrix}$$

Hint: Part (b) requires no scratch work.

Solution. Recall that the function $A\mathbf{x}$ is injective exactly when the columns of A are linearly independent. Therefore, the answer for independence will be the same as the answer for injectivity.

(a) For this matrix, there aren't any clear signs that this is independent. However, we can see that if we take 4 of column 2, we'll need -3 of column 1 and 1 of column 3 to cancel out the first two entries.

$$-3\begin{bmatrix} -4 \\ 0 \\ 1 \\ 5 \end{bmatrix} + 4\begin{bmatrix} -3 \\ -1 \\ 0 \\ 4 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 4 \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 7 \end{bmatrix} \neq \vec{0}$$

So it is independent and injective.

However, if you couldn't see that immediately (that the first and third column would make it easy to cancel out entries of the remaining second column), then we can row reduce.

If we swap rows 1 and 3 (and multiply row 2 by -1), we get

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -4 \\ -4 & -3 & 0 \\ 5 & 4 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -4 \\ 0 & -3 & 12 \\ 0 & 4 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

We can see now that we'll get a pivot in every column

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

so the vectors are independent. Theerefore, the function $A\mathbf{x}$ will also be injective.

(b)

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

Here we don't need to do any work. There are too many columns for every column to be a pivot column! With only three rows, we can only have three pivots maximum. With four variables, that means we have at least one free variable.

Someone asked in discussion "which is the dependent column?" In general, that is a very good question to ask, and a good thing to know. But for this problem, the answer is: "I don't know and I don't care!" If all we need to know is if they are independent, then the fact that there is a free variable is all I need to know.

∎

Remark 7. This logic implies that if you have more than m vectors from $\mathbb{R}^m$, then they are *automatically linearly dependent.*

4. For what values of h is
   (i) $\mathbf{v}_3$ in Span$\{\mathbf{v}_1, \mathbf{v}_2\}$?
   (ii) $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = \mathbf{v}_3$ consistent?
   (iii) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ linearly dependent? Justify each answer.

$$(a) \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 9 \\ -6 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 5 \\ -15 \\ h \end{bmatrix}$$

$$(b) \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ -5 \\ -3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 10 \\ 7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ -9 \\ h \end{bmatrix}$$

Solution. First, a remark: (ii) is exactly equivalent to (i). So whatever we say for (i) is also the answer for (ii).

(a) Notice that $\mathbf{v}_2 = -3\mathbf{v}_1$. So, this set will always be linearly dependent.

   (i) We can notice that in both $\mathbf{v}_1$ and $\mathbf{v}_2$, the second entry is -3 times the first, and the third entry is 2 times the first. This property will hold for any linear combination, so if $\mathbf{v}_3$ is in the span (it is a linear combination), then the third entry must be twice the first. That is,

   $$h = 2(5) = 10$$

   Notice that this is only possible because the second entry is $-15 = -3(5)$. If this wasn't the case, $\mathbf{v}_3$ couldn't possibly be in the span.

   (ii) Same for (i), as stated above.

   (iii) As we also said above, since $\mathbf{v}_2 = -3\mathbf{v}_1$, any set of vectors containing both $\mathbf{v}_1, \mathbf{v}_2$ will always be linearly dependent. It doesn't matter what the other vectors are.
   This confused some students, but basically, being linearly independent is a property about a group of vectors saying that none of them are combinations of the others. If any one of them *is* a combination of any of the others, then the whole group is tainted. And I can add whatever vectors I want. The group is still tainted.

(b)

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -5 \\ -3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 10 \\ 7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ -9 \\ h \end{bmatrix}$$

We find a similar behavior from the previous part that the second entry is -5 times the first entry. This will also be preserved through linear combinations.

(i) Unfortunately, $\mathbf{v}_3$ doesn't have this property that the second entry is -5 the first, so there's no way for there to be a linear combination of $\mathbf{v}_1, \mathbf{v}_2$ to get $\mathbf{v}_3$. This is actually sufficient for the problem.
You would also see this if you were trying to solve the system in (ii):

$$\left[\begin{array}{cc|c} 1 & -2 & 2 \\ -5 & 10 & -9 \\ -3 & 7 & h \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -2 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & h+6 \end{array}\right]$$

This is inconsistent.

(ii) Same for (i), as stated above.

(iii) At this point, we can conclude that since $\mathbf{v}_1$ and $\mathbf{v}_2$ are not scalar multiples of each other (meaning they are independent), and $\mathbf{v}_3$ is independent from $\{\mathbf{v}_1, \mathbf{v}_2\}$, then the whole set of vectors must be independent always. Though, I am not sure you learned in lecture that adjoining an independent vector to an independent set results in an independent set.
Examining the reduced matrix (though it's not in EF or REF)

$$\left[\begin{array}{cc|c} 1 & -2 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & h+6 \end{array}\right]$$

we can spot that we'll have three pivots:

$$\left[\begin{array}{cc|c} \boxed{1} & -2 & 2 \\ 0 & 0 & \boxed{1} \\ 0 & \boxed{1} & h+6 \end{array}\right]$$

Since every column is a pivot column, they are linearly independent. In this case, it doesn't matter what h is, so there is no h such that the set is dependent.

■

3.2 Determinant Introduction

5. Consider the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and the homogeneous system

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2}$$

(a) Explain why if there exists a nontrivial solution to (2), then that directly implies the columns of A, $\left\{ \begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix} \right\}$ must be linearly dependent.

Solution. This is basically just definition. If there's a nontrivial solution $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, then we're saying for at least one x_1, x_2 nonzero

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} a \\ c \end{bmatrix} + x_2 \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Meaning there is a nontrivial linear combination that yields zero. That's the definition of linear dependence! ■

(b) Show that if the columns of A, $\left\{\begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}\right\}$, are linearly dependent, then $ad - bc = 0$.

Hint: Suppose for k_1, k_2 not both zero that

$$k_1 \begin{bmatrix} a \\ c \end{bmatrix} + k_2 \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \begin{cases} ak_1 + bk_2 = 0 \\ ck_1 + dk_2 = 0 \end{cases}$$

Use the fact that both

$$\begin{cases} d(ak_1 + bk_2) - b(ck_1 + dk_2) = d(0) - b(0) = 0 \\ -c(ak_1 + bk_2) + a(ck_1 + dk_2) = -c(0) + a(0) = 0 \end{cases}$$

must be satisfied.

<u>*Solution.*</u> If the vectors are dependent, then we can find k_1, k_2 not both zero such that

$$k_1 \begin{bmatrix} a \\ c \end{bmatrix} + k_2 \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \begin{cases} ak_1 + bk_2 = 0 \\ ck_1 + dk_2 = 0 \end{cases}$$

Examining the two equations in the hint:

$$\begin{cases} d(ak_1 + \cancel{bk_2}) - b(ck_1 + \cancel{dk_2}) = k_1(ad - bc) = 0 \\ -c(\cancel{ak_1} + bk_2) + a(\cancel{ck_1} + dk_2) = k_2(ad - bc) = 0 \end{cases}$$

Well, if at least one of the k's is nonzero (say, without loss of generality, it was k_1), then

$$k_1(ad - bc) = 0 \implies ad - bc = 0$$

The only way two (real) numbers can multiply to zero is if at least one of them is zero.

Similarly if k_2 is nonzero the same thing is implied. Therefore, we've shown that dependent columns implied $ad - bc = 0$. ■

(c) Show that if $ad - bc = 0$, then both $\begin{bmatrix} d \\ -c \end{bmatrix}$ and $\begin{bmatrix} -b \\ a \end{bmatrix}$ are homogeneous solutions to

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Solution. To be homogeneous solutions, that just means that plugging them in should give us zero! So let's check that.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d \\ -c \end{bmatrix} = \begin{bmatrix} ad - bc \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \checkmark$$
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} -b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ ad - bc \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \checkmark$$

Again, these are zero *because* $ad - bc = 0$. If it was nonzero, this wouldn't work. (Note this is explored in (4.3)) ■

(d) Assume that a, b, c, d are not *all* zero. This means that at least one of $\begin{bmatrix} d \\ -c \end{bmatrix}$ or $\begin{bmatrix} -b \\ a \end{bmatrix}$ are nonzero vectors. Explain how the the previous part implies that if $ad - bc = 0$, then there is a nontrivial homogeneous solution.

Solution. The only way for both $\begin{bmatrix} d \\ -c \end{bmatrix}$ and $\begin{bmatrix} -b \\ a \end{bmatrix}$ to be the zero vectors are if $a = b = c = d = 0$. Therefore, at least one of these vectors is nonzero. Since we showed previously that they are homogeneous solutions, at least one of these vetcors is a nonzero (nontrivial) homogeneous solution. ■

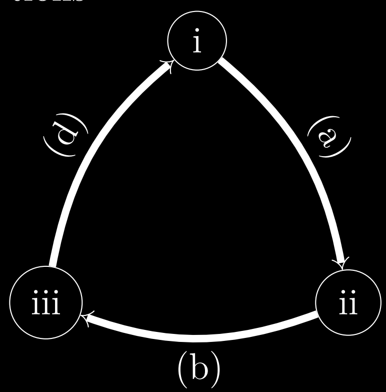
(e) Based on all these parts together, explain how we have showed that (for a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$) the following are equivalent

i. There exists a nontrivial homogeneous solution

ii. The columns are linearly dependent

iii. $ad - bc = 0$

Solution. Essentially we have created a sort of circle of implications



Therefore, either all of them are true, or none of them are true. ■

(f) Explain how knowing $ad - bc \neq 0$ implies

i. The only homogeneous solution is the trivial solution.

ii. The columns are linearly independent

Remark 8. Eventually, we will show that $ad - bc \neq 0$ is equivalent to $A\mathbf{x}$ being injective and surjective. But it turns out these two things are equivalent to being at least injective.

Solution. Like I said in the previous part, either all of them are true, or none of them are true (they are equivalent after all!). Therefore, if iii is false, then the other properties must also be false. What we have here are just the opposites. ■

Week #4 Worksheet Solutions

4.1 Linear Transformations (Injectivity / Surjectivity)

1. Define the linear transformation

$$T(\mathbf{x}) = \begin{bmatrix} 1 & 1 & 1 & 3 & -1 \\ 1 & 2 & 2 & 5 & 0 \\ 1 & 2 & 3 & 6 & 0 \end{bmatrix} \mathbf{x}$$

   (a) For what α and β is T a function $T\colon \mathbb{R}^\alpha \to \mathbb{R}^\beta$?

   Solution. This is a 3×5 matrix. Thus, for $A\mathbf{x}$ to be defined, $\mathbf{x}$ needs to have 5 rows. That is, the inputs to T need to be in $\mathbb{R}^5$. The outputs will be linear combinations of the columns of A, which have three rows. Thus, the outputs are in $\mathbb{R}^3$, so

$$T\colon \mathbb{R}^5 \to \mathbb{R}^3$$

   In general, an $m \times n$ matrix defines a transformation $\mathbb{R}^n \to \mathbb{R}^m$. ■

   (b) Determine if T is injective, surjective, neither, or both.

   Solution. This matrix has too many columns to have a pivot in every column, so it's definitely *not* injective.
   After performing $R_2 - R_1$ and $R_3 - R_2$ we get three pivots

$$\begin{bmatrix} \boxed{1} & 1 & 1 & 3 & -1 \\ 0 & \boxed{1} & 1 & * & * \\ 0 & 0 & \boxed{1} & * & * \end{bmatrix}$$

   (the $*$ are entries we don't care about). The important thing we're looking for is that we get a pivot in every row. Since we do have that, this transformation is surjective. ■

2. Assume that T is a linear transformation. Find the standard matrix of T.

   (a)
   $$T\colon \mathbb{R}^2 \to \mathbb{R}^4, \quad T(\mathbf{e}_1) = (3,1,3,1), \quad T(\mathbf{e}_2) = (-5,2,0,0)$$
   where $\mathbf{e}_1 = (1,0)$ and $\mathbf{e}_2 = (0,1)$.

   (b)
   $$T\colon \mathbb{R}^3 \to \mathbb{R}^2, \quad T(\mathbf{e}_1) = (1,3), \quad T(\mathbf{e}_2) = (4,-7), \quad T(\mathbf{e}_3) = (-5,4)$$
   where $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are the columns of the 3×3 identity matrix.

   Solution. The standard matrix is just given by placing the images of the standard basis vectors (the $\mathbf{e}$'s) as the columns. The ith column is the image of $\mathbf{e}_i$ ($T(\mathbf{e}_i)$). Thus, the matrices are just

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

   (b) $\begin{bmatrix} 1 & 4 & -5 \\ 3 & -7 & 4 \end{bmatrix}$

   ■

   Remark 9. For how to find the matrix of a transformation where you don't have the images of the standard basis vectors, see (5.1).

3. Decide if T maps $\mathbb{R}^5$ onto $\mathbb{R}^5$ (i.e. if T is surjective).

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

Solution. This has a pivot in every row so it's surjective. ■

(b)

$$\begin{bmatrix} 1 & -5 & 0 & 0 & 25 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Solution. There is a row of zeros, so it's not surjective. ■

4.2 Eigenvectors and Determinants

4. Last week (3.2) we showed that (for a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$) the following are equivalent

(i). The only solution to $A\mathbf{x} = \vec{0}$ is the trivial solution (A has a trivial kernel or $\ker(A) = \{\vec{0}\}$)

(ii). The columns of A are linearly independent

(iii). $ad - bc \neq 0$

The week before (2.2), I gave you the matrix $A = \begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix}$ and we looked for vectors which satisfied $A\mathbf{v} = \lambda\mathbf{v}$. For $\lambda = 1$, we were able to find a nonzero vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, but not for $\lambda = 2$. What we basically did was rearrange the equation to

$$\begin{aligned}
\begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix} \\
\implies x \begin{bmatrix} 2 \\ -3 \end{bmatrix} + y \begin{bmatrix} -1 \\ 4 \end{bmatrix} = x \begin{bmatrix} \lambda \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ \lambda \end{bmatrix} \\
\implies x \begin{bmatrix} 2-\lambda \\ -3 \end{bmatrix} + y \begin{bmatrix} -1 \\ 4-\lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\implies \begin{bmatrix} 2-\lambda & -1 \\ -3 & 4-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\end{aligned}$$

(a) Based on our equivalent statements, we can understand $ad - bc$ as a "nontrivial kernel detector". Calculate $ad - bc$ for the matrix $\begin{bmatrix} 2-\lambda & -1 \\ -3 & 4-\lambda \end{bmatrix}$ (you will get a quadratic equation) and find the values of λ such that $ad - bc = 0$.

Hint: We know that $\lambda = 1$ should work, so that should be one of your solutions. $\lambda = 2$ should also *not* be a solution. We will call the other value λ_1.

Solution.

$$\begin{aligned}
(2-\lambda)(4-\lambda) - (-1)(-3) \\
= \lambda^2 - 6\lambda + 5 \\
= (\lambda - 1)(\lambda - 5) = 0 \\
\implies \lambda = 1, 5
\end{aligned}$$

Like we expected, 1 is a solution, but 2 is not. We also see that $\lambda_1 = 5$ is our other magical value (eigenvalue is the technical term). ■

(b) For the other value λ_1, find a nonzero vector $\mathbf{v}$ in the kernel.

Hint: For consistency with following parts, I suggest picking $\mathbf{v}$ such that the first entry/component is 1.

Solution. Plugging in $\lambda = 5$ we get

$$\begin{bmatrix} -3 & -1 \\ -3 & -1 \end{bmatrix}$$

If we want a nontrivial kernel vector (with a 1 in the first entry), that means we need to find a value such that

$$\begin{bmatrix} -3 & -1 \\ -3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ ? \end{bmatrix} = 1 \begin{bmatrix} -3 \\ -3 \end{bmatrix} + ? \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Clearly $? = -3$ works, so our nontrivial kernel vector is

$$\mathbf{v} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

■

(c) For that nonzero vector $\mathbf{v}$, verify that multiplying by A gives you $\lambda_1 \mathbf{v}$.

Solution.

$$\begin{aligned} A\mathbf{v} &= \begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \end{bmatrix} \\ &= 1 \begin{bmatrix} 2 \\ -3 \end{bmatrix} - 3 \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ -15 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ -3 \end{bmatrix} \quad \checkmark \end{aligned}$$

■

4.2.1 Diagonalization Introduction

(d) Define $T(\mathbf{x}) = \begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix} \mathbf{x}$ (i.e. $T(\mathbf{x}) = A\mathbf{x}$). We showed two weeks ago (2.2) that $T \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Compute

$$T^{100}\left(3\begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 1 \\ -3 \end{bmatrix}\right)$$

Leave your answer as a linear combination.

Solution. Observe that

$$\begin{aligned} T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) &= 1\begin{bmatrix} 1 \\ 1 \end{bmatrix} &\implies T^{100}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) &= 1^{100}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ T\left(\begin{bmatrix} 1 \\ -3 \end{bmatrix}\right) &= 5\begin{bmatrix} 1 \\ -3 \end{bmatrix} &\implies T^{100}\left(\begin{bmatrix} 1 \\ -3 \end{bmatrix}\right) &= 5^{100}\begin{bmatrix} 1 \\ -3 \end{bmatrix} \end{aligned}$$

The defining (and wonderful) property of linear transformations is that they preserve linear combinations. So, since we know the image of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ under T^{100}, we know the image of any linear combination (it's just the linear combination of the images!).

$$\begin{aligned} &T^{100}\left(3\begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 1 \\ -3 \end{bmatrix}\right) \\ &\qquad = 3T^{100}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) - 2T^{100}\left(\begin{bmatrix} 1 \\ -3 \end{bmatrix}\right) \\ &\qquad\qquad = 3\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) - 2\left(5^{100}\begin{bmatrix} 1 \\ -3 \end{bmatrix}\right) \end{aligned}$$

$$\boxed{= 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2 \cdot 5^{100} \begin{bmatrix} 1 \\ -3 \end{bmatrix}}$$

■

(e) What is $A^{100} \left(3 \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ -3 \end{bmatrix} \right)$?

Hint: Do *not* calculate A^{100}. This part requires no further work.

Solution. Applying T is the same as multiplying by A so the answer to this part is exactly the same as as the previous part. ■

4.3 Determinants (Part II)

5. We denote the calculation of $ad - bc$ as the determinant: $\det(A)$. We often shorthand the det by using straight lines for the matrix.

$$ad - bc := \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

It turns out there are a lot more things that the determinant can tell us.

(a) Last week (3.2), we showed that if $ad - bc = 0$, then $\begin{bmatrix} d \\ -c \end{bmatrix}$ and $\begin{bmatrix} -b \\ a \end{bmatrix}$ are homogeneous solutions. Suppose that $ad - bc$ is NOT zero. Find preimages of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Hint: Calculate $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d \\ -c \end{bmatrix}$ and $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} -b \\ a \end{bmatrix}$ and assume $ad - bc \neq 0$.

Solution. We calculated last week (3.2) that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d \\ -c \end{bmatrix} = \begin{bmatrix} ad-bc \\ 0 \end{bmatrix} = (ad-bc) \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} -b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ ad-bc \end{bmatrix} = (ad-bc) \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

We have that $\begin{bmatrix} d \\ -c \end{bmatrix}$ is a preimage of $(ad-bc)\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, but we want the preimage of just $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Well, since linear transformations preserve scalar multiplication, then scaling a preimage scales the image by the exact same amount. So if we scale the preimage by $\frac{1}{ad-bc}$, then that will scale the image by that same amount, giving us what we want! Notice that it's important $ad-bc$ is nonzero here, otherwise we can't divide it.

Hence, a preimage of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is

$$\mathbf{v}_1 = \frac{1}{ad-bc} \begin{bmatrix} d \\ -c \end{bmatrix}$$

since $A\left(\frac{1}{ad-bc}\begin{bmatrix} d \\ -c \end{bmatrix}\right) = \frac{1}{ad-bc} A \begin{bmatrix} d \\ -c \end{bmatrix} = \frac{ad-bc}{ad-bc}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ gives us what we want.

and by the same logic a preimage of $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is

$$\mathbf{v}_2 = \frac{1}{ad-bc} \begin{bmatrix} -b \\ a \end{bmatrix}$$

■

(b) Let's call the preimage of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ you found in the previous part $\mathbf{v}_1$, and the preimage of $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ we will call $\mathbf{v}_2$.

i. Using $\mathbf{v}_1$ and $\mathbf{v}_2$, find a preimage of an arbitrary vector $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$. You must keep it in terms of arbitrary z_1, z_2. You can't set them to specific values. You do not have to write the vector entries explicitly.
Hint: It's going to be some linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

Solution. If $\mathbf{v}_1$ gives us $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2$ gives us $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then we can take z_1 of $\mathbf{v}_1$ and z_2 of $\mathbf{v}_2$ to get an image of $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$.

$$\boxed{z_1\mathbf{v}_1 + z_2\mathbf{v}_2}$$

Another way to think of it is utilizing that the image of a linear combination is the same linear combination of the images.

$$T(c_1\mathbf{w}_1 + \ldots + c_k\mathbf{w}_k) = c_1T(\mathbf{w}_1) + \ldots + c_kT(\mathbf{w}_k)$$

We can read this right to left as a preimage of a linear combination of known images, is that same linear combination of preimages. Thus,

$$\begin{aligned} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} &= z_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + z_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= z_1T(\mathbf{v}_1) + z_2T(\mathbf{v}_2) \\ &= T(z_1\mathbf{v}_1 + z_2\mathbf{v}_2) \end{aligned}$$

So a preimage is $z_1\mathbf{v}_1 + z_2\mathbf{v}_2$. ■

ii. Explain how this implies $T(\mathbf{x}) = A\mathbf{x}$ is surjective if $ad - bc \neq 0$.

Solution. To be surjective means, by definition, that every vector in the codomain has a preimage. Since the codomain is $\mathbb{R}^2$, and we found a preimage of a general vector in $\mathbb{R}^2$ in the previous part, then we have not only shown that every vector has a preimage, but actually *found it explicitly*! ■

(c) You can assume that the following are equivalent.

i. $A\mathbf{x}$ is injective.

ii. If $A\mathbf{x} = \mathbf{b}$ has a solution, that solution is unique.

iii. The only solution to $A\mathbf{x} = \vec{0}$ is the trivial solution.

iv. The columns of A are linearly independent.

v. $\det(A) \neq 0$

vi. The rows of A are linearly independent.

vii. $A\mathbf{x} = \mathbf{b}$ is always consistent.

viii. $A\mathbf{x}$ is surjective.

Based on this, explain why we know that the following system

$$\begin{bmatrix} 23\pi\sqrt{3} & -e^{-7}\ln(2) \\ e^{7}\ln(2) & \pi\sqrt{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \sqrt{5} \\ \tan(1) \end{bmatrix}$$

is consistent *and* has a unique solution (without solving it) based on the calculation

$$\left(23\pi\sqrt{3}\right)\left(\pi\sqrt{3}\right) - \left(-e^{-7}\ln(2)\right)\left(e^{7}\ln(2)\right) = 69\pi^2 + \ln(2)^2 \neq 0$$

Solution. The computation is the determinant, which is clearly nonzero. Thus, we have satisfied v., and by extension all the other properties. Then vii. tells us the system must be consistent, and ii. tells us the solution is unique. ■

4.4 Cramer's Rule

(d) Verify that if $ad - bc \neq 0$, then

$$x_1 = \frac{\det \begin{bmatrix} z_1 & b \\ z_2 & d \end{bmatrix}}{\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}}, \quad x_2 = \frac{\det \begin{bmatrix} a & z_1 \\ c & z_2 \end{bmatrix}}{\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}}$$

is the solution to the system of equations

$$\begin{array}{ccccc} ax_1 & + & bx_2 & = & z_1 \\ cx_1 & + & dx_2 & = & z_2 \end{array}$$

by plugging them into the system.

How does this compare to your answer to part b?

<u>*Solution.*</u>

$$x_1 = \frac{z_1 d - z_2 b}{ad - bc}, \quad x_2 = \frac{z_2 a - z_1 c}{ad - bc}$$

Plugging in,

$$ax_1 + bx_2 = \frac{a(z_1 d - \cancel{z_2 b}) + b(\cancel{z_2 a} - z_1 c)}{ad - bc} = z_1 \frac{ad - bc}{ad - bc} = z_1 \quad \checkmark$$

$$cx_1 + dx_2 = \frac{c(\cancel{z_1 d} - z_2 b) + d(z_2 a - \cancel{z_1 c})}{ad - bc} = z_2 \frac{-bc + ad}{ad - bc} = z_2 \quad \checkmark$$

Notice that

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{z_1 d - z_2 b}{ad - bc} \\ \frac{z_2 a - z_1 c}{ad - bc} \end{bmatrix}$$

$$= z_1 \left(\frac{1}{ad - bc} \begin{bmatrix} d \\ -c \end{bmatrix} \right) + z_2 \left(\frac{1}{ad - bc} \begin{bmatrix} -b \\ a \end{bmatrix} \right)$$

$$= z_1 \mathbf{v}_1 + z_2 \mathbf{v}_2$$

Exactly like part b. ■

Remark 10. This is called Cramer's rule, which gives a general and explicit formula for each variable in a system $A\mathbf{x} = \mathbf{b}$ (where A is square and has a nonzero determinant).

Week #5 Worksheet Solutions

5.1 Linear Transformations Continued

1. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation that maps

$$\vec{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \mapsto T(\mathbf{u}) = \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix}$$

$$\vec{v} = \begin{bmatrix} 2 \\ 7 \end{bmatrix} \mapsto T(\mathbf{v}) = \begin{bmatrix} 9 \\ 2 \\ -5 \end{bmatrix}$$

(a) Use the fact that T is linear to find the images under T of $7\vec{u}$, $-3\vec{v}$, and $7\vec{u} - 3\vec{v}$.

Solution. By linearity

$$T(7\mathbf{u}) = 7T(\mathbf{u}) = \begin{bmatrix} 28 \\ 7 \\ -14 \end{bmatrix}$$

$$T(-3\mathbf{v}) = -3T(\mathbf{u}) = \begin{bmatrix} -27 \\ -6 \\ 15 \end{bmatrix}$$

$$T(7\mathbf{u} - 3\mathbf{v}) = 7T(\mathbf{u}) - 3T(\mathbf{u}) = \begin{bmatrix} 28 \\ 7 \\ -14 \end{bmatrix} + \begin{bmatrix} -27 \\ -6 \\ 15 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

■

(b) Row reduce the matrix $\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 7 & 0 & 1 \end{array}\right]$ into REF

Solution. $\sim \left[\begin{array}{cc|cc} 1 & 0 & 7 & -2 \\ 0 & 1 & -3 & 1 \end{array}\right]$ ■

Remark 11. Notice this is actually computing the inverse of the matrix $\begin{bmatrix}\mathbf{u} & \mathbf{v}\end{bmatrix}$.

(c) Find linear combinations of **u** and **v** to get the standard basis vectors. i.e.

$$c_1 \begin{bmatrix}1\\3\end{bmatrix} + c_2 \begin{bmatrix}2\\7\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}, \quad k_1 \begin{bmatrix}1\\3\end{bmatrix} + k_2 \begin{bmatrix}2\\7\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}$$

You may use the results of part (b).

Solution. The system for the c's, is just the first three columns of the augmented matrix in part (b). $\left[\begin{array}{cc|c} 1 & 0 & 7 \\ 0 & 1 & -3 \end{array}\right]$ So, we just look at the third column to get the solution $c_1 = 7, \quad c_2 = -3$.
Similarly, the fourth column gives us the solution to $\left[\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 1 \end{array}\right]$ which is $k_1 = -2, \quad k_2 = 1$.
We can verify that

$$7\begin{bmatrix}1\\3\end{bmatrix} - 3\begin{bmatrix}2\\7\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}, \quad -2\begin{bmatrix}1\\3\end{bmatrix} + \begin{bmatrix}2\\7\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}$$

■

5.1.1 Finding Standard Matrices

(d) Find the standard matrix for T.

Solution. The standard matrix is just the image of the standard basis vectors. But now that we have linear combinations of **v** and

u to get them, we can just take the same linear combinations of the images.

$$\begin{bmatrix}1\\0\end{bmatrix} = 7\mathbf{u} - 3\mathbf{v}$$

$$\implies T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = T\left(7\mathbf{u} - 3\mathbf{v}\right) = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

as we calculated in part (a). This will be the first column of A. Similarly,

$$\begin{bmatrix}0\\1\end{bmatrix} = -2\mathbf{u} + \mathbf{v}$$

$$\implies T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = -2T\left(\mathbf{u}\right) + T\left(\mathbf{v}\right)$$

$$= -2\begin{bmatrix}4\\1\\-2\end{bmatrix} + \begin{bmatrix}9\\2\\-5\end{bmatrix} = \begin{bmatrix}1\\0\\-1\end{bmatrix}$$

which is the second column. Thus,

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

You can confirm that, indeed,

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

■

(e) Row reduce the matrix $\left[\begin{array}{cc|ccc} 1 & 3 & 4 & 1 & -2 \\ 2 & 7 & 9 & 2 & -5 \end{array}\right]$ into REF and compare it to your result in part (d).

Solution. $\sim \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & -1 \end{array}\right]$ Which just gives us I augmented with the standard matrix (transposed)! This gives us a much faster method to compute the standard matrix.

So, why does this work? Well, it's because linear combinations of the preimages are preserved with the images. So, if we put both $\mathbf{u}, \mathbf{v}$ and the images $T(\mathbf{u}), T(\mathbf{v})$ as a single vector,

$$\begin{bmatrix} \mathbf{u} \\ T(\mathbf{u}) \end{bmatrix}, \quad \begin{bmatrix} \mathbf{v} \\ T(\mathbf{v}) \end{bmatrix}$$

Then, any linear combination will preserve the image

$$c_1 \begin{bmatrix} \mathbf{u} \\ T(\mathbf{u}) \end{bmatrix} + c_2 \begin{bmatrix} \mathbf{v} \\ T(\mathbf{v}) \end{bmatrix} = \begin{bmatrix} c_1\mathbf{u} + c_2\mathbf{v} \\ T(c_1\mathbf{u} + c_2\mathbf{v}) \end{bmatrix}$$

This is actually a good thing to make sure you can understand and justify!

When we row reduce a matrix, we are just taking linear combinations of the rows. And, for square matrices specifically, our goal is usually to try to get the standard basis vectors as the rows (and, this is possible exactly when the matrix is invertible).

Remark 12. This method is actually equivalent to the following method using inverses

$$A \begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix} = \begin{bmatrix} T(\mathbf{u}) & T(\mathbf{v}) \end{bmatrix} \implies A = \begin{bmatrix} T(\mathbf{u}) & T(\mathbf{v}) \end{bmatrix} \begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix}^{-1}$$

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

■

5.2 Matrix Multiplication

2. Compute each matrix sum or product if it is defined. If an expression is undefined, explain why. Let

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 4 & -5 & 2 \end{bmatrix}, B = \begin{bmatrix} -7 & -5 & 1 \\ 1 & -4 & -3 \end{bmatrix}, C = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}, D = \begin{bmatrix} 3 & 5 \\ -1 & 4 \end{bmatrix}$$

$-2A, \quad B - 2A, \quad AC, \quad CD$

Solution.
- $-2A = \begin{bmatrix} -4 & 0 & 2 \\ -8 & 10 & -4 \end{bmatrix}$
- $B - 2A = \begin{bmatrix} -7 & -5 & 1 \\ 1 & -4 & -3 \end{bmatrix} + \begin{bmatrix} -4 & 0 & 2 \\ -8 & 10 & -4 \end{bmatrix} = \begin{bmatrix} -11 & -5 & 3 \\ -7 & 6 & -7 \end{bmatrix}$
- AC is not defined, because the dimensions don't match.
- $$CD = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ -1 & 4 \end{bmatrix}$$
  $$= \begin{bmatrix} 3(1) - (2) & 5(1) + 4(2) \\ 3(-2) - (1) & 5(-2) + 4(1) \end{bmatrix} = \begin{bmatrix} 1 & 13 \\ -7 & -6 \end{bmatrix}$$

■

3. Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 69 & 420 \\ 5 & 5 \end{bmatrix}$, and $C = \begin{bmatrix} \pi & \ln(2) \\ 5 & 5 \end{bmatrix}$. Verify that $AB = AC$ and yet $B \neq C$.

Solution.

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 69 & 420 \\ 5 & 5 \end{bmatrix} = \begin{bmatrix} 5 & 5 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 69 & 420 \\ 5 & 5 \end{bmatrix}$$

Therefore, $AB = AC \not\Rightarrow B = C$, so we don't get a nice cancellation property with matrices. ■

Remark 13. The idea behind this exercise is to caution you that matrix multiplication is not as nice as multiplication of simple numbers. For example, we know $3x = 3y \implies x = y$. And, in general, if $a \neq 0$ is a real number, then

$$ab = ac \implies b = c$$

But we can see this is not so for matrices. This is because some nonzero matrices can "lose information" (though having a nontrivial kernel).

In this case, the difference between B and C happened to be *in* the kernel of A, so the difference was lost and AB was identical of AC.

This is why *invertible matrices* are so important. Because then we *can* just cancel them! They are the ultra important matrices that *never* lose information.

The matrix A in this exercise is actually more sinister than just any arbitrary non-invertible square matrix. It's called "nilpotent" because $A^2 = 0$. See more in (7.3).

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

   (a) Compute $3I_2 - A$ and $(3I_2)A$.

   <u>*Solution.*</u> $3I_2 - A = \begin{bmatrix} -1 & 1 \\ -5 & 5 \end{bmatrix}$, $(3I_2)A = 3A = \begin{bmatrix} 12 & -3 \\ 15 & -6 \end{bmatrix}$ ■

   (b) Compute $B - 5I_3$ and $(5I_3)B$

   <u>*Solution.*</u>

$$B - 5I_3 = \begin{bmatrix} -4 & 1 & 1 \\ 1 & -3 & 2 \\ 1 & 2 & -2 \end{bmatrix}, \quad (5I_3)B = 5B = \begin{bmatrix} 5 & 5 & 5 \\ 5 & 10 & 10 \\ 5 & 10 & 15 \end{bmatrix}$$

■

5. Compute the product AB in two ways:

   (a) by the definition, where $A\mathbf{b}_1$ and $A\mathbf{b}_2$ are computed separately, and

   (b) by the row–column rule for computing AB

   with $A = \begin{bmatrix} 4 & -2 \\ -3 & 0 \\ 3 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}$.

   <u>*Solution*</u>.

$$A\mathbf{b}_1 = \begin{bmatrix} 4 & -2 \\ -3 & 0 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1 \begin{bmatrix} 4 \\ -3 \\ 3 \end{bmatrix} + 2 \begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 13 \end{bmatrix}$$

$$A\mathbf{b}_2 = \begin{bmatrix} 4 & -2 \\ -3 & 0 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = 3 \begin{bmatrix} 4 \\ -3 \\ 3 \end{bmatrix} - \begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 15 \\ -9 \\ 4 \end{bmatrix}$$

   Thus,

$$AB = \begin{bmatrix} 0 & 15 \\ -3 & -9 \\ 13 & 4 \end{bmatrix}$$

   With the alternate computation being

$$AB = \begin{bmatrix} 1(4) & + & 2(-2) & 3(4) & - & (-2) \\ 1(-3) & + & 2(0) & 3(-3) & - & (0) \\ 1(3) & + & 2(5) & 3(3) & - & (5) \end{bmatrix} = \begin{bmatrix} 0 & 15 \\ -3 & -9 \\ 13 & 4 \end{bmatrix}$$

   Observe how these methods are, in fact, equivalent. ■

6. Assume that each matrix expression is defined. That is, the sizes of the matrices (and vectors) involved "match" appropriately.

(a) If a matrix A is 5×3 and the product AB is 5×7, what is the size of B?

Solution. Like before with vectors, we need the dimensions to match.

$$(5 \times 3) \cdot (? \times ?) = (5 \times 7)$$

The inner ones have to match so the number of rows must be 3. For the output to have 7 columns, the right matrix must also have 7 columns. Thus, B must be a 3×7. ■

(b) If $T_1 \colon \mathbb{R}^3 \to \mathbb{R}^5$ and $T_1 \circ T_2 \colon \mathbb{R}^7 \to \mathbb{R}^5$ what is $T_2 \colon \mathbb{R}^? \to \mathbb{R}^?$

Solution. One can observe, that this problem is exactly equivalent to the previous problem, but let's pretend we don't know that.

Imagine you're taking two trains T_2 and T_1 (in that order). You know that T_1 starts in Paris and ends in Berlin. Further, you know that after taking both trains, you will have started in Madrid and ended up in Berlin. Where must the first train have gone between? Well, obviously it would have to be from Madrid to Paris. This same logic can be used for this problem.

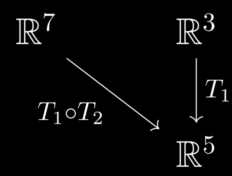
We have the following pieces of information:

- T_1 takes things from $\mathbb{R}^3$ to $\mathbb{R}^5$
- Doing T_1 after T_2 results in a net transformation from $\mathbb{R}^7$ to $\mathbb{R}^5$.

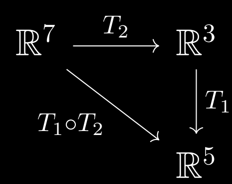
Essentially, here $\mathbb{R}^3$ is Paris, $\mathbb{R}^5$ is Berlin, and $\mathbb{R}^7$ is Madrid. So we expect $T_2 \colon \mathbb{R}^7 \to \mathbb{R}^3$ based on our logic from earlier. But let's reason it out with the exact spaces given.

The codomain makes sense for the composition, because we know the outputs of T_1 are $\mathbb{R}^5$. Since the inputs of T_1 are vectors in $\mathbb{R}^3$, and $T_1 \circ T_2$ being defined means the outputs of T_2 are valid inputs for T_1, that tells us the outputs of T_2 have to be in $\mathbb{R}^3$. Similarly, the inputs to T_2 have to be the same as the composition $T_1 \circ T_2$. So the inputs for T_2 have to be $\mathbb{R}^7$.

One way to think of it is to draw a picture:



For this to make sense, we need a line from $\mathbb{R}^7$ to $\mathbb{R}^3$. Thus, the diagram is



■

(c) How many rows does D have if DC is a 3×4 matrix?

Solution. The rows of a product are the same as the rows of the left matrix, so D has three rows.

In terms of linear transformations, if DC outputs vectors in $\mathbb{R}^3$ (which it does because it has three rows), and D is the last part of the transformation, then D must output things to $\mathbb{R}^3$ as well, so it has three rows. ■

5.3 Inverses

7. Find the inverse of $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}$ without doing *any* scratch work, given that

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

Solution. The key insight for this problem is that the columns of the inverse of a matrix are the preimages of the standard basis vectors. Notice the symmetry we have here.

- The columns of A are the images of the standard basis vectors under A.
- The columns of A^{-1} are the preimages of the standard basis vectors under A.

I encourage you to think about why this is true. As a hint, think about what happens if you do $A \circ A^{-1}$. What happens to the standard basis vectors?

Thus, the answer is just

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

You can verify multiplying these matrices does result in I. ■

8. Find the inverses of the following matrices. Show work.

(a) $\begin{bmatrix}0 & 1 & 0 & 0\\1 & 0 & 0 & 0\\-1 & 2 & 0 & 1\\0 & 0 & \frac{1}{2} & 0\end{bmatrix}$ using the row reduction algorithm (page 110)

Solution.

Remark 14. i don't want to type a solution for this.):

I'm just gonna do some intense row reduction.

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

$$\sim \begin{array}{c} R_1 \to R_2 \\ R_2 \to R_1 \\ R_3 \to 2R_4 \\ R_4 \to -2R_1 + R_2 + R_3 \end{array} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & -2 & 1 & 1 & 0 \end{array}\right]$$

You would need about 6 elementary operations to do this, and I don't want to type all that. Anyway, we get

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

■

Remark 15. I want to point out the similarities between the inverse we got and the row operations that got us there.

$$\begin{array}{c} R_1 \to R_2 \\ R_2 \to R_1 \\ R_3 \to 2R_4 \\ R_4 \to -2R_1 + R_2 + R_3 \end{array} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ -2 & 1 & 1 & 0 \end{bmatrix}$$

If we read the operations as

$$\text{(the row we go into} \to c_1R_1 + c_2R_2 + c_3R_3 + c_4R_4) \iff \begin{bmatrix} c_1 & c_2 & c_3 & c_4 \end{bmatrix}$$

We get this correspondence.

$$c_1R_1 + c_2R_2 + c_3R_3 + c_4R_4 \iff \begin{bmatrix} c_1 & c_2 & c_3 & c_4 \end{bmatrix}$$

And, in fact, this highlights the correspondence in calculating AB between linear combinations of the rows of the right matrix and the rows of the left matrix. That is, the rows of A tell you what linear combination of the rows of B to take. You can read more here.

There was going to be an exercise about this, but I cut it. You can see it (here).

(b) $\begin{bmatrix} \pi & -e \\ e & \pi \end{bmatrix}$ using the formula for a 2×2 inverse (page 105)

Solution. The formula is

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

$ad - bc = \pi^2 + e^2$, so

$$\begin{bmatrix} \pi & -e \\ e & \pi \end{bmatrix}^{-1} = \frac{1}{\pi^2 + e^2} \begin{bmatrix} \pi & e \\ -e & \pi \end{bmatrix}$$

■

Use the inverse found above to solve the system

$$\pi x_1 - e x_2 = \sqrt{2}$$
$$e x_1 + \pi x_2 = \ln(2)$$

Solution.

$$\frac{1}{\pi^2 + e^2} \begin{bmatrix} \pi & e \\ -e & \pi \end{bmatrix} \left(\begin{bmatrix} \pi & -e \\ e & \pi \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
$$= \frac{1}{\pi^2 + e^2} \begin{bmatrix} \pi & e \\ -e & \pi \end{bmatrix} \left(\begin{bmatrix} \sqrt{2} \\ \ln(2) \end{bmatrix} \right) = \frac{1}{\pi^2 + e^2} \begin{bmatrix} \pi\sqrt{2} + e\ln(2) \\ \pi\ln(2) - e\sqrt{2} \end{bmatrix}$$

■

5.3.1 Orthogonal Matrices Preview

9. Suppose that a, b are any two real numbers such that $a^2 + b^2 = 1$. Consider the matrices

$$S = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}, \quad Q = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

Show that the inverse for both these matrices is just their transpose. That is, $S^{-1} = S^T$ and $Q^{-1} = Q^T$.

Solution.

$$S^T S = \begin{bmatrix} a & b \\ b & -a \end{bmatrix} \begin{bmatrix} a & b \\ b & -a \end{bmatrix} = \begin{bmatrix} a^2+b^2 & 0 \\ 0 & a^2+b^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \checkmark$$

$$Q^T Q = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} a^2+b^2 & 0 \\ 0 & a^2+b^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \checkmark$$

■

5.4 Cayley-Hamilton

10. Last week (4.2) we looked at the matrix

$$A = \begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix}$$

and when looking at $A\mathbf{v} = \lambda\mathbf{v}$ we ended up with $\begin{bmatrix} 2-\lambda & -1 \\ -3 & 4-\lambda \end{bmatrix}$ (which we now recognize as $A - \lambda I$). The determinant of that matrix was $\lambda^2 - 6\lambda + 5$.

(a) Evaluate $A^2 - 6A + 5I$.

Solution.

$$\begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix}^2 - 6\begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix} + 5\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 7 & -6 \\ -18 & 19 \end{bmatrix} + \begin{bmatrix} -12 & 6 \\ 18 & -24 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Cool. ■

(b) For a general 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, you can take my word that

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \lambda I\right) = \lambda^2 - (a+d)\lambda + ad - bc$$

Compute $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^2 - (a+d)\begin{bmatrix} a & b \\ c & d \end{bmatrix} + (ad - bc)I$

Remark 16. Notice that the constant term is the determinant and the second highest term is the sum of the diagonal elements. That is actually always true (for $\det(\lambda I - A)$).

Solution.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^2 - (a+d)\begin{bmatrix} a & b \\ c & d \end{bmatrix} + (ad-bc)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} =$$
$$\begin{bmatrix} a^2 + bc & b(a+d) \\ c(a+d) & d^2 + bc \end{bmatrix}$$
$$+ \begin{bmatrix} -a^2 - ad & -b(a+d) \\ -c(a+d) & -ad - d^2 \end{bmatrix}$$
$$+ \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

■

Remark 17. This does actually always happen. It's called the Cayley-Hamilton theorem which is far beyond the scope of this course, but I think it's a cool magic trick.

(c) Based on part (a), show that $B = \frac{1}{5}(6I - A)$ is the inverse of A. I will accept either an algebraic justification or explicit verification that AB or BA give I.

Solution. I'll give the algebraic justification. The explicit computation is simple enough. We have shown in part (a) that

$$A^2 - 6A + 5I = 0 \implies I = \frac{1}{5}(6A - A^2) = \frac{1}{5}(6I - A)A$$

Well, that's the definition of an inverse, isn't it? That multiplying them together gives I (assuming they're *square*)! Thus,

$$\frac{1}{5}(6I - A)A = I \implies \frac{1}{5}(6I - A) = A^{-1}$$

And, just to double check,

$$\frac{1}{5}(6I - A) = \frac{1}{5}\begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix}$$

which is actually equal to the $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ matrix (what is called the *adjugate*) of $\begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix}$ divided by the determinant $2(4) - (-1)(-3)$. ■

Week #6 Worksheet Solutions

6.1 Invertibility of Transformations

1. Determine if the following matrices are invertible. Explain why.

(a) $\begin{bmatrix} \pi & 42 & e \\ \pi & 42 & e \\ \pi & 42 & e \end{bmatrix}$

(b) $\begin{bmatrix} 19 & 0 & 3 \\ 6 & 0 & 9 \\ 42 & 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 8 & -6 \\ -4 & 3 \end{bmatrix}$

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

(e) $\begin{bmatrix} \pi & 42 & e & 17 \\ 0 & \ln(2) & 42 & e \\ 0 & 0 & 55 & \sin(1) \\ 0 & 0 & 0 & 1984 \end{bmatrix}$

Solution.

(a) can't be invertible because the rows are linearly dependent (they are identical). The columns are also all scalar multiples, and thus linearly dependent, so the matrix can't be invertible.

(b) Also cannot be invertible. The columns are dependent (since one of them is zero). We can additionally see that $A\mathbf{e}_2 = \mathbf{0}$, so the kernel is nontrivial.

(c) Row 1 is -2 times row 2, so the rows are dependent. Thus, it's not invertible.

(d) This isn't even square! Not invertible.

(e) Triangular and has all nonzero entries on the diagonal so it's invertible.

■

2. Suppose A is an $n \times n$ invertible matrix. Are any of the following situations possible?

(a) A has two identical columns

(b) A has two identical rows

(c) The columns of A do not span $\mathbb{R}^n$

(d) A^{-1} has linearly independent columns

(e) $AB\mathbf{x} = \mathbf{0}$ has a nontrivial solution for some $n \times n$ matrix B.

Solution.

(a) No, then the columns are dependent.

(b) No, then the rows are dependent.

(c) No, then the function isn't surjective (this also means the rows are dependent)

(d) Yes, in fact this *has* to be true. If A is invertible, then so is A^{-1}.

(e) Yes, if B has a nontrivial kernel (that is, a nonzero solution to $Bx = 0$), then that will also be a solution to $ABx = 0$, since $A0 = 0$. For example, if B is the zero matrix, then *everything* is a solution.

■

6.2 LU Factorization

3. Verify the LU factorization

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

and then use it to solve

$$\begin{bmatrix} 2 & 5 \\ 6 & 9 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

Solution. $\begin{bmatrix}1 & 0\\3 & 1\end{bmatrix}\begin{bmatrix}2 & 5\\0 & -6\end{bmatrix}$ does indeed multiply to $A = \begin{bmatrix}2 & 5\\6 & 9\end{bmatrix}$.

Then to solve the system $A\mathbf{x} = LU\mathbf{x} = \mathbf{b}$, the process is as follows:

- Let $\mathbf{y} = U\mathbf{x}$
- Solve $L(U\mathbf{x}) = L\mathbf{y} = \mathbf{b}$
- Solve $U\mathbf{x} = \mathbf{y}$

Doing that we have

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

Thus, $\mathbf{y} = \begin{bmatrix}3\\-6\end{bmatrix}$. Next we solve

$$\left[\begin{array}{cc|c}2 & 5 & 3\\0 & -6 & -6\end{array}\right] \sim \left[\begin{array}{cc|c}2 & 0 & -2\\0 & 1 & 1\end{array}\right] \sim \left[\begin{array}{cc|c}1 & 0 & -1\\0 & 1 & 1\end{array}\right]$$

Therefore, $\mathbf{x} = \begin{bmatrix}-1\\1\end{bmatrix}$. ■

6.3 Orthogonality and Orthogonal Matrices

4. Suppose for $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{b} \in \mathbb{R}^3$

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b} \tag{3}$$

is consistent, and the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ have the very nice property that

$$\mathbf{v}_i^T\mathbf{v}_j = \begin{cases}0, & i \neq j\\1, & i = j\end{cases} \tag{4}$$

where we assume that a 1×1 matrix is a scalar value. For example, $\mathbf{v}_1^T\mathbf{v}_2 = \mathbf{v}_1^T\mathbf{v}_3 = 0$, but $\mathbf{v}_1^T\mathbf{v}_1 = 1$. We call a set of vectors satisfying (4) "orthonormal".

Remark 18. We have a mathematical way to say $\begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$ very compactly. It's

$$\delta_{ij} := \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \tag{5}$$

So we can say $\mathbf{v}_i^T \mathbf{v}_j = \delta_{ij}$.

Note that the identity matrix I satisfies $I_{ij} = \delta_{ij}$.

(a) Show that $x_i = \mathbf{v}_i^T \mathbf{b}$ and use that to write the solution of the system (3).

Solution. If we multiply $\mathbf{v}_i^T$ to both sides of (3) then the RHS is $\mathbf{v}_i^T \mathbf{b}$. But on the LHS, (4) tells us that $\mathbf{v}_i^T$ in a sense "deletes" all the $\mathbf{v}_j$'s that don't match $j = i$. The, we get

$$x_i \left(\mathbf{v}_i^T \mathbf{v}_i \right) = x_i(1) = \mathbf{v}_i^T \mathbf{b}$$

Essentially, multiplying by $\mathbf{v}_i^T$ isolates x_i. ■

(b) Based on the previous part, conclude that any orthonormal set is linearly independent.

Solution. If $\mathbf{b} = \vec{0}$, then we have that $\mathbf{v}_i^T \vec{0} = 0$, so

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + x_3 \mathbf{v}_3 = \mathbf{0} \implies x_i = 0$$

Which is the definition of being linearly independent! ■

(c) If Q is a 3×3 matrix with columns $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, implying the system (3) is just $Q\mathbf{x} = \mathbf{b}$. Then explain how the previous parts have shown

- $\mathbf{x} = Q^T \mathbf{b}$
- Q is invertible

- $Q\mathbf{x} = \mathbf{b}$ is consistent for all $\mathbf{b}$
- $Q^{-1} = Q^T$

Solution. We have that

$$\mathbf{x} = \begin{bmatrix} \mathbf{v}_1^T\mathbf{b} \\ \mathbf{v}_2^T\mathbf{b} \\ \mathbf{v}_3^T\mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \mathbf{v}_3^T \end{bmatrix} \mathbf{b} = Q^T\mathbf{b}$$

so that is the first point.

Part (b) showed the columns of this square matrix Q are linearly independent, so Q is invertible.

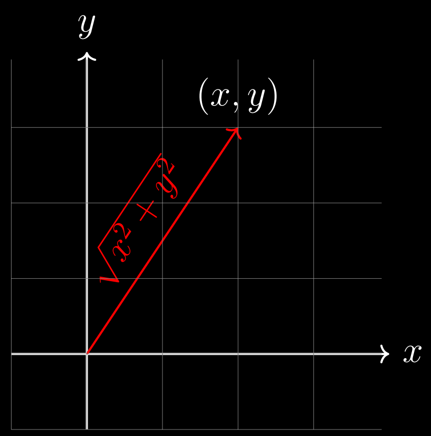
Since Q is invertible, $Q\mathbf{x} = \mathbf{b}$ is consistent for all $\mathbf{b}$.

And, since preimages of invertible matrices are given exactly by the inverse, and Q^T gives us our preimages, then Q^T has to be Q^{-1}. You can show this through other means like how through the first bullet we get $Q\mathbf{x} = \mathbf{b} = Q(Q^T\mathbf{b}) \implies \mathbf{b} = QQ^T\mathbf{b}$ for all $\mathbf{b}$, thus Q^T is the inverse. But whatever. ■

6.4 Magnitudes and Angles of Vectors (Inner Products)

5. One thing that we are often very concerned with is the *angle* between vectors and their *size*. The way we measure these things in $\mathbb{R}^n$ is typically with the *dot product.*

   Starting with magnitude (or the *size*) of vectors in $\mathbb{R}^2$, we can use the Pythagorean theorem to say the magnitude of $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$, which we denote $\|\mathbf{v}\|$ is $\sqrt{x^2 + y^2}$.



(a) Show that

$$\|\mathbf{v}\|^2 = \mathbf{v}^T\mathbf{v} \tag{6}$$

We can use this as a general definition for magnitude in $\mathbb{R}^n$ (though one can also prove this result geometrically). Note that if a vector has magnitude one, then we call it a "unit vector".

Solution.

$$\mathbf{v}^T\mathbf{v} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + y^2 = \|\mathbf{v}\|^2$$

■

(b) One property we should expect for magnitude is that a nonzero vector should have a nonzero (positive) magnitude! Similarly, we should only get a magnitude of zero when the vector is zero. This may seem trivial, but this means we have to be careful with our definition. Show that if $\mathbf{v} = \begin{bmatrix} 1 \\ i \end{bmatrix} \in \mathbb{C}^2$ (where i is the imaginary number such that $i^2 = -1$), then

$$\mathbf{v}^T\mathbf{v} = \begin{bmatrix} 1 & i \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix}$$

does not give a sensible magnitude for the vector. We do have to modify this definition slightly for $\mathbb{C}^n$ (but not by much!), but we will focus on $\mathbb{R}^n$ here, though. See (9.6) for how we fix it.

<u>*Solution.*</u>

$$\begin{bmatrix} 1 & i \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = 1^2 + i^2 = 0$$

but the vector $\mathbf{v} \neq 0$, so this doesn't work. ■

Remark 19. The proper way to define magnitude in $\mathbb{C}^n$ is with

$$\left\| \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \right\|^2 = |v_1|^2 + \ldots + |v_n|^2$$

where $|z|$ denotes the magnitude of the complex number z. Since $|z|^2 = \overline{z}z$, where $\overline{z}$ denotes the *conjugate*, this gives us

$$|\mathbf{v}|^2 = \mathbf{v}^* \mathbf{v}$$

where $\mathbf{v}^* := (\overline{\mathbf{v}})^T$ is the "conjugate transpose". Note that if $\mathbf{v}$ is real, this does give us our previous definition that $\|\mathbf{v}\|^2 = \mathbf{v}^T \mathbf{v}$. So the *correct* definition of magnitude of a vector in $\mathbb{C}^n$ and $\mathbb{R}^n$ is

$$\|\mathbf{v}\|^2 = \mathbf{v}^* \mathbf{v}$$

In general, the conjugate transpose (often called "adjoint") is the *correct* transpose. All the nice things that you get with the transpose in $\mathbb{R}^n$ are obtained by the conjugate transpose in $\mathbb{C}^n$ (and, again, the conjugate transpose can be used interchangably with the tranpose in $\mathbb{R}^n$), so you can pretty much always use it!

(c) In general, for $\mathbb{R}^n$, we can define the dot product as follows

$$\mathbf{v} \cdot \mathbf{w} := \mathbf{v}^T \mathbf{w} = v_1 w_1 + \ldots + v_n w_n \tag{7}$$

Note that based on this definition $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$.

One can prove geometrically that

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\| \cos(\theta) \tag{8}$$

where θ is the angle between the vectors. Use (8) to explain why vectors which are perpendicular (have an angle of 90°) should have a dot product of zero.

Remark 20. In mathematics, the term "orthogonal" is preferred over "perpendicular". Advanced linear algebra often focuses on inner products, which explore various methods for measuring magnitudes and angles, though the dot product is the most common. Generally, vectors are called orthogonal if they have an inner product of zero. For this class and assignment, we will consider vectors orthogonal if their dot product is zero, as defined in (7).

Solution. $\cos(90°) = 0$, so that expression should be zero if the angle is 90°. ■

6.4.1 Projectors

(d) One can also geometrically (or even with calculus) prove that the "projection" of one vector $\mathbf{w}$ onto $\mathbf{v}$ (i.e. the vector on the span of $\mathbf{v}$ closest to $\mathbf{w}$) is

$$\text{proj}_{\mathbf{v}}(\mathbf{w}) = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|^2}\mathbf{v} \tag{9}$$

Show that

$$\text{proj}_{\mathbf{v}}(\mathbf{w}) = \left(\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}\right)\mathbf{w}$$

implying that $\left(\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}\right)$ is the matrix which projects onto $\mathbf{v}$. Explain why if $\mathbf{v}$ is a unit vector then the projector matrix onto $\mathbf{v}$ is just $\mathbf{v}\mathbf{v}^T$.

Remark 21. We assume that expressions like $\mathbf{w}^T\mathbf{v}$ or $\mathbf{v}^T\mathbf{v}$ are scalars that can be divided (if they are nonzero) rather than 1×1 matrices (since we cannot divide by a matrix). Technically, we can rewrite

$$\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} = \mathbf{v}\left(\mathbf{v}^T\mathbf{v}\right)^{-1}\mathbf{v}^T$$

but for simplicity we will just treat dot products as scalars.

Solution.

$$\frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{v}\|^2}\mathbf{v} = \mathbf{v}\frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}^T\mathbf{v}} = \frac{\mathbf{v}}{\mathbf{v}^T\mathbf{v}}\left(\mathbf{v}^T\mathbf{w}\right) = \left(\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}\right)\mathbf{w}$$

■

(e) Verify that $P^2 = P$ for

$$P = \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$$

Solution.

$$P^2 = \left(\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}\right)\left(\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}\right) = \frac{\mathbf{v}\cancel{(\mathbf{v}^T\mathbf{v})}\mathbf{v}^T}{(\mathbf{v}^T\mathbf{v})^{\cancel{2}}} = \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} = P$$

This cancellation is okay because $\mathbf{v}^T\mathbf{v}$ is just a scalar we can pull in and out of the matrix/vector multiplication. ■

6.5 Orthogonal Matrices (Part II)

(f) Last week (5.3.1), I showed you the matrices

$$S = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}, \quad Q = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

where we assumed that $a^2 + b^2 = 1$, and asked you to show that their inverses were just their transpose.

i. Show that the columns of these matrices are unit vectors which are orthogonal

Solution. $\|(a,b)\|^2 = a^2 + b^2 = 1$. Similarly,

$$b^2 + (-a)^2 = (-b)^2 + a^2 = a^2 + b^2 = 1$$

so the columns are all unit vectors.

$$\begin{bmatrix} a \\ b \end{bmatrix} \cdot \begin{bmatrix} b \\ -a \end{bmatrix} = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} b \\ -a \end{bmatrix} = ab - ba = 0$$

so the columns of S are orthogonal.

Similarly, $\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} -b \\ a \end{bmatrix} = -ab + ba = 0$ so the columns of Q are orthogonal. ■

ii. Explain how that explains *why* their inverse is just their transpose.

Hint: the i, j entry of AB is $\mathbf{a}_i^T \mathbf{b}_j$ where $\mathbf{a}_i$ is the ith row of A (as a colummn vector) and $\mathbf{b}_j$ is the jth column of B.

Solution. Like the hint says, if we let $\mathbf{a}_i$ be the ith column of A, then

$$(A^T A)_{ij} = \mathbf{a}_i^T \mathbf{a}_j = \mathbf{a}_i \cdot \mathbf{a}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} = \delta_{ij}$$

Which is saying that the entries on the diagonal (when $i = j$) are all 1, and everything off the diagonal ($i \neq j$) is zero. That means $A^T A$ is the idenity matrix, meaning A^T is the inverse of A (since they are square). ■

From this conclude that any matrix such that $A^T A = I$ must have columns which are orthogonal unit vectors (i.e. "othonormal").

Solution. Indeed,

$$\begin{aligned} A^T A = I &\iff (A^T A)_{ij} = (I)_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \\ &\iff \mathbf{a}_i \cdot \mathbf{a}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \\ &\iff \{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \text{ are orthonormal} \end{aligned}$$

■

(g) We call a real square matrix with orthonormal columns (whose inverse must be its transpose) an "orthogonal matrix". Show that an orthogonal matrix (that is, a matrix satisfying $Q^{-1} = Q^T$) preserves angles and magnitudes by showing that

i. $\|Q\mathbf{v}\|^2 = \|\mathbf{v}\|^2$

Proof.

$$\|Q\mathbf{v}\|^2 = (Q\mathbf{v})^T(Q\mathbf{v}) = \mathbf{v}^T Q^T Q\mathbf{v} = \mathbf{v}^T I\mathbf{v} = \|\mathbf{v}\|^2$$

□

ii. $(Q\mathbf{v}) \cdot (Q\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$

Proof.

$$(Q\mathbf{v}) \cdot (Q\mathbf{w}) = (Q\mathbf{v})^T(Q\mathbf{w}) = \mathbf{v}^T Q^T Q\mathbf{w} = \mathbf{v}^T I\mathbf{w} = \mathbf{v} \cdot \mathbf{w}$$

□

Remark 22. The second property actually implies the first. Additionally, the property $\|Q\mathbf{v}\|^2 = \|\mathbf{v}\|^2$ actually proves that Q is injective (has a trivial kernel), since if $Q\mathbf{v} = 0$, then $\|\mathbf{v}\| = \|Q\mathbf{v}\| = \|0\| = 0$ implying that $\mathbf{v} = 0$ (since only the zero vector has a magnitude zero).

6.6 Eigenvectors of Symmetric Matrices

(h) I mentioned before that you have *no idea* how amazing Symmetric matrices are. This is one the main reasons.

Suppose $S = S^T$ is symmetric and has two nonzero magical (eigen)vectors $\mathbf{v}$ and $\mathbf{w}$ with two different magical (eigen)values.

$$S\mathbf{v} = \lambda\mathbf{v}, \quad S\mathbf{w} = \mu\mathbf{w}$$

where $\lambda \neq \mu$. Show that $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

Hint: Use the fact that, since it's a 1×1 scalar value (and thus symmetric),

$$\mathbf{w}^T S\mathbf{v} = \left(\mathbf{w}^T S\mathbf{v}\right)^T = \mathbf{v}^T S^T (\mathbf{w}^T)^T = \mathbf{v}^T S\mathbf{w}$$

Solution. We have that

$$\mathbf{w}^T S\mathbf{v} = \mathbf{v}^T S\mathbf{w}$$

Then,

$$\begin{aligned}
\mathbf{w}^T(\lambda\mathbf{v}) &= \mathbf{v}^T(\mu\mathbf{w}) \\
\lambda\mathbf{w}^T\mathbf{v} &= \mu\mathbf{v}^T\mathbf{w} \\
\lambda\left(\mathbf{w}^T\mathbf{v}\right) &= \mu\left(\mathbf{v}^T\mathbf{w}\right) \\
\lambda\left(\mathbf{w}\cdot\mathbf{v}\right) &= \mu\left(\mathbf{v}\cdot\mathbf{w}\right) \\
\lambda\left(\mathbf{w}\cdot\mathbf{v}\right) &= \mu\left(\mathbf{w}\cdot\mathbf{v}\right)
\end{aligned}$$

as we said before, the dot product is commutative, so we can change the order. However, we're assuming that $\lambda \neq \mu$. So the only way that

$$\lambda\left(\mathbf{w}\cdot\mathbf{v}\right) = \mu\left(\mathbf{w}\cdot\mathbf{v}\right)$$

is if $\mathbf{w}\cdot\mathbf{v} = 0$. Therefore, they must be orthogonal. ■

Remark 23. This is an incredibly powerful property. My intention for this exercise was to show you that orthogonality makes your life very easy. The fact that the magical (eigen)vectors of a symmetric matrix are orthogonal means any computations that have to do with the magical (eigen)vectors is easier than it would be with any other arbitrary matrix.

Week #7 Worksheet Solutions

7.1 Vector Spaces, Subspaces

1. Which of the following are subspaces over $\mathbb{R}$? If one is *not* a subspace, give a linear combination of elements that is not in the set.

   Remark 24. Many of these which are not subspaces break multiple or all three requirements to be a subspace. I won't always mention why all three don't work, and sometimes instead just mention the broken rule I think is easiest to see or justify (ex. sometimes it's faster to say "doesn't contain a zero vector", but maybe it's more intuitive to see that it isn't closed under scalar multiplication by 0).

   (a) $\{x + 0i \in \mathbb{C} : x \in \mathbb{R}\}$ (the real numbers)

   Solution. Yes, a linear combination of real numbers is still a real number. ■

   (b) $\{0 + xi \in \mathbb{C} : x \in \mathbb{R}\}$ (purely imaginary numbers)

   Solution. Yes, a linear combination of real numbers times i is still a real number times i.

   $$c_1(ai) + c_2(bi) = (c_1a + c_2b)i$$

   ■

   (c) Smooth functions with inputs in $\mathbb{R}$ (i.e. functions for which every derivative is continuous), denoted $C^\infty(\mathbb{R})$ under standard addition and scalar multiplication

   Solution. Yes, smooth functions are just about as well behaved as you can hope for. The proof that they form a subspace is more advanced, and a question for real analysis, but here we'll just take it as "proof by obvious". ■

(d) Functions $f \in C^\infty(\mathbb{R})$ such that $f(0) = 0$.

Solution. Yes, if $f(0) = g(0) = 0$, then

$$(c_1 f + c_2 g)(0) = c_1 f(0) + c_2 g(0) = 0 + 0 = 0 \quad \checkmark$$

■

(e) Functions $f \in C^\infty(\mathbb{R})$ such that $f(1) = 1$.

Solution. No. Two possible justifications are

- $(2f)(1) = 2f(1) = 2 \neq 1$. Therefore, it's not closed under scalar multiplication.
- The additive identity (the function that is 0 everywhere) is not a part of the set.

It is also not closed under addition. ■

(f) $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$ such that $x \geq 0$ and $y \geq 0$ (the first quadrant)

Solution. No, even though it contains $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and is closed under addition, it's not closed under scalar multiplication (by *negative* scalars). $-\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ is not in the set even though $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is. ■

(g) Polynomials of degree exactly n in $\mathbb{R}_n[x]$

Solution. No. This breaks all three subspace rules, but the easiest to see is that even though x^n is in the set, $0x^n = 0$ is not, since it's degree 0 not n. ■

(h) Polynomials of the form $at^2 \in \mathbb{R}_n[t]$ for any $a \in \mathbb{R}$. For which n is this a subspace (if any)?

Solution. Yes. $c_1(at^2)+c_2(bt^2)=(c_1a+c_2b)t^2$ so it's closed under linear combinations. $n \geq 2$ is necessary. If $n < 2$, then t^2 isn't even a member of the set! If $n \geq 2$, then everything is still fine, since at^2 is either degree 0 or 2 (if $a = 0$ or not respectively), which is sufficiently less than or equal to n if $n \geq 2$. ■

(i) For some fixed matrix A, vectors of the form $\mathbf{x} = A\mathbf{y}$

Solution. Yes. $c_1(A\mathbf{x}_1) + c_2(A\mathbf{x}_2) = A(c_1\mathbf{x}_1 + c_2\mathbf{x}_2)$ so it's closed under linear combinations. ■

Remark 25. Thus, the image of A is a subspace.

(j) For some fixed matrix A, vectors $\mathbf{x}$ such that $A\mathbf{x} = \vec{0}$

Solution. Yes. $A(c_1\mathbf{x}_1 + c_2\mathbf{x}_2) = c_1A\mathbf{x}_1 + c_2A\mathbf{x}_2 = \vec{0} + \vec{0} = \vec{0}$ so it's closed under linear combinations. ■

Remark 26. Thus, the kernel of A (the solution space to a homogeneous equation) is a subspace.

(k) For some fixed $A \in \mathbb{R}^{2\times 2}$, vectors $\mathbf{x}$ such that $A\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Solution. No, the set does not contain $\vec{0}$ (and it's not closed under scalar mult. or addition). ■

Remark 27. Thus, the solution space to a nonhomogeneous equation is not a subspace, while a homogeneous solution space *is*.

(l) $\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ for $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$

Solution. Yes. Linear combinations are closed under linear combinations! ■

Remark 28. Every subspace can be written as the span of some set of vectors. In a way, span is the *only* kind of subspace there is!

7.1.1 Differential Equations Teaser

(m) Solutions of the differential equation $y'' + y = 0$.

<u>*Solution.*</u> Yes, because the derivative is linear. We can see $0''+0 = 0$, so the additive identity is in the set. Next, say y_1, y_2 are both solutions.

$$\begin{aligned}(y_1 + ky_2)'' + (y_1 + ky_2) &= y_1'' + ky_2'' + y_1 + ky_2 \\ &= (y_1'' + y_1) + k(y_2'' + y_2) = 0 + k0 = 0\end{aligned}$$

Thus, it's closed under linear combinations. ■

Remark 29. In differential equations, this is called the solution space of a "linear <u>homogeneous</u> differential equation".

(n) Solutions of the differential equation $y'' + y = 1$.

<u>*Solution.*</u> If y is such a solution, then

$$(2y)'' + (2y) = 2y'' + 2y = 2(y'' + y) = 2 \neq 1$$

Thus, it is not closed under linear combinations. ■

Remark 30. In differential equations, this is called the solution space of a "linear <u>*non*homogeneous</u> differential equation".

It's not a coincidence that the solution space to a homogeneous differential equation is a subspace and not so for a nonhomogeneous equation, just like it is for linear systems.

(o) Vectors in $\mathbb{R}^3$ with at least one entry that is zero.

<u>*Solution.*</u> No. Though closed under scalar multiplication, and though it contains $\vec{0}$, it is not closed under addition.

$$\begin{bmatrix}1\\0\\0\end{bmatrix} + \begin{bmatrix}0\\1\\1\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

Though both vectors on the left are in the set, their sum is not. ■

(p) Is $\mathbb{Q}$ (the rational numbers) a subspace of the real numbers?

Solution. 0 is a rational number, and rational numbers are closed under addition, but if our scalars are allowed to be real, then it cannot possibly be closed under scalar multiplication. For example,

$$\underbrace{\sqrt{2}}_{\text{scalar}} \cdot \underbrace{1}_{\text{vector}} = \underbrace{\sqrt{2}}_{\text{vector}}$$

■

2. The following IS a vector space (you don't have to prove it). Determine what the additive identity is (Hint: It is *not* the real number 0).

$$V = \mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}, \quad \begin{cases} x \oplus y := x \cdot y \\ c \otimes x := x^c \end{cases}$$

That is, *strictly positive* (this means nonzero, nonnegative real numbers) where "vector addition" is defined as real number *multiplication* and "scalar multiplication" is defined as real number exponentiation.

Solution. If "addition" is multiplication, then the thing that you "add" (multiply) which won't change things is the number 1.

$$\underbrace{1}_{\text{vector}} \oplus \underbrace{x}_{\text{vector}} = \underbrace{1 \cdot x}_{\text{vector}} = \underbrace{x}_{\text{vector}}$$

Another way to check (or do the problem, I suppose) is to use the theorem that $\underbrace{0}_{\text{scalar}} \cdot \underbrace{\mathbf{v}}_{\text{vector}} = \underbrace{\vec{0}}_{\text{vector}}$. Then,

$$\underbrace{0}_{\text{scalar}} \otimes \underbrace{2}_{\text{vector}} = \underbrace{2^0}_{\text{vector}} = \underbrace{1}_{\text{vector}}$$

■

7.2 Column Space, Null Space

3. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$, and $A = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3]$.

(a) How many linearly independent vectors are in $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

(b) What is the dimension of Col A?

(c) Is $\mathbf{p}$ in Col A? Why or why not?

(d) Is $\mathbf{p}$ in the subspace generated by $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$

Solution. Parts a,b,c,d can all be answered by row reducing the matrix A augmented with $\mathbf{p}$, but I'd like to make two observations

- (a) and (b) are the same question. The dimension of Col A, which is the span of the columns, is the number of linearly independent vectors. In general, the dimension of the span of some set of vectors is the number of linearly independent vectors there are in the set.
- (c) and (d) are the same question. Again, Col A is the span of the columns, so being in Col A is to be in the span of the columns. And the "subspace generated by" a set of vectors is just the span of those set of vectors!

So, row reducing,

$$\left[\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & 1 & 1 & 1 \\ -1 & 1 & -1 & 2 \end{array}\right] \sim \left[\begin{array}{ccc|c} \boxed{1} & 0 & 2 & -1 \\ 0 & \boxed{1} & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

(a)&(b): A has $\boxed{2}$ pivots, which means there are $\boxed{2}$ linearly independent columns which means the dimension of Col A is $\boxed{2}$.

(c)&(d): The system $A\mathbf{x} = \mathbf{p}$ is clearly consistent, based on the REF matrix not having any $0 = 1$ rows, so $\mathbf{p}$ must be in the column space. Specifically, based on the solution to the system, we can see that

$$\mathbf{p} = -\mathbf{v}_1 + \mathbf{v}_2$$

(which you can verify). Therefore, $\mathbf{p}$ is in the span of the $\mathbf{v}$ vectors, which is exactly what it means to be in the subspace generated by the $\mathbf{v}$ vectors.

■

(e) Is $\mathbf{p}$ is in Nul A?

Solution. This isn't directly answered by row reduction, but this is the easiest one to check.
$\mathbf{p} \in \ker(A)$ is the same as saying $A\mathbf{p} = \vec{0}$.
Just looking at the first entry of

$$\begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ \\ \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

then it can't be in the null space/kernel. ■

4. Consider the conjugation operation

$$f \colon \mathbb{C} \to \mathbb{C}, \quad f(a+bi) = a - bi$$

(a) Show f IS a linear transformation on $\mathbb{C}$ as a real vector space (with *real* scalars).

Solution.

$$\begin{aligned} &f\big((a+bi) + k(c+di)\big) \\ &\quad = f\big((a+kc) + (b+kd)i\big) \\ &\quad = \big(a+kc\big) - \big(b+kd\big)i \\ &\quad = \big(a-bi\big) + k\big(c-di\big) \\ &\quad = f\big(a+bi\big) + kf\big(c+di\big) \end{aligned}$$

■

(b) Show that f is NOT a linear transformation if our scalars are allowed to be complex.

Hint: The only thing that changed is our ability to use i as a scalar. Check $f(i\,(a+bi))$!

Solution. If this is a linear transformation, then $f\Big(i\big(a+bi\big)\Big)$ should be equal to $i\,f\big(a+bi\big)$.

$$\begin{array}{ccc} f\Big(i\big(a+bi\big)\Big) & & i\,f\big(a+bi\big) \\ =f\big(-b+ai\big) & & =i\ (a-bi) \\ -b-ai & \neq & =b+ai \end{array}$$

Thus, if the scalars are allowed to be complex, then this conjugation operation doesn't preserve scalar multiplication. So it's not a linear transformation if the scalars can be complex.

Alternate quicker solution: We should expect $f(-1)=f(i(i))=if(i)$. But

$$f(-1)=-1\neq i(-i)=1$$

■

7.3 The Derivative Operator

7.3.1 On Polynomials

5. Consider the linear transformation

$$D\colon \mathbb{R}_2[x]\to\mathbb{R}_2[x],\quad D\left(a_0+a_1x+a_2x^2\right)=a_1+2a_2x+0x^2$$

(a) Which familiar operation from calculus is this linear transformation equivalent to?

Solution. This is the derivative operation. ■

(b) Find the kernel and image of D

Solution. If something is in the kernel of D, then its derivative must be zero. We know from calculus that a constant has a derivative of zero, so

$$\ker(D) = \{\text{Constant polynomials}\} = \operatorname{span}\{1\}$$

We can also see that the the image appears to be all first degree polynomials or lower ($\operatorname{span}\{1, x\}$). One way to see this is that 1 has a preimage (x), and x has a preimage $\frac{x^2}{2}$, but x^2 has no preimage (since we are restricted to degree 2 or lower polynomials in our domain, we cannot use $\frac{x^3}{3}$). ■

Remark 31. Notice that the vector 1 is both in the image *and* the kernel.

(c) Find a matrix A such that

$$A \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix}$$

and describe its relation to D.

Solution. First of all, we can see that A is in pretty much every way the same thing as D, but on $\mathbb{R}^3$ vectors instead of the polynomials of $\mathbb{R}_2[x]$. If we rewrite D in the following way, we can see it's the same thing except we are just not writing the powers of x or $+$ signs.

$$D \begin{bmatrix} a_0 \\ +a_1 x \\ +a_2 x^2 \end{bmatrix} = \begin{bmatrix} a_1 \\ +2a_2 x \\ +0x^2 \end{bmatrix}, \quad A \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix}$$

One way to obtain A is to use the fact that the ith column is given by the image of $\mathbf{e}_i$. So, based on the definition,

$$A\begin{bmatrix}a_0\\a_1\\a_2\end{bmatrix}=\begin{bmatrix}a_1\\2a_2\\0\end{bmatrix}$$

- $A\begin{bmatrix}1\\0\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$ is the first col
- $A\begin{bmatrix}0\\1\\0\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}$ is the second col $\qquad \Longrightarrow A=\begin{bmatrix}0&1&0\\0&0&2\\0&0&0\end{bmatrix}$
- $A\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}0\\2\\0\end{bmatrix}$ is the third col

■

(d) Compute both A^2 and D^2 and compare them

<u>*Solution.*</u>

$$A^2=\begin{bmatrix}0&0&2\\0&0&0\\0&0&0\end{bmatrix}\qquad \begin{aligned}&D^2\left(a_0+a_1x+a_2x^2\right)\\&=D\left(a_1+2a_2x+0x^2\right)\\&=\left(2a_2+0x+0x^2\right)\end{aligned}$$

and, indeed, $A^2\begin{bmatrix}a_0\\a_1\\a_2\end{bmatrix}=\begin{bmatrix}2a_2\\0\\0\end{bmatrix}$. Thus, squaring A tells us exactly what applying D twice does. ■

(e) Compute A^3 and describe its image and kernel. Confirm you get the same result as D^3. Based on the operation you likened D to in part (a), why is the result of A^3 expected?

<u>*Solution*</u>.

$$A^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} &D^3\left(a_0 + a_1 x + a_2 x^2\right) \\ &= D\left(2a_2 + 0x + 0x^2\right) \\ &= \left(0 + 0x + 0x^2\right) \end{aligned}$$

The image of the zero matrix is just $\left\{\vec{0}\right\}$, and $\ker(A^3) = \mathbb{R}^3$. This is because everything gets sent to $\vec{0}$, so the image is only that vector, and the kernel, which is all the vectors that get sent to $\vec{0}$ is thus *everything* (i.e. $\mathbb{R}^3$ since it's a 3×3 matrix).

This is exactly what we would expect for "what happens" when you take three derivatives of any polynomial of degree 2 or less. It's always going to be zero! ■

Remark 32. A matrix for which some power gives the zero matrix is called "<u>nilpotent</u>". Nilpotent operators have a lot of interesting properties. But the primary characterization is that the image is "contained" in the kernel in such a way that the kernel "eats up" the image until eventually there's nothing left and it all goes to zero. Which makes sense for the derivative on polynomials, right? The degree always goes down by one until it's zero.

If you consider eigenvectors as "the good guys" of linear algebra, then, in a way, nilpotent operators are "the bad guys". Nilpotent is essentially the ingredient necessary for there to *not* be enough eigenvectors. And when an operator doesn't have enough eigenvectors, it develops "nilpotency pneumonia" (a phrase I literally just made up that means it is not "diagonalizable").

MATH132 is a course that spends a large majority of its content on how to "fight back" against operators that have a bad case of "nilpotency pneumonia", using the "Jordan form". The Jordan form is a way to deal with "bad matrices" that is as close to as nice as matrices which have a full set of eigenvectors (i.e. a form where they are diagonalizable).

7.3.2 Differential Equations Introduction

6. We can consider the derivative operator D in general on the set of smooth functions $C^\infty(\mathbb{R})$ by

$$D(f(x)) = f'(x)$$

(a) What is the kernel of D?

Solution. Again, all constant functions. So, span $\{1\}$. ■

(b) Find one example of a nonzero magical vector (i.e. function) of D that stays the same after applying D

$$D(f(x)) = f(x)$$

Hint: What is a nonzero function that is equal to its own derivative?

Solution. $f(x) = e^x$ is a great example. ■

(c) Use the previous part to find a nonzero solution $y_1(x)$ to

$$y'(x) - y(x) = 0$$

(this requires no scratch work)

Solution. $y_1(x) = e^x$ is a solution since $y' = y \iff y' - y = 0$ ■

(d) If we define the operator $L = D - I$ (you can assume that it is linear), then

$$L[y] = y' - y$$

Explain why the function you found above $y_1 \in \ker(L)$.

Solution. If $y' - y = 0$, then that literally means $L[y] = 0$. So it must be in $\ker(L)$. Thus, $e^x \in \ker(L)$. ■

(e) Since the kernel is a subspace, give an expression that yields *infinitely* many solutions to the differential equation $y' - y = 0$.

Hint: Being a subspace means it's closed under linear combinations (specifically, scalar multiplication)!

Solution. The kernel is closed under scalar multiplication, so $2e^x$, $3e^x$, $-e^x$, πe^x, etc. will all be in the kernel. Thus, $y = Ce^x$ for any $C \in \mathbb{R}$ will be in the kernel.

Since the kernel is the homogeneous solutions space, then

$$y = Ce^x, \quad C \in \mathbb{R}$$

is an expression for infinitely many solutions. ■

Remark 33. This is actually the "general solution". That is, every solution to $y' - y = 0$ is of the form Ce^x. So, congrats, you may have just solved your first differential equation!

Week #8 Worksheet Solutions

8.1 Dimension and Basis

1. Find a basis of Col A where $A = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}$.

   Use the standard method detailed in section 4.3 of the textbook so that everyone gets the same answer.

   Solution.
   $$A \sim \begin{bmatrix} \boxed{1} & 4 & 0 & 2 & 0 \\ 0 & 0 & \boxed{1} & -1 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
   so the pivot columns are columns 1, 3, 5, so the first, third, and fifth column form a basis for Col A.
   $$\text{Col}\, A = \text{span}\left\{ \begin{bmatrix} 1 \\ 3 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 2 \\ 8 \end{bmatrix} \right\}$$
   ■

2. Consider the subspace $W = \left\{ \begin{bmatrix} s \\ s \\ 0 \end{bmatrix} : s \in \mathbb{R} \right\}$. We can say that
   $$\begin{bmatrix} s \\ s \\ 0 \end{bmatrix} = s \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Hence, every element of W is a linear combination of $\mathbf{e}_1, \mathbf{e}_2$.

(a) Is W contained in span $\{\mathbf{e}_1, \mathbf{e}_2\}$?

Solution. Yes. Every element has been shown to be a linear combination of $\mathbf{e}_1, \mathbf{e}_2$, so W must be contained in the span. ■

(b) Is $\{\mathbf{e}_1, \mathbf{e}_2\}$ a basis for W? If not, find a correct basis for W.

Solution. No, neither $\mathbf{e}_1$ nor $\mathbf{e}_2$ are actually in W! There's no value of s that can make $\begin{bmatrix} s \\ s \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. One can see this as W being the set of vectors with identical first and second entries, and a zero in the third entry. Neither $\mathbf{e}_1, \mathbf{e}_2$ have identical first entries so they aren't part of the set. A basis must be made up of vector in the space!

However, $\begin{bmatrix} s \\ s \\ 0 \end{bmatrix} = s \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, so it seems we can write every element of W as a linear combination of $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ uniquely! Thus, a basis is

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

(Any nonzero scalar multiple would also be a valid basis) ■

3. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 7 \\ -9 \end{bmatrix}$. Determine if $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $\mathbb{R}^3$. Is it a basis for $\mathbb{R}^2$?

Solution. The vectors are certainly linearly independent, but they cannot span $\mathbb{R}^3$ because no set of only two vectors can span a three dimensional space!

However, it is also definitely not a basis for $\mathbb{R}^2$, because the vectors aren't in $\mathbb{R}^2$! So it's $\boxed{\textbf{NOT A BASIS FOR } \mathbb{R}^2}$. ■

Remark 34. Was I expecting you to say it was a basis for $\mathbb{R}^2$? Yes, pretty much. But don't worry, every single student I've ever worked with (and that is a lot of students) has said the same thing. As I said in discussion, this is just a pet peeve of mine. But while strictly incorrect, your intuition is actually in the right place. Here is an explanation for how:

The span of the two vectors is two dimensional, so the span is isomorphic to $\mathbb{R}^2$, but not equal to $\mathbb{R}^2$. One such isomorphism is the coordinate map from $\mathbb{R}^2$ to $W = \text{span}\left\{\begin{bmatrix}1\\-2\\3\end{bmatrix}, \begin{bmatrix}-2\\7\\-9\end{bmatrix}\right\}$

$$\phi\colon \mathbb{R}^2 \to W, \quad \phi\left(\begin{bmatrix}c_1\\c_2\end{bmatrix}\right) = c_1\begin{bmatrix}1\\-2\\3\end{bmatrix} + c_2\begin{bmatrix}-2\\7\\-9\end{bmatrix}$$

It can be shown this map is a bijection.

4. The following is a matrix A and an echelon form of A.

$$A = \begin{bmatrix}-3 & 9 & -2 & -10\\-3 & 9 & -2 & -10\\2 & -6 & 4 & 4\\3 & -9 & -2 & 14\\3 & -9 & -2 & 14\end{bmatrix} \sim \begin{bmatrix}1 & -3 & 0 & 4\\0 & 0 & 1 & -1\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix}$$

(a) Find a basis for Col A, Row A, and Nul A.

Solution. The pivot columns are columns one and three, so a basis for Col A is

$$\text{Col } A = \text{span}\left\{\begin{bmatrix}-3\\-3\\2\\3\\3\end{bmatrix}, \begin{bmatrix}-2\\-2\\4\\-2\\-2\end{bmatrix}\right\}$$

the nonzero rows of the row echelon form of A give the basis

$$\text{Row } A = \text{span}\left\{\begin{bmatrix}1\\-3\\0\\4\end{bmatrix}, \begin{bmatrix}0\\0\\1\\-1\end{bmatrix}\right\}$$

And we can, by inspection, see from taking linear combinations of the columns that a basis for the Null space (kernel) is

$$\text{Nul } A = \text{span}\left\{\begin{bmatrix}3\\1\\0\\0\end{bmatrix}, \begin{bmatrix}-4\\0\\1\\1\end{bmatrix}\right\}$$

■

(b) State the dimension of those subspaces and the rank and nullity of the matrix.

Solution. All three subspaces have dimension 2, since their basis each has two elements. The rank is subsequently 2 (the dimension of the column space) and the nullity is 2 (the dimension of the kernel/null space). ■

(c) The following is A^T and an echelon form of A^T.

$$A^T = \begin{bmatrix}-3 & -3 & 2 & 3 & 3\\ 9 & 9 & -6 & -9 & -9\\ -2 & -2 & 4 & -2 & -2\\ -10 & -10 & 4 & 14 & 14\end{bmatrix} \sim \begin{bmatrix}1 & 1 & 0 & -2 & -2\\ 0 & 0 & 2 & -3 & -3\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

Explain why

i. $\left\{ \begin{bmatrix} -3 \\ 9 \\ -2 \\ -10 \end{bmatrix}, \begin{bmatrix} 2 \\ -6 \\ 4 \\ 4 \end{bmatrix} \right\}$ is another basis for Row A

ii. $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ -3 \\ -3 \end{bmatrix} \right\}$ is another basis for Col A

Solution. The pivot columns of A^T give a basis for Col A^T (which is Row A), and the nonzero rows of a sufficiently reduced echelon form (with a minimal number of nonzero rows/maximal number of zero rows) of A^T give a basis for Row A^T (which is Col A). ■

Remark 35. Note that this basis for Col A is MUCH nicer. But it's harder to verify that it is actually in the column space (so, not as good for an exam where your professor has to be able to check your work. but it is an objectively better basis that's easier to work with).

iii. If we augment A^T with I and row reduce, we can obtain

$$\left[\begin{array}{ccccc|cccc} 1 & 1 & 0 & -2 & -2 & 0 & 0 & 1/8 & -1/8 \\ 0 & 0 & 1 & -3/2 & -3/2 & 0 & 0 & 5/16 & -1/16 \\ 0 & 0 & 0 & 0 & 0 & 4 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 3 & 3 \end{array}\right]$$

Consider the rows of the right matrix corresponding to the zero rows of the left matrix,

$$\left\{ \begin{bmatrix} 4 \\ 0 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 3 \\ 3 \end{bmatrix} \right\}$$

verify that this is another basis for Nul A.

Solution. You can confirm that

$$\begin{bmatrix} -3 & 9 & -2 & -10 \\ -3 & 9 & -2 & -10 \\ 2 & -6 & 4 & 4 \\ 3 & -9 & -2 & 14 \\ 3 & -9 & -2 & 14 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 4 \\ -1 & 3 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Since these are two clearly linearly independent vectors in the kernel (dimension 2), then it's also a basis. This is sufficient. However, if we compare the bases,

$$\left\{ \begin{bmatrix} 4 \\ 0 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 3 \\ 3 \end{bmatrix} \right\}_{new} \quad \text{vs.} \quad \left\{ \begin{bmatrix} 3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}_{old}$$

we can see some interesting similarities. The first of the new basis is -1 of the second of the original basis. And we can also see pretty easily that the second of the new basis is 4 of the first of the old basis $+3$ of the second of the old basis.

I just think it's cool you can get a basis for the Row, Col, and Nul spaces of A all in one shot by row reducing A^T. ■

Remark 36. Why does this work, though? Well, when row reducing A, we find which columns are independent by identifying pivot columns, and we get a basis for the row space by taking linear combinations of the rows of A (that is what row reduction is) until we get rows that are clearly independent.

When we row reduce A^T augmented with I, we get

$$\left[\ A^T \mid I\ \right] \to \left[\ \text{REF}(A^T) \mid B\ \right]$$

where B is a matrix such that $BA^T = \text{REF}(A^T)$. By row perspective, this means the rows of B corresponding to the rows of A^T

that become zero are coefficients for linear combinations of rows of A^T (columns of A) that give zero. But that's just the kernel of A! It's linear combinations of the columns of A that give zero. So this is one way to do it all very quickly and algorithmically.

8.2 Coordinate Vectors

5. Find the vector $\mathbf{x}$ determined by the given coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ and the given basis $\mathcal{B}$. (Think of this as using a dictionary to translate a sentence in a foreign language to your own.)

$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}, [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

Solution. The coordinate vector just tells us the coefficients for the linear combination so

$$\mathbf{x} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$$

■

6. The vector $\mathbf{x}$ is in a subspace H with a basis $\mathcal{B} = \{b_1, b_2\}$. Find the $\mathcal{B}$-coordinate vector of $\mathbf{x}$. (Think of this as using a dictionary to translate words in your language to a foreign language.)

$b_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \quad b_2 = \begin{bmatrix} -3 \\ 5 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} -7 \\ 13 \end{bmatrix}$

Solution. Here we need to find the linear combination, so we need to find c_1, c_2 such that

$$c_1 \begin{bmatrix} 1 \\ -3 \end{bmatrix} + c_2 \begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} -7 \\ 13 \end{bmatrix}$$

Solving this system by row reducing or by inspection yields $c_1 = 1, c_2 = -2$. So

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

■

7. Find a basis for the following subspaces, and determine the dimension.

(a) Symmetric 2×2 matrices

Solution. A generic symmetric 2×2 matrix is of the form

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

The *only* necessary requirement is that the 12 and 21 entry are the same. The diagonal entries don't matter at all. We can write this generic element as

$$= a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Which gives us a basis

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

This is a basis because it definitely spans the space, and the representations are unique (equivalently, they are clearly linearly independent). ■

(b) Vector in $\mathbb{R}^3$ such that $x_1 - 2x_2 + x_3 = 0$

Solution. We can see that $x_1 = 2x_2 - x_3$.

If we call this subspace W, then a generic element of this subspace must be of the form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_2 - x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Which gives one potential basis of

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

This is not the only basis possible. Other possible solutions can be found by solving for x_3, or considering this a 1×3 system of equations (but this list is certainly not exhaustive). ■

8. In one sentence, each, explain why each set is or is not a basis of $\mathbb{R}^3$. No scratch work is required for any of these.

   Remark 37. To solve this, we are going to use our trick that we need the "right number of vectors" (the dimension, which is 3 in this case), and that they are linearly independent. If either one of those is not satisfied, it's not a basis. If both are satisfied, then it is a basis. This is nice because counting and determining independence is much easier than verifying span.

   (a) $\begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 9 \end{bmatrix}$

   Solution. Clearly linearly independent because it has a triangular form, and there are the right number of vectors, so it is a basis. ■

   (b) $\begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 9 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Solution. No set with the zero vector is linearly independent, so this is not a basis. ■

(c) $\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\2\\4\end{bmatrix}, \begin{bmatrix}1\\3\\9\end{bmatrix}, \begin{bmatrix}1\\4\\16\end{bmatrix}$

Solution. Too many vectors. Not a basis. ■

(d) $\begin{bmatrix}1\\2\\-3\end{bmatrix}, \begin{bmatrix}-4\\-5\\6\end{bmatrix}$

Solution. Too few vectors. Not a basis. ■

9. Using a row reduction calculator, find a basis for the space spanned by the given vectors, $\mathbf{v}_1, \ldots, \mathbf{v}_5$.

$$\begin{bmatrix}1\\0\\-3\\2\end{bmatrix}, \begin{bmatrix}0\\1\\2\\-3\end{bmatrix}, \begin{bmatrix}-3\\-4\\1\\6\end{bmatrix}, \begin{bmatrix}1\\-3\\-8\\7\end{bmatrix}, \begin{bmatrix}2\\1\\-6\\9\end{bmatrix}$$

Solution. The row reduced matrix gives pivot columns 1,2,4. Therefore a basis for the span is $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ ■

10. Use coordinate vectors to test if the sets of polynomials

(i) Are linearly independent

(ii) Span $\mathbb{R}_2[x]$

(iii) Form a basis of $\mathbb{R}_2[x]$

Explain your work.

$$\begin{cases} x - 2x^2, \\ 1 - 2x^2, \\ 1 + 10x + x^2, \end{cases}$$

Solution. The coordinate vectors are

$$\begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 10 \\ 1 \end{bmatrix}$$

From here, we can determine these things in many ways. The fastest way here is to note that the first two vectors are clearly linearly independent. Thus, the only way for the whole set to be dependent is if we can write the third vector as a linear combination of the first two. We would need 1 of the first vector and 10 of the second. But

$$1 \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} + 10 \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 10 \\ 1 \end{bmatrix}$$

so this set is linearly independent.

Since $\mathbb{R}_2[x]$ is three dimensional, and we have three linearly independent vectors, we know that it spans the space and is a basis (by one of the provided theorems). ■

8.3 Eigenvectors Proper

11.

Definition 38. A nonzero vector which is merely scaled by a linear transformation is called an eigenvector (what we have been calling magical vectors). The factor by which it scales is called its eigenvalue. Ex. if

$$\boxed{T(\mathbf{v}) = 7\mathbf{v}}$$

and $\mathbf{v} \neq 0$, then $\mathbf{v}$ is an "eigenvector of T with eigenvalue 7".

Suppose we have a transformation $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ with three nonzero magical eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$

$$\begin{cases} T(\mathbf{v}_1) = \frac{1}{1}\mathbf{v}_1 \\ T(\mathbf{v}_2) = \frac{1}{2}\mathbf{v}_2 \\ T(\mathbf{v}_3) = \frac{1}{3}\mathbf{v}_3 \end{cases}$$

We are going to show the following facts:

- The eigenvectors are linearly independent because they have different eigenvalues
- The eigenvectors form a basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ we call the eigenbasis.
- This basis makes it very easy to calculate multiple applications of T.

(a) Suppose we have a linear combination that yields zero

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \vec{0} \tag{*}$$

i. Apply $T - I$ to both sides, and let that be a second equation.
Solution. Clearly $(T - I)\vec{0} = \vec{0}$, and we can also see that

$$(T - I)\mathbf{v}_j = \left(\frac{1}{j} - 1\right)\mathbf{v}_j$$

So we get

$$\vec{0} - \frac{1}{2}c_2\mathbf{v}_2 - \frac{2}{3}c_3\mathbf{v}_3 = \vec{0} \tag{**}$$

■

ii. Apply $T - \frac{1}{2}I$ to both sides of your second equation and let that be a third equation. What must c_3 be?

Solution. Similarly, $\left(T - \frac{1}{2}I\right)\mathbf{v}_j = \left(\frac{1}{j} - \frac{1}{2}\right)\mathbf{v}_j$ so

$$\vec{0} + -\frac{2}{3}c_3\left(-\frac{1}{6}\mathbf{v}_3\right) = \vec{0} \qquad (***)$$

$\implies c_3 = 0$ ■

iii. Based on what you found c_3 to be, what does your second equation imply about c_2?

Solution. In (∗∗), we can see that if $c_3 = 0$, then $c_2 = 0$. ■

iv. Conclude the value of c_1 and that $\mathcal{B}$ must be linearly independent.

Solution. Plugging in $c_2 = c_3 = 0$ in (*) gives $c_1 = 0$. Therefore, we have shown that the set is linearly independent. ■

v. Why is $\mathcal{B}$ a basis?

Solution. We have a linearly independent set of three vectors in $\mathbb{R}^3$ (dimension 3), so it's a basis. ■

(b) Based on the linear independence of $\mathcal{B}$, we can say

$$c_1\mathbf{v}_1 + \frac{1}{2}c_2\mathbf{v}_2 + \frac{1}{3}c_3\mathbf{v}_3 = \vec{0} \implies c_1 = c_2 = c_3 = 0$$

Explain how this proves that T is invertible.
Hint: What is $T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3)$?

Solution. If $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ is in $\ker(T)$ (since $\mathcal{B}$ is a basis, we know that any $\mathbf{x} \in \ker(T)$ can be written as a linear combination of the $\mathbf{v}$ vectors uniquely), then $T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3) = c_1\mathbf{v}_1 + \frac{1}{2}c_2\mathbf{v}_2 + \frac{1}{3}c_3\mathbf{v}_3 = \vec{0}$ But we're given that this implies $c_i = 0$ (since the vectors are linearly independent). So $\mathbf{x} = \vec{0}$ and thus all elements of the kernel are $\vec{0}$. Therefore, T has a trivial kernel, so it's invertible (since it's an operator, we can use the invertible matrix theorem or whatever). ■

8.3.1 Preimage of Eigenvectors

(c) Clearly, $\mathbf{v}_1$ is a preimage of itself. Find a preimage of $\mathbf{v}_2$ (hint: it's a scalar multiple of $\mathbf{v}_2$). Use the same logic to find a preimage of $\mathbf{v}_3$ without doing any scratch work.

Solution. We can see that if

$$T(k\mathbf{v}_2) = \frac{k}{2}\mathbf{v}_2 = \mathbf{v}_2 \implies k = 2 \implies 2\mathbf{v}_2 = \frac{\mathbf{v}_2}{\frac{1}{2}}$$

is a preimage of $\mathbf{v}_2$. Similarly, $3\mathbf{v}_3 = \frac{\mathbf{v}_3}{\frac{1}{3}}$ is a preimage of $\mathbf{v}_3$. That is, we just divide by the eigenvalue! This is also consistent with $\mathbf{v}_1 = \frac{\mathbf{v}_1}{1}$ ■

8.3.2 Diagonalization Introduction (Part II)

(d) Suppose that $\mathbf{x} \in \mathbb{R}^3$ and $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$.

i. Compute $[T(\mathbf{x})]_{\mathcal{B}}$

Solution. $T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3) = c_1\mathbf{v}_1 + \frac{1}{2}c_2\mathbf{v}_2 + \frac{1}{3}c_3\mathbf{v}_3$ has coordinate vector

$$[T(\mathbf{x})]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \frac{1}{2}c_2 \\ \frac{1}{3}c_3 \end{bmatrix}$$

■

ii. Compute $[T^n(\mathbf{x})]_{\mathcal{B}}$ (where $n \geq 1$).

Solution. $[T^n(\mathbf{x})]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \left(\frac{1}{2}\right)^n c_2 \\ \left(\frac{1}{3}\right)^n c_3 \end{bmatrix}$ ■

iii. Does your formula above work for $n = 0$ (assuming $T^0 := I$)? What about negative values of n?

Solution. Yes. For $n = 0$, we retrieve the original $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$, for $n = -1$, we get an answer consistent with our preimages $[T^{-1}(\mathbf{v})]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ 2c_2 \\ 3c_3 \end{bmatrix}$, and we can see that this will work for all $n \in \mathbb{N}$. ■

iv. Based on what we've done, show that

$$[T^n(\mathbf{x})]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}^n [\mathbf{x}]_{\mathcal{B}}$$

Solution.

$$\begin{aligned} [T^n(\mathbf{x})]_{\mathcal{B}} &= \begin{bmatrix} 1^n c_1 \\ \left(\frac{1}{2}\right)^n c_2 \\ \left(\frac{1}{3}\right)^n c_3 \end{bmatrix} \\ &= \begin{bmatrix} 1^n & 0 & 0 \\ 0 & (\frac{1}{2})^n & 0 \\ 0 & 0 & (\frac{1}{3})^n \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}^n [\mathbf{x}]_{\mathcal{B}} \end{aligned}$$

■

8.3.3 Markov Chains Introduction

v. What is $\lim_{n\to\infty} T^n(\mathbf{x})$? Conclude that you only need to know the coordinate on $\mathbf{v}_1$ to know the "end behavior" of repeated application of T.

Solution. As $n \to \infty$, $\left(\frac{1}{2}\right)^n, \left(\frac{1}{3}\right)^n \to 0$ so

$$[T^n(\mathbf{x})]_{\mathcal{B}} \to \begin{bmatrix} c_1 \\ 0 \\ 0 \end{bmatrix}$$

Thus,

$$T^n(\mathbf{x}) \to c_1\mathbf{v}_1$$

and $c_2, c_3, \mathbf{v}_2, \mathbf{v}_3$ have no impact on the limit. Therefore, we need only find $\mathbf{v}_1$ and c_1 to know the limiting behavior. ■

(e) The following matrix DOES have eigenvalues $1, \frac{1}{2}, \frac{1}{3}$.

$$A = \begin{bmatrix} 2/3 & 1/3 & 0 \\ 1/6 & 1/2 & 1/3 \\ 1/6 & 1/6 & 2/3 \end{bmatrix}$$

This is actually a very special type of matrix called a <u>Markov</u> or <u>Stochastic</u> matrix, which is used in a number of applications.

An eigenvector of A with eigenvalue 1, called "the steady state vector" is $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Verify this.

Solution.

$$\begin{bmatrix} 2/3 & 1/3 & 0 \\ 1/6 & 1/2 & 1/3 \\ 1/6 & 1/6 & 2/3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 + 1/3 \\ 1/6 + 1/2 + 1/3 \\ 1/6 + 1/6 + 2/3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \checkmark$$

■

(f) Suppose we have three initial vectors, $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ with coordinate vectors

- $[\mathbf{x}_1]_{\mathcal{B}} = \begin{bmatrix} 1/3 \\ 0 \\ 0.27 \end{bmatrix}$
- $[\mathbf{x}_2]_{\mathcal{B}} = \begin{bmatrix} 1/3 \\ \pi/8 \\ 0.13 \end{bmatrix}$
- $[\mathbf{x}_3]_{\mathcal{B}} = \begin{bmatrix} 1/3 \\ 0.12315 \\ 3/7 \end{bmatrix}$

with respect to an eigenbasis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ (where $\mathbf{v}_2$ is an eigenvector of A with eigenvalue $\frac{1}{2}$ and $\mathbf{v}_3$ is an eigenvector of A with eigenvalue $\frac{1}{3}$ just like before).

Find $\lim_{n\to\infty} T^n(\mathbf{x}_j)$ for $j = 1, 2, 3$.

Solution. We know the limit will be $c_1\mathbf{v}_1$ by part d v., and $c_1 = \frac{1}{3}$ in all three cases. Therefore, the limits are all

$$\lim_{n\to\infty} T^n(\mathbf{x}_j) = \frac{1}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix}$$

■

Remark 39. We can conclude that this matrix, over time, spreads out the entries evenly. Something not entirely obvious just by looking at it. In fact, we can say

$$\lim_{n\to\infty}\begin{bmatrix}2/3 & 1/3 & 0\\1/6 & 1/2 & 1/3\\1/6 & 1/6 & 2/3\end{bmatrix}^n = \begin{bmatrix}1/3 & 1/3 & 1/3\\1/3 & 1/3 & 1/3\\1/3 & 1/3 & 1/3\end{bmatrix}$$

We can also conclude that all probability vectors (vectors with positive entries that add up to 1) will approach $\begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix}$ in this Markov chain.

Week #10 Worksheet Solutions

9.1 Rank Nullity

1. Suppose A is a 3×8. If the rank of A is 2, then what is

   (a) the rank of A^T?

   (b) the nullity of A?

   (c) the nullity of A^T?

   Solution. The rank of A and A^T are always the same, so $\text{rank}(A^T) = 2$.

   A has 8 columns, so the rank and nullity must add up to 8. If A has rank 2, then it must have nullity 6.

   A^T has 3 columns, so the nullity and rank must add up to 3. Rank is 2, so nullity is 1. ■

9.2 Eigenvectors and Eigenvalues

Remark 40. A nice shorthand for the words "eigenvalue" and "eigenvector" are "ew" and "ev" respectively. This comes from the German origin of the words

$$\text{Eigenwert (ew)} \iff \text{Eigenvalue}$$
$$\text{Eigenvektor (ev)} \iff \text{Eigenvector}$$

9.2.1 Preimages of Eigenvectors

2. Suppose that $\mathbf{v} \neq 0$ is an eigenvector of T with eigenvalue λ.

$$T(\mathbf{v}) = \lambda\mathbf{v}$$

(a) Find a simple formula for a preimage of $\mathbf{v}$ under T in terms of $\mathbf{v}$ assuming $\lambda \neq 0$. Does it work for $\lambda = 0$?

<u>*Solution.*</u> Just divide by the eigenvalue! We saw this in (8.3.1).

$$\mathbf{x} = \frac{\mathbf{v}}{\lambda}$$

This is easy to derive. If we suppose our preimage is $k\mathbf{v}$, then

$$\mathbf{v} = T(k\mathbf{v}) = k\lambda\mathbf{v} \implies k = \frac{1}{\lambda}$$

No, this does not work for $\lambda = 0$. If $\lambda = 0$, then there is no simple preimage in terms of $\mathbf{v}$ (in fact, you should hope there is no preimage at all! If a nonzero kernel vector is in the image, that means the operator is trouble (nondiagonalizable)).

If you are interested in why this is the case, suppose that there is an eigenbasis for T, and suppose $\mathbf{x} \in \ker(T) \cap \operatorname{im}(T)$, and show that $\mathbf{x} = \vec{0}$. ■

(b) Show that this is equivalent to $\mathbf{v} \in \ker(T - \lambda I)$.

We call $E_\lambda(T) := \ker(T - \lambda I)$ the <u>eigenspace</u> of T with eigenvalue λ, which is the subspace that contains all eigenvectors with eigenvalue λ (and $\vec{0}$).

<u>*Solution.*</u>

$$\begin{aligned} T\mathbf{v} = \lambda\mathbf{v} \iff T\mathbf{v} - \lambda I\mathbf{v} = \vec{0} \iff (T - \lambda I)\mathbf{v} = \vec{0} \\ \iff \mathbf{v} \in \ker(T - \lambda I) \iff \mathbf{v} \in E_\lambda(T) \end{aligned}$$

■

9.2.2 Polynomials in Operators

(c) Show that $\mathbf{v}$ is also an eigenvector of T^2 with eigenvalue λ^2.

<u>*Solution.*</u>

$$T^2(\mathbf{v}) = T(T\mathbf{v}) = T(\lambda\mathbf{v}) = \lambda T(\mathbf{v}) = \lambda\lambda\mathbf{v} = \lambda^2\mathbf{v}$$

■

(d) Use the same logic to conclude that $\mathbf{v}$ is an eigenvector of T^n with eigenvalue λ^n (if $n \geq 0$).

<u>*Solution.*</u> If applying T scales by λ, then applying T n times is the same as scaling by λ n times. So

$$T^n(\mathbf{v}) = \lambda^n\mathbf{v}$$

■

Remark 41. This actually works for all n if T is invertible.

(e) Given a polynomial $p(x) = p_0 + p_1x + \ldots + p_nx^n$, we define the polynomial operator

$$\begin{aligned} &p(T) := p_0I + p_1T + \ldots + p_nT^n \\ &\qquad\implies p(T)\mathbf{v} := p_0\mathbf{v} + p_1T(\mathbf{v}) + \ldots + p_nT^n(\mathbf{v}) \end{aligned}$$

Show that $\mathbf{v}$ is an eigenvalue of $p(T)$ with eigenvalue $p(\lambda)$.

<u>*Solution.*</u> We can write $p(x) = \sum_{j=0}^{n} p_nx^n$. So $p(T)\mathbf{v} = \sum_{j=0}^{n} p_nT^n(\mathbf{v})$. Then,

$$p(T)\mathbf{v} = \sum_{j=0}^{n} p_n\lambda^n\mathbf{v} = \left(\sum_{j=0}^{n} p_n\lambda^n\right)\mathbf{v} = p(\lambda)\mathbf{v}$$

■

(f) If $p(\lambda) \neq 0$, use part (a) to find a preimage of $\mathbf{v}$ under $p(T)$.

Solution. Divide by the eigenvalue!

$$\mathbf{x} = \frac{\mathbf{v}}{p(\lambda)}$$

■

(g) Suppose that the polynomial p has a root at $x = \lambda$ (i.e. $p(\lambda) = 0$). Show that $\mathbf{v} \in \ker(p(T))$.

Solution. If $p(\lambda) = 0$, then

$$p(T)\mathbf{v} = p(\lambda)\mathbf{v} = 0\mathbf{v} = \vec{0} \implies \mathbf{v} \in \ker(p(T))$$

■

(h) Suppose that $p(x) = (x-1)(x-2)$ and $\mathbf{v}_1$ is an eigenvector of T with eigenvalue 1 and $\mathbf{v}_2$ is an eigenvector of T with eigenvalue 2. Show that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \in \ker(p(T))$$

Solution. $p(T)\mathbf{v}_1 = p(1)\mathbf{v}_1 = 0\mathbf{v}_1 = \vec{0}$. Since $p(1) = (1-1)(1-2) = 0$. Thus, it's in the kernel. Similar for $\mathbf{v}_2$ since $p(2) = 0$. Since the kernel is a subspace (closed under linear combinations), then any linear combination of $\mathbf{v}_1, \mathbf{v}_2$ is in the kernel. ■

9.3 Differential Equations Introduction (Part II)

3. A couple weeks ago (7.3.2), we looked at the differential (derivative) operator on the space of smooth functions $D\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}), \quad D(y) = y'$, and we found that e^t was an eigenvector with eigenvalue 1.

(a) Show that $e^{\lambda t}$ is an eigenvector of D with eigenvalue λ, for any $\lambda \in \mathbb{R}$.

Solution. $D\left(e^{\lambda t}\right) = \frac{\mathrm{d}}{\mathrm{d}t} e^{\lambda t} = \lambda \left(e^{\lambda t}\right)$
Thus, applying D just scales it by λ.
Since $e^{\lambda t} \neq 0$ (it's not just the zero function), then it's an eigenvector. Also, this works for any λ. ■

9.3.1 Polynomial Differential Operators

(b) The most important form of differential equation is $p(D)y$. $p(D)$ is called a "polynomial differential operator", and it describes a lot of physical systems like masses on a spring, circuits, kinematics, and more. Using the previous problem 9.2.2 on the worksheet, explain why $e^{\lambda t}$ is an eigenvector of $p(D)$ with eigenvalue $p(\lambda)$.

$$p(D)e^{\lambda t} = p(\lambda)e^{\lambda t}$$

Solution. Since $e^{\lambda t}$ is an eigenvector of D, it's an eigenvector of $p(D)$ (as we saw in the previous question). With eigenvalue $p(\lambda)$.

$$p(D)e^{\lambda t} = p(\lambda)e^{\lambda t}$$

■

(c) Using question 2h (in the previous problem), find two independent nonzero solutions y_1, y_2 to

$$(D-1)(D-2)y = 0$$

in terms of eigenvectors of D. Observe that

$$y = c_1 y_1 + c_2 y_2$$

is also a solution (since the kernel is a subspace).

Solution. Like we saw before, if we have an eigenvector of T with eigenvalue $1, 2$, then it's in the kernel of $(T - I)(T - 2I)$. This is the same for $T = D$, so we just take eigenvector of D with eigenvalue 1 and 2 (e^{1t} and e^{2t}) then those are two solutions (which are independent).

Additionally, any linear combination will be in the kernel since it is a subspace.

$$y = c_1 e^t + c_2 e^{2t} \in \ker(p(D))$$

So this gives an infinite family of solutions. ■

9.3.2 Exponential Response Formula

(d) The "exponential response formula" (ERF) is an amazingly useful technique in differential equations, and with linear algebra we can prove it in one line using part (3b). Show that if $p(\alpha) \neq 0$, then $y = \frac{Be^{\alpha t}}{p(\alpha)}$ is a particular solution to

$$p(D)y = Be^{\alpha t}$$

Use the ERF to find one solution to

$$y'' - \sqrt{2}y' + \sqrt{10}y = 5e^{\sqrt{5}t}$$

Hint: $p(x) = x^2 - \sqrt{2}x + \sqrt{10}$. What is the eigenvalue of $5e^{\sqrt{5}t}$?

Solution. Divide by the eigenvalue! In this case, that's $p(D)e^{\alpha t} = p(\alpha)e^{\alpha t}$ so the eigenvalue is $p(\alpha)$. Thus,

$$y_p = \frac{Be^{\alpha t}}{p(\alpha)}$$

For the given ODE, the RHS is $5e^{\sqrt{5}t}$, which has eigenvalue $\sqrt{5}$ (under D). Thus, the eigenvalue of $p(D)$ is

$p(\sqrt{5}) = \sqrt{5}^2 - \sqrt{2}\sqrt{5} + \sqrt{10} = 5$. So a particular solution is given by dividing by 5.

$$y_p = \frac{5e^{\sqrt{5}}}{5} = e^{\sqrt{5}t}$$

You can plug it in and see that it is indeed a solution. ■

(e) Use the logic of the previous parts and the ERF to find a two parameter family of solutions to

$$(D-1)(D-3)y = y'' - 4y' + 3y = e^{2t}$$

and solve for the constants to satisfy $y(0) = 1$, $y'(0) = 2$
Hint: Use the ideas of parts (c) and (d) to find a general form for y (you should get two constants) and then plug in $t = 0$ for y and y' to solve for the initial conditions (you'll get a 2×2 system).

Solution. The ERF tells us a particular solution is

$$y_p = \frac{e^{2t}}{p(2)} = \frac{e^{2t}}{-1} = -e^{2t}$$

The roots are $1, 3$, so we can just take eigenvectors of D with eigenvalue 1 and 3 to get the kernel (i.e. e^t and e^{3t})

$$y_h = c_1 e^t + c_2 e^{3t}$$

Thus the solution is of the form

$$y = c_1 e^t + c_2 e^{3t} - e^{2t}$$

We want to satisfy the initial conditions so we plug in 0 to y and y':

$$y(0) = c_1 + c_2 - 1 = 1 \implies c_1 + c_2 = 2$$

and $y' = c_1 e^t + 3c_2 e^{3t} - 2e^{2t}$ so

$$y'(0) = c_1 + 3c_2 - 2 = 2 \implies c_1 + 3c_2 = 4$$

One can find the solution is $c_1 = c_2 = 1$. Thus, the solution is

$$y = e^t + e^{3t} - e^{2t}$$

You can plug it in to verify it satisfies everything. Wolfram Alpha confirms this solution is correct. ■

9.4 Bonus Solutions: Solving $p(D)y = 0$

The following were never in a worksheet (this quarter), but if you are curious about how we can use linear algebra to solve $p(D)y = 0$ in general, you can read through the following solutions and/or this blog post.

9.4.1 The Wronskian

(f) In differential equations, we call two functions $y_1(t), y_2(t)$ linearly (in)dependent on some interval if there is some point t_0 in the interval such that

$$\det \begin{bmatrix} y_1(t_0) & y_2(t_0) \\ y_1'(t_0) & y_2'(t_0) \end{bmatrix} \neq 0$$

This is called the Wronskian.

Using the point $t_0 = 0$, show that $y_1 = e^{at}$ and $y_2(t) = e^{bt}$ are linearly independent if and only if $a \neq b$.

Solution.

$$\det \begin{bmatrix} e^{at} & e^{bt} \\ ae^{at} & be^{bt} \end{bmatrix} = e^{(a+b)t}(b-a)$$

At $t = 0$, we get $b - a$. The result is nonzero if and only if $a \neq b$, so they are independent on $\mathbb{R}$ if and only if $a \neq b$. If $b = a$, then the result is actually zero everywhere (so they are dependent on all of $\mathbb{R}$ (not surprising, since they would be identical!)). ■

(g) Show that e^{it} and e^{-it} are solutions to $y'' + y = 0$.

Solution. $\pm i$ are indeed roots of $D^2 + 1 = (D+i)(D-i)$. So $e^{\pm it}$ are solutions. You can also plug them in to verify. ■

(h) Show that $\cos(t)$ and $\sin(t)$ are eigenvectors of D^2 with eigenvalue -1. Use this to get another pair of solutions to

$$y'' + y = 0, \quad y(0) = a, \quad y'(0) = b$$

Solution. The derivative of $\cos(t)$ is $-\sin(t)$, and the derivative of $\cos(t)$ is $-\sin(t)$. Thus,

$$D^2 \cos(t) = -\cos(t) \implies (D^2 + 1)\cos(t) = 0$$

Similarly, $D^2 \sin(t) = -\sin(t) \implies (D^2 + 1)\sin(t) = 0$. So

$$y = c_1 \cos(t) + c_2 \sin(t) \in \ker(D^2 + 1)$$

so this is a solution to $y'' + y = 0$. Similar to previous parts, we can solve for the constants to solve $y(0) = a, \quad y'(0) = b$. We get

$$y(0) = c_1 = a$$

and

$$y'(0) = c_2 = b$$

So the solution is

$$y = a\cos(t) + b\sin(t)$$

■

9.4.2 Euler's Formula

(i) It turns out the town of $\ker(D^2+1)$ ain't big enough for the two of these solution sets $\left\{e^{it}, e^{-it}\right\}$ and $\{\cos(t), sin(t)\}$. That is, they are linearly dependent! This is based on a more advanced existence and uniqueness theorem for linear differential equations that says an nth order equation has n linearly independent solutions (which is based on linear algebra!).

Find the coefficients on $\cos(t)$ and $\sin(t)$ to form e^{it} and e^{-it}.

Hint: Suppose $y = c_1 \cos(t) + c_2 \sin(t) = e^{\pm it}$. And make sure $y(0)$ and $y'(0)$ are consistent with the values for $e^{\pm it}$.

Solution.

$$y(0) = c_1 = e^{\pm i0} = e^0 = 1 \implies c_1 = 1$$

$y' = -c_1 \sin(t) + c_2 \cos(t) = \pm i e^{\pm it}$, so

$$y'(0) = c_2 = \pm i$$

Thus, we apparently get

$$e^{\pm it} = \cos(t) \pm i \sin(t)$$

This gives us **Euler's Formula**,

$$e^{it} = \cos(t) + i \sin(t)$$

■

9.4.3 Repeated Roots in Differential Equations

(j) Show that $t \in \ker(D^2)$. Extend this logic to show

$$\left\{1, t, t^2, \ldots, t^{n-1}\right\} \in \ker(D^n)$$

Solution. $D^2 t = D1 = 0$. We can conclude similarly that t^k becomes zero after $k+1$ derivatives. So all positive integer powers of t strictly below n will be sent to zero after n derivatives. ■

9.4.4 Exponential Shift

(k) Show that $D(e^{\lambda t}y) = e^{\lambda t}(D+\lambda)y$.

Solution.

$$\begin{aligned} D(e^{\lambda t}y) = e^{\lambda t}y' + \lambda e^{\lambda t}y = e^{\lambda t}(y' + \lambda y) \\ = e^{\lambda t}(Dy + \lambda y) = e^{\lambda t}(D+\lambda)y \end{aligned}$$

■

(l) It's possible to prove through induction that

$$D^n(e^{\lambda t}y) = e^{\lambda t}(D+\lambda)^n y$$

Use this to conclude that

$$p(D)(e^{\lambda t}y) = e^{\lambda t}p(D+\lambda)y$$

Solution. If $p(x) = \sum_{n=0}^{\infty} p_n x^n$, then

$$\begin{aligned} p(D)(e^{\lambda t}y) = \sum_{n=0}^{\infty} p_n D^n(e^{\lambda t}y) = \sum_{n=0}^{\infty} p_n e^{\lambda t}(D+\lambda)^n y \\ = e^{\lambda t}\left(\sum_{n=0}^{\infty} p_n (D+\lambda)^n\right) y = e^{\lambda t}p(D+\lambda)y \end{aligned}$$

■

Remark. An nth order linear ODE (that is, one with an nth derivative $y^{(n)}$ (or when $p(D)$ is degree n) has n linearly independent homogeneous solutions.

(m) If we try to use our previous methods to solve

$$y''' + 3y'' + 3y' + y = (D+1)^3 y = 0$$

we run into a problem trying to find more than one independent solution. We call this a "repeated root".

Sure, we get $y = e^{-t}$ since -1 is a root. But what about the other two we should get? We can use the last few parts to help us out. If we suppose $y = e^{-t}u$, then we get

$$(D+1)^3 e^{-t}u = e^{-t}(D+1-1)^3 u = e^{-t}D^3 u = 0$$

Use the previous parts to show that

$$y = e^{-t}(c_1 + c_2 t + c_3 t^2)$$

is a solution to $y''' + 3y'' + 3y' + y = 0$.

Solution. $e^{-t}D^3u = 0$ implies that $u \in \ker(D^3)$. We know that $\{1, t, t^2\} \in \ker(D^3)$, so that gives us three solutions: $u = 1, t, t^2$. Therefore, we get three solutions

$$y = e^{-t}1, e^{-t}t, e^{-t}t^2 \in \ker\left((D+1)^3\right)$$

So we can take a linear combination to get three independent solutions

$$y = e^{-t}(c_1 + c_2 t + c_3 t^2)$$

■

9.4.5 Complex Roots in Differential Equations

(n) In the theory of differential equations, when we have complex roots to $p(D) \in \mathbb{R}[x]$, such as $\lambda = \alpha \pm i\beta$, then the solutions we pick are not $y = e^{(\alpha \pm i\beta)t}$, but

$$y = e^{\alpha t}\cos(\beta t), e^{\alpha t}\sin(\beta t)$$

Note that by Euler's formula, these are just the real and imaginary parts of $e^{(\alpha+\beta i)t}$. Verify that

$$e^{\alpha t}\cos(\beta t), e^{\alpha t}\sin(\beta t) \in \ker\left((D-\alpha)^2 + \beta^2\right)$$

Hint: You can use that $\cos(\beta t), \sin(\beta t) \in \ker(D^2 + \beta^2)$

Solution.

$$\left((D-\alpha)^2+\beta^2\right)e^{\alpha t}\cos(\beta t) = e^{\alpha t}\left(D^2+\beta^2\right)\cos(\beta t) = 0$$

Similarly,

$$\left((D-\alpha)^2+\beta^2\right)e^{\alpha t}\sin(\beta t) = e^{\alpha t}\left(D^2+\beta^2\right)\sin(\beta t) = 0$$

■

Remark 42. We can combine all of this to solve $p(D)y = 0$ in general. If the distinct roots of p are $\lambda_1, \ldots, \lambda_k$ and

$$p(D) = (D-\lambda_1)^{m_1}\ldots(D-\lambda_k)^{m_k}$$

then a general solution is

$$y = e^{\lambda_1 t}\left(c_{11}+c_{12}t+\ldots+c_{1(m_1-1)}t^{m_1-1}\right) + \ldots + e^{\lambda_k t}\left(c_{k1}+c_{k2}t+\ldots+c_{k(m_k-1)}t^{m_k-1}\right)$$

And if those roots are real, then this is the solution we use. If there are complex roots, then we adapt it similarly. If you have a root $\alpha \pm i\beta$ repeated m times, (that is, $((D-\alpha)^2+\beta^2)^m$ divides $p(D)$, then we get $2m$ solutions from

$$y = e^{\alpha t}\left[\left(c_{11}+c_{12}t+\ldots+c_{1(m-1)}t^{m-1}\right)\cos(\beta t) \left(c_{21}+c_{22}t+\ldots+c_{2(m-1)}t^{m-1}\right)\sin(\beta t)\right]$$

9.5 Eigenvalue / Eigenvector Shortcuts

Here's a blog post on some general tricks, but these next few problems show some of the results in that blog post.

4. Show that for a 2×2 matrix, the "characteristic polynomial" $\det(A - \lambda I)$ is

$$\det(A - \lambda I) = \lambda^2 - \operatorname{tr}(A)\lambda + \det(A) \tag{10}$$

Recall that the trace $\operatorname{tr}(A)$ is the sum of the diagonal entries.

Hint: use the generic matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

<u>*Solution*</u>.

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - (a + d)\lambda + ad - bc = \lambda^2 - \operatorname{tr}(A)\lambda + \det(A) \end{aligned}$$

■

9.5.1 Trace and Determinant

5. (a) Using the above formula for the characteristic polynomial of a 2×2, show that the eigenvalues λ_1, λ_2 of A must satisfy

$$\begin{aligned} \lambda_1 + \lambda_2 &= \operatorname{tr}(A) \\ \lambda_1 \lambda_2 &= \det(A) \end{aligned}$$

Hint: Use that the characteristic polynomial must factor into $(\lambda - \lambda_1)(\lambda - \lambda_2)$ and equate coefficients for the formula in the previous problem.

<u>*Solution*</u>.

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2$$

If

$$\begin{array}{cccccccc} & \lambda^2 & - & \operatorname{tr}(A) & \lambda & + & \det(A) \\ = & \lambda^2 & - & (\lambda_1 + \lambda_2) & \lambda & + & \lambda_1\lambda_2 \end{array}$$

Then we clearly need

$$\lambda_1 + \lambda_2 = \operatorname{tr}(A)$$
$$\lambda_1 \lambda_2 = \det(A)$$

■

9.5.2 Black Magic

(b) Consider the matrix $A = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}$

i. Explain why the eigenvalues are $\lambda = 3, -1$, *without* computing the characteristic polynomial, using the trace and determinant (use question 5).
Hint: What are two numbers that add up to the trace and multiply to the determinant?

Solution. Trace is 2 and determinant is -3. Two numbers that add up to 2 and multiply to -3 are $3, -1$. ■

ii. Find the eigenvectors for $\lambda = 3$ and $\lambda = -1$.

Solution.

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

If we take 1 of column 2 and 3 of column 1, then that will give zero. So $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue 3.

$$A + I = \begin{bmatrix} 3 & 3 \\ 1 & 1 \end{bmatrix}$$

If we take 1 of column 2 and -1 of column 1 then we get zero, so $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue -1. ■

iii. Find a basis for the column space of $A-3I$ and $A+I$. How do the bases relate to the eigenvectors you found in the previous part?

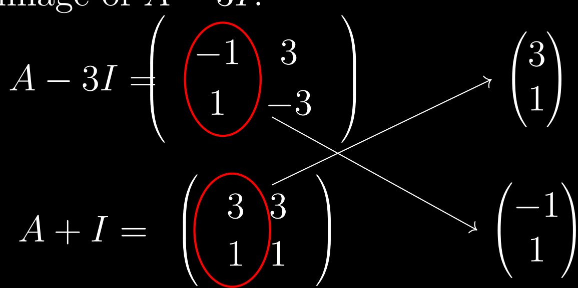
<u>*Solution.*</u> You can go through the steps of row reduction and taking the pivot columns (there will always be one), but we can observe that since nullity is 1 for both, then rank is 1 for both, then any nonzero column (like the first in this particular case) will always be a basis for the column space (in general, any nonzero vector in a dimension 1 subspace is a basis for that subspace). So

$$\mathrm{Col}(A-3I) = \mathrm{span}\left\{\begin{bmatrix}-1\\1\end{bmatrix}\right\}, \quad \mathrm{Col}(A+I) = \mathrm{span}\left\{\begin{bmatrix}3\\1\end{bmatrix}\right\}$$

and

$$\ker(A-3I) = \mathrm{span}\left\{\begin{bmatrix}3\\1\end{bmatrix}\right\}, \quad \ker(A+I) = \mathrm{span}\left\{\begin{bmatrix}-1\\1\end{bmatrix}\right\}$$

Wait... they're just switched! The eigenvector for 3 generates the image of $A+I$ and the eigenvector for -1 generates the image of $A-3I$.



■

Remark 43. No, this is not a coincidence. This always works for 2×2 matrices that aren't just scalar multiples of the identity. Any nonzero column of $A-\lambda_1 I$ will be an eigenvector with eigenvalue λ_2 and vice versa.
This works in general for all matrices with a minimal polynomial of degree exactly 2.
I call this the <u>eigenvector columns theorem</u>.

Here's a quick proof for why this is true (using some advanced linear algebra) if you're curious.

Proof. The characteristic polynomial of A is $x^2 - 2x - 3 = (x-3)(x+1)$. By the Cayley Hamilton theorem (See (5.4)), A satisfies

$$(A - 3I)(A + I) = 0$$

Thus, the images of $A+I$ are contained in the kernel of $A-3I$ (the eigenspace of 3). So any nonzero column of $A+I$ is an eigenvector with eigenvector 3. Similarly, since

$$(A - 3I)(A + I) = (A + I)(A - 3I) = 0$$

then the image of $A-3I$ is contained in the kernel of $A+I$, so any nonzero column of $A+I$ is an eigenvector with eigenvalue -1. □

(c) In general, the trace is always the sum of the eigenvalues, and the determinant is always the product for any square matrix.

If you can find n numbers that add up to the trace and multiply to the determinant of an $n \times n$ matrix, those could be the eigenvalues. For a 2×2 those *will* be the eigenvalues without exception. The solution to that system of equations is unique (and given exactly by (10)).

Find the eigenvalues of the matrix

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

Hint: $\det(A) = 6$. What three numbers add up to the trace and multiply to the determinant?

Solution. Three numbers that add up to the trace $1 + 4 + 1 = 6$ and multiply to 6 are $1, 2, 3$. And those are in fact the eigenvalues. We can know this for sure since 1 is clearly an eigenvalue by inspection (we can see $A\mathbf{e}_1 = \mathbf{e}_1$, so A definitely has eigenvalue 1 (and $\mathbf{e}_1$ is an eigenvector!). ■

9.6 Inner Products on Complex Vector Spaces

6. We showed a few weeks ago (6.4) that $\|\mathbf{v}\|^2 = \mathbf{v}^T\mathbf{v}$ was not a good way to define magnitudes when vectors can have complex entries. For example, before, we saw $\begin{bmatrix} 1 \\ i \end{bmatrix}^T \begin{bmatrix} 1 \\ i \end{bmatrix} = 1^2 + i^2 = 0$. Which gives a "magnitude" of 0 even though the vector is nonzero! The way to fix this is to use the "conjugate transpose", or "adjoint"

$$\mathbf{v}^* := \overline{\mathbf{v}^T}$$

where $\overline{\mathbf{v}^T}$ defines the transpose of the vector such that each entry is changes to its conjugate (note this means that if $\mathbf{v}$ is real, then $\mathbf{v}^* = \mathbf{v}^T$). For this reason, I call the conjugate transpose * the one true transpose (pretty much all the nice properties you get involving the transpose with real vectors/matrices are generalized with the conjugate transpose when you move to complex stuff).

For a complex number, the magnitude is given by

$$|a + bi|^2 = a^2 + b^2 = (a - bi)(a + bi)$$

Therefore, we can say $|z|^2 = \overline{z}z$, or $|z|^2 = z^*z$ if we allow $\overline{z} = z^*$ for scalars. We generalize this to vectors by

$$\|\mathbf{v}\|^2 = \mathbf{v}^*\mathbf{v}$$

Which has the effect of defining

$$\begin{aligned} \left\| \begin{bmatrix} a_1 + b_1 i \\ \vdots \\ a_n + b_n i \end{bmatrix} \right\|^2 &= \begin{bmatrix} (a_1 - b_1 i) & \cdots & (a_n - b_n i) \end{bmatrix} \begin{bmatrix} a_1 + b_1 i \\ \vdots \\ a_n + b_n i \end{bmatrix} \\ &= a_1^2 + b_1^2 + \ldots + a_n^2 + b_n^2 \\ &= |a_1 + b_1 i|^2 + \ldots + |a_n + b_n i|^2 \end{aligned}$$

This is consistent with our idea of magnitude being the sum of the squares of each entry. For complex numbers we just need to be a bit more careful about how we get those "squares". Note, also, this means that all magnitudes are real.

For our vector $\begin{bmatrix} 1 \\ i \end{bmatrix}$,

$$\begin{bmatrix} 1 \\ i \end{bmatrix}^* \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} 1 & -i \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = 1^2 + (-i)(i) = 1 + 1 = 2$$

which gives a much more reasonable magnitude for $\begin{bmatrix} 1 \\ i \end{bmatrix}$ of $\sqrt{2}$.

9.6.1 Eigenvalues of Orthogonal / Symmetric Matrices

We define the adjoint (conjugate transpose) A^* for a matrix A in a similar way. $A^* = \overline{A^T}$ is just the conjugate of the transpose! You may need to use the following facts in this problem: ($\mathbf{v}$ a vector and α is a scalar)

(i) $(\mathbf{v}^*)^* = \mathbf{v}$

(ii) $|\alpha|^2 = \overline{\alpha}\alpha$

(iii) $(\alpha\mathbf{v})^* = \overline{\alpha}\mathbf{v}^*$

(iv) If $\alpha = \overline{\alpha}$, then α is real.

(a) Recall a (real) matrix is called "orthogonal" if its inverse is its transpose ($Q^{-1} = Q^T = Q^*$). (See (6.3) and (6.5))

Show that if λ is an eigenvalue of Q, then $|\lambda| = 1$.

Hint: We showed previously that Q "preserves" angles and magnitudes. That is, $\|\mathbf{v}\|^2 = \|Q\mathbf{v}\|^2$.

Solution. Suppose $Q\mathbf{v} = \lambda\mathbf{v}$. Then we know

$$\|Q\mathbf{v}\|^2 = \|\mathbf{v}\|^2$$

But

$$\|Q\mathbf{v}\|^2 = (Q\mathbf{v})^* (Q\mathbf{v}) = (\lambda\mathbf{v})^* \lambda\mathbf{v} = \overline{\lambda}\mathbf{v}^*\lambda\mathbf{v} = |\lambda|^2\|\mathbf{v}\|^2$$

Thus,

$$|\lambda|^2\|\mathbf{v}\|^2 = \|\mathbf{v}\|^2 \implies |\lambda| = 1$$

■

Remark 44. This doesn't necessarily mean $\lambda = \pm 1$, since the eigenvalues can be complex (but, if they *are* real, then they are ± 1). But it does imply that all eigenvalue of an orthogonal (and, more generally, unitary transformation) are on the unit circle in the complex plane, and can be written as $\lambda_j = e^{i\theta_j}$.

One crazy fact is that this means if $U^{-1} = U^*$, then

$$U = e^{iH}$$

where $H^* = H$ and e^{iH} is the matrix exponential (see problem 7e).

(b) We showed before (6.6) that symmetric matrices have orthogonal eigenvectors (when their eigenvalues are different). We can also show their eigenvalues are all real too using a similar method.

Suppose that S is a real symmetric matrix (so $S = S^T = S^*$) and S has an eigenvector $\mathbf{v}$ with eigenvalue λ. Show λ is real by considering

$$\alpha = \mathbf{v}^* S\mathbf{v}$$

Hint: $\overline{\alpha} = \alpha^* = (\mathbf{v}^* S\mathbf{v})^* = \mathbf{v}^* S^*\mathbf{v}$. Since S is real and symmetric, $S^* = S$.

Solution.

$$\alpha = \mathbf{v}^*(S\mathbf{v}) = \mathbf{v}^*(\lambda\mathbf{v}) = \lambda\|\mathbf{v}\|^2$$

We have from the hint that

$$\overline{\alpha} = \mathbf{v}^* S^*\mathbf{v} = \mathbf{v}^* S\mathbf{v} = \alpha$$

Then, by the given property (iv), α is real! But then we have that $\alpha = \lambda\|\mathbf{v}\|^2$ is real. Since $\|\mathbf{v}\|^2$ is real, then λ must be real. ■

Remark 45. You may notice that what's really necessary for these problems are

- $Q^{-1} = Q^*$
- $S = S^*$

In general, if $U^{-1} = U^*$, then U is called "Unitary" (Orthogonal matrices are just the unitary matrices which are real), and if $H = H^*$, then H is called "Hermitian" (and symmetric matrices are just the Hermitian matrices which are real).

Another type is called "skew-hermitian" (or "skew-symmetric" if it's real), where $A^* = -A$ (or $A^T = -A$ if A is real). Then the eigenvalues of A are purely imaginary (or zero). In fact, one other interesting fact is that if n is odd, then any $n \times n$ skew-hermitian matrix is never invertible. This is because $\det(A) = \det(A^*) = \det(-A) = (-1)^n \det(A) = -\det(A) \implies \det(A) = 0$.

These general classes of matrices are often called "normal" and they are *amazing*. Some of the reasons they're so fascinating are covered in MATH132. But the main kicker is that they *always* have an orthonormal eigenbasis (at least over $\mathbb{C}$).

9.7 Diagonalization

7. Consider

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

(a) Find the eigenvectors and eigenvalues of A, and compute the diagonalization

$$A = PDP^{-1}$$

You may use the trace and determinant to find the eigenvalues, and the evil technique of problem 5b, or the fact that A is triangular (or whatever technique you want).

Solution. This matrix is triangular, so the diagonal entries are the eigenvalues (that's the faster optimal way for this problem). But otherwise, the trace is 0 and determinant is -1. Two numbers that add up to 0 and multiply to -1 are ± 1, so those are the eigenvalues.

Looking at either the columns or kernel of $A \pm I$ (and dividing by 2 if doing the columns) gives eigenvectors

$$\begin{cases} \lambda = 1 & \implies \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \lambda = -1 & \implies \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{cases}$$

The diagonalization is then given by putting the eigenvalues in a matrix $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ (putting -1 first is fine but you have to be careful with the order of the columns of P).

Then P's first column is the eigenvector with eigenvalue 1, and the second column is the eigenvector for the second eigenvalue: $P = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Thus,

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

■

(b) Show that

$$A^k = 1^k \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + (-1)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix}$$

Note: This implies $A = 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + (-1) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix}$.

Hint: Use the diagonalization! What is $(PDP^{-1})^k$? You may also use that

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}^k = a^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + b^k \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix}$$

Solution. We know $A^k = \left(PDP^{-1}\right)^k = PD^kP^{-1}$, so

$$A^k = P \begin{bmatrix} 1^k & 0 \\ 0 & (-1)^k \end{bmatrix} P^{-1}$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \left[1^k \mathbf{e}_1\mathbf{e}_1^T + (-1)^k \mathbf{e}_2\mathbf{e}_2^T\right] \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$
$$= 1^k \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + (-1)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix}$$

Selecting $k = 1$ shows the first requirement. ■

9.8 Systems of Differential Equations

(c) In differential equations, we are often concerned with problems that can be somehow "reduced" to a problem of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = A\mathbf{x}(t)$$

That is, finding vector functions $\mathbf{x}(t)$ where the derivative is the same as multiplying by the matrix A (the derivative of a vector function is calculated by differentiating each entry individually).

Verify that $\mathbf{x}_1(t) = e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{x}_2(t) = e^{-t} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are solutions to this equation.

Hint: For both, calculate $\mathbf{x}'(t)$ and $A\mathbf{x}$ and just confirm they are the same.

Solution. $\mathbf{x}_1' = \frac{\mathrm{d}}{\mathrm{d}t} e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} = e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$$A\mathbf{x}_1 = Ae^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} = e^t \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{x}_1' \quad \checkmark$$

Similarly, $\mathbf{x}_2' = \frac{\mathrm{d}}{\mathrm{d}t} e^{-t} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -e^{-t} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

$$A\mathbf{x}_2 = Ae^{-t} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = e^{-t} \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = e^{-t} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = -e^{-t} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$= \mathbf{x}_2' \quad \checkmark$ ■

Remark 46. In general, if $\mathbf{v}$ is an eigenvector of A with eigenvalue λ,

$$\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$$

is a solution to $\mathbf{x}'(t) = A\mathbf{x}(t)$. The idea is that taking the derivative only affects the exponential, which gets scaled by λ, while A only affects $\mathbf{v}$, which also gets scaled by λ. Thus, the action of taking the derivative is the same as multiplying by A.

(d) We are often concerned about the stability of the solutions to $\mathbf{x}' = A\mathbf{x}$. Will the solution explode to infinity, stay stable, or tend to zero? Clearly, if $\lambda > 0$, then the solution $\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$ will explode to infinity, since $e^{\lambda t} \to \infty$ as $t \to \infty$ when λ is positive. Using question 5, explain why if $\mathrm{tr}(A) > 0$ then the system is automatically *unstable* (that is, at least one solution will explode to infinity over time). For simplicity, assume the eigenvalues are real (though this is true even if they are complex).

Hint: explain why at least one eigenvalue must be positive, and then explain why that is sufficient.

Solution. Since the trace is the sum of eigenvalues, if the sum is positive, then at least one eigenvalue $\lambda > 0$ is positive. Then the solution corresponding to that eigenvalue $\mathbf{x} = e^{\lambda t}\mathbf{v}$ (where $\mathbf{v}$ is any eigenvector with eigenvalue λ) will explode to infinity, so the system is unstable. ∎

9.8.1 Matrix Exponentials: e^{At}

(e) Recall from calculus that $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$. We can compute the matrix exponential as

$$e^{At} := \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$$

Show that for $A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$,

$$e^{At} = e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + e^{-t} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix} = \begin{bmatrix} e^t & 0 \\ e^t - e^{-t} & e^{-t} \end{bmatrix}$$

Hint: Assume you can just distribute the sum.

Solution.

$$\begin{aligned} e^{At} &= \sum_{k=0}^{\infty} \frac{t^k}{k!} \left[1^k \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + (-1)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix} \right] \\ &= \left(\sum_{k=0}^{\infty} \frac{t^k}{k!} 1^k \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + \left(\sum_{k=0}^{\infty} \frac{t^k}{k!} (-1)^k \right) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix} \\ &= e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} + e^{-t} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix} = \begin{bmatrix} e^t & 0 \\ e^t - e^{-t} & e^{-t} \end{bmatrix} \end{aligned}$$

∎

Remark 47. As a matter of fact, $e^{At} = Pe^{Dt}P^{-1}$. And e^{Dt} is easy to calculate: it's just the exponential of each entry.

(f) In the same way that $\frac{\mathrm{d}}{\mathrm{d}t}e^{at} = ae^{at}$, matrix exponentials satisfy a similar property:

   i. Multiply Ae^{At}

   ii. Take the derivative $\frac{\mathrm{d}}{\mathrm{d}t}e^{At}$ (take the derivative of each entry individually)

   and verify you get the same answer.

   Solution.

$$Ae^{At} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ e^t - e^{-t} & e^{-t} \end{bmatrix} = \begin{bmatrix} e^t & 0 \\ e^t + e^{-t} & -e^{-t} \end{bmatrix}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{At} = \frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} e^t & 0 \\ e^t - e^{-t} & e^{-t} \end{bmatrix} = \begin{bmatrix} e^t & 0 \\ e^t + e^{-t} & -e^{-t} \end{bmatrix} = Ae^{At} \quad \checkmark$$

■

(g) Verify that if you plug in $t = 0$ to e^{At}, you get I.

   Solution. $\begin{bmatrix} e^0 & 0 \\ e^0 - e^{-0} & e^{-0} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 - 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \quad \checkmark$ ■

(h) Based on the above, verify that

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0$$

   is a solution to

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = A\mathbf{x}(t), \quad \mathbf{x}(0) = \mathbf{x}_0$$

   Hint: Just explain why it satisfies both $\mathbf{x}' = A\mathbf{x}$ and $\mathbf{x}(0) = \mathbf{x}_0$.

   Solution.

$$\mathbf{x}' = (e^{At})'\mathbf{x}_0 = (Ae^{At})\mathbf{x}_0 = A(e^{At}\mathbf{x}_0) = A\mathbf{x} \quad \checkmark$$

$$\mathbf{x}(0) = e^{A0}\mathbf{x}_0 = I\mathbf{x}_0 = \mathbf{x}_0 \quad \checkmark$$

■

9.9 Orthogonal Diagonalization

8. Recall a matrix A is "diagonalizable" if there exists a matrix P such that $P^{-1}AP = D \iff A = PDP^{-1}$, where D is diagonal. It follows that the columns of P must form an eigenbasis for A.

   (a) Suppose that S is a real matrix with an orthonormal eigenbasis. That is, the matrix Q with the eigenbasis as its columns is an orthogonal matrix ($Q^{-1} = Q^T$).

   $$S = QDQ^{-1} = QDQ^T$$

   Show that S is symmetric.

   Solution. $S^T = \left(QDQ^T\right)^T = \left(Q^T\right)^T D^T Q^T = QDQ^T = S$
   since D is diagonal (and thus, symmetric). Thus, $S = S^T$ so it's symmetric. ■

   Remark 48. By doing this, you will have shown that if a (real) matrix has an orthonormal eigenbasis, then it is symmetric. We say that an orthogonal matrix that diagonalizes S ($Q^T SQ = D$) "orthogonally diagonalizes S".

   It turns out that all real symmetric matrices have an orthonormal eigenbasis (though this is very hard to prove). That is, real symmetric matrices are *exactly* the *only* kind of (real) matrix guaranteed to have an orthonormal eigenbasis. This is why real symmetric matrices are objectively the best matrices.

9.9.1 Spectral Decomposition

   (b) Using problems 7b and 7e, explain why if the columns of Q are $\mathbf{v}_j$, then

      i.
      $$S = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T + \ldots + \lambda_n \mathbf{v}_n \mathbf{v}_n^T$$

ii.

$$S^k = \lambda_1^k \mathbf{v}_1 \mathbf{v}_1^T + \ldots + \lambda_n^k \mathbf{v}_n \mathbf{v}_n^T$$

iii.

$$e^{St} = e^{\lambda_1 t} \mathbf{v}_1 \mathbf{v}_1^T + \ldots + e^{\lambda_n t} \mathbf{v}_n \mathbf{v}_n^T$$

Hint: Use the idea of diagonalization to show ii. (since that implies i. directly). For iii., you can use the logic of 7e (please don't do it all over again). I'm not looking for a lengthy explanation.

Remark. This is called the "spectral decomposition" of A. Each $\mathbf{v}_j \mathbf{v}_j^T$ is a projector onto the eigenvector $\mathbf{v}_j$.

Solution. We've already done the groundwork in problem 7.

$$D = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \ldots + \lambda_n \mathbf{e}_n \mathbf{e}_n^T$$

so

$$D^k = \lambda_1^k \mathbf{e}_1 \mathbf{e}_1^T + \ldots + \lambda_n^k \mathbf{e}_n \mathbf{e}_n^T$$

and thus all the same formulas from problem 7 hold, except it'll be $\mathbf{v}_j$ and then $\mathbf{e}_j^T Q^T = \mathbf{v}_j^T$.

$$\begin{aligned} QD^kQ^T = \lambda_1^k Q\mathbf{e}_1 \mathbf{e}_1^T Q^T + \ldots + \lambda_n^k Q\mathbf{e}_n \mathbf{e}_n^T Q^T \\ = \lambda_1^k \mathbf{v}_1 \mathbf{v}_1^T + \ldots + \lambda_n^k \mathbf{v}_n \mathbf{v}_n^T \end{aligned}$$

Further, since $e^{St} = Qe^{Dt}Q^T$ (and also just based on what we did in problem 7), we get

$$e^{St} = e^{\lambda_1 t} \mathbf{v}_1 \mathbf{v}_1^T + \ldots + e^{\lambda_n t} \mathbf{v}_n \mathbf{v}_n^T$$

■

9.9.2 Singular Value Decomposition / PCA

(c) We can do a similar decomposition into a sum of things in terms of orthonormal bases with the Singular Value Decomposition (SVD)

$$A = USV^T$$

where

- A and S are $m \times n$
- U is an $m \times m$ orthogonal matrix
- V is an $n \times n$ orthogonal matrix.

S is a rectangular diagonal matrix with nonnegative "diagonal" entries σ_k (the singular values in descending order) on the "diagonal". The singular values squared σ_k^2 are the nonzero eigenvalues of $A^T A$ and AA^T, the columns of U are the eigenvectors of AA^T and the columns of V are the eigenvectors of $A^T A$.

The following are examples of S matrices

$$\begin{bmatrix} 7 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix}, \quad \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Then we can write

$$S = \sigma_1 \mathbf{e}_1^{(m)} {\mathbf{e}_1^{(n)}}^T + \ldots + \sigma_r \mathbf{e}_r^{(m)} {\mathbf{e}_r^{(n)}}^T$$

where the $\mathbf{e}_j^{(m)}$ means the jth standard basis vector of $\mathbb{R}^m$ (and $\mathbf{e}_j^{(n)}$ is in $\mathbb{R}^n$), and r is the rank (number of nonzero singular values). Then, if the columns of U are $\mathbf{u}_j$, and the columns of V are $\mathbf{v}_j$, then show we can write

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \ldots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T$$

Solution. Given the sum formula for S above we have

$$USV^T = \sigma_1 U \mathbf{e}_1^{(m)} {\mathbf{e}_1^{(n)}}^T V^T + \ldots + \sigma_r U \mathbf{e}_r^{(m)} {\mathbf{e}_r^{(n)}}^T V^T$$

And $U\mathbf{e}_j = \mathbf{u}_j$ and ${\mathbf{e}_r^{(n)}}^T V^T = \mathbf{v}_j^T$, so we just get the desired

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \ldots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T$$

■

Remark 49. PCA is essentially just taking some subset of the singular values (the largest ones). This gives you the "most important" contributing factors of the matrix.

If we convert an image into a matrix A (say it has 2,000 singular values), then taking the first few hundred will result in a compressed but largely recognizable version of the image. That is, if

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \ldots + \sigma_{2000} \mathbf{u}_{2000} \mathbf{v}_{2000}^T$$

then a pretty good approximation is something like

$$A \approx \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \ldots + \sigma_{300} \mathbf{u}_{300} \mathbf{v}_{300}^T$$

but the latter takes up much less storage size.

When I did a project detecting if pages were black or white, I found that white pages have one singular value much larger than the second highest (they are much easier to approximate well) while the largest singular values of a black page are closer together.

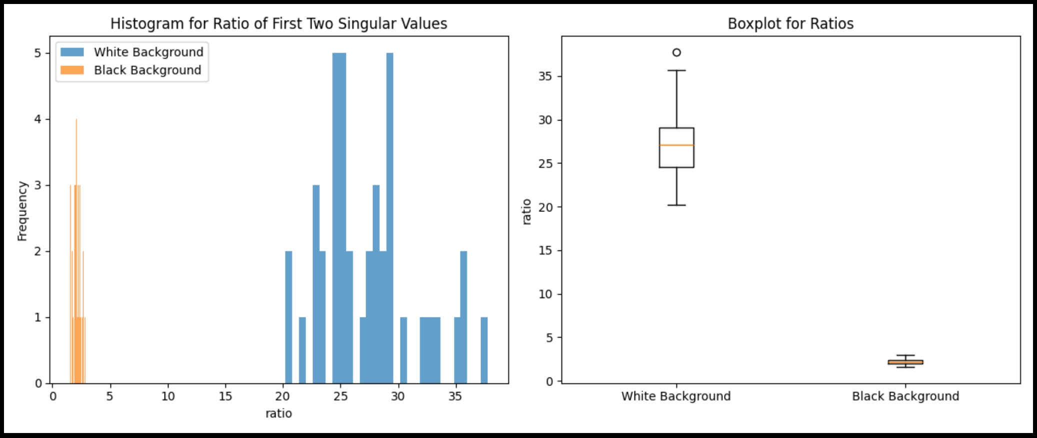


Figure 1: Ratio of Singular Values of PDFs with Black or White pages

Remark 50. Here are some facts about this decomposition that can help you calculate it:

i. $A\mathbf{v}_j = \sigma_j \mathbf{u}_j$
Thus, if you have all your $\mathbf{v}_j$ vectors corresponding to the nonzero singular values, you can get $\mathbf{u}$ from

$$\mathbf{u}_j = \frac{A\mathbf{v}_j}{\sigma_j}$$

ii. $A^T A = \sigma_1^2 \mathbf{v}_1 \mathbf{v}_1^T + \ldots + \sigma_r^2 \mathbf{v}_r \mathbf{v}_r^T$
Thus, you can get your $\mathbf{v}$ vectors by doing the (orthogonal) diagonalization of $A^T A$. Since it's symmetric, we know this decomposition into $A^T A = VDV^T$ exists. More specifically, $A^T A = VS^T SV^T$, and $S^T S$ (and SS^T) are both diagonal matrices.

iii. Similarly, as $AA^T = USS^T U^T$,

$$AA^T = \sigma_1^2 \mathbf{u}_1 \mathbf{u}_1^T + \ldots + \sigma_r^2 \mathbf{u}_r \mathbf{u}_r^T$$

Some questions that naturally come up:

- Are the nonzero eigenvalues of AA^T always the same as $A^T A$?
- Is the rank of A always the same as $A^T A$ and AA^T?
- Does the SVD always exist?

The answer to all of these questions is a (possibly surprising) yes! But they aren't easy questions to answer.

9.10 Non-Diagonalizable Matrices

9. Consider the matrix

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

(a) Find the eigenvalues and corresponding eigenvectors.

Solution. Trace is 4 and so is the determinant. The only two numbers that add up and multiply to 4 are 2 and 2, so the eigenvalues are $\lambda = 2, 2$ (2 is repeated!).

$$A - 2I = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$$

A basis for the kernel is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and that's the only linearly independent vector we can get, since there's only one free variable. ■

(b) Show A is not diagonalizable.

Solution. Since we can't form a basis of eigenvectors (there's only one linearly independent eigenvector), then it's not diagonalizable. ■

9.10.1 Jordan Decomposition

10. Optional problem:

Continuing off from the previous problem with $A = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}$.

(a) Let λ be an eigenvalue of A and $\mathbf{v}$ be a corresponding eigenvector. Show $(A - \lambda I)\mathbf{x} = \mathbf{v}$ is consistent. Let $\mathbf{w}$ be a solution.

Solution. We have $\lambda = 2$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then

$$\left[\ A - \lambda I \mid \mathbf{v}\ \right] = \left[\begin{array}{cc|c} 1 & -1 & 1 \\ 1 & -1 & 1 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -1 & 1 \\ 0 & 0 & 0 \end{array}\right]$$

The general solution is

$$\mathbf{w} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The best choice is $\mathbf{w} = \begin{bmatrix}1\\0\end{bmatrix}$. ■

(b) Show that $(A - \lambda I)^2 = 0$.

Solution.

$$\begin{bmatrix}1 & -1\\1 & -1\end{bmatrix}\begin{bmatrix}1 & -1\\1 & -1\end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix} \quad \checkmark$$

■

(c) Verify that the system $c_1\mathbf{v} + 1\begin{bmatrix}1\\0\end{bmatrix} = \mathbf{w}$ is consistent.

Solution. With our choice of $\mathbf{w} = \begin{bmatrix}1\\0\end{bmatrix}$, then this is clearly consistent with $c_1 = 0$. But no matter what we picked for $\mathbf{w}$, this would be consistent. Our general solution for $\mathbf{w}$ was

$$\mathbf{w} = t\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}1\\0\end{bmatrix}$$

which is exactly what this question is asking about (letting $t = c_1$). ■

(d) Verify that

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

Hint: For ease of calculation, you can use that this expression is equal to

$$\lambda I + \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}^{-1} = \lambda I + \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}0 & 1\\1 & -1\end{bmatrix}$$

Solution.

$$2I + \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = 2I + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}$$

■

Remark 51. The matrix $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$ is called a Jordan matrix, and would be A's Jordan Canonical Form. It's the best eigendecomposition you can do for a nondiagonalizable matrix (i.e. as close as you can get to a diagonal matrix). As we can see, $(A - \lambda I)^2 = 0$ shows that A has some "nilpotency" to it, which is actually why it isn't diagonalizable. MATH132 is also the class that teaches you primarily about this form of decomposition.

Theorem List

Vector spaces are real and finite dimensional unless otherwise specified.

Definition 52. The span of a set of vectors

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \{c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k : c_i \in \mathbb{R}\}$$

is the set of all linear combinations of that set of vectors.

The span can be understood as a subspace that the vectors generate. It is the smallest subspace that contains all the vectors.

Definition 53. A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is **linearly dependent** if one of the vectors (say $\mathbf{v}_i$) can be written as a linear combination of the other vectors

$$\mathbf{v}_i = c_1\mathbf{v}_1 + \ldots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \ldots + c_n\mathbf{v}_n$$

i.e. at least one vector $\mathbf{v}_i$ is in the span of the other vectors. Note that this means we can remove $\mathbf{v}_i$ from the set of vectors, and the span will be unchanged. In other words, the set isn't a "minimal generating set", since we can get the same span with fewer vectors.

Equivalently, there is a linear combination such that

$$c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$$

where at least one $c_i \neq 0$.

Definition 54. A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is **linearly independent** if none of the vectors can be written as a linear combination of the other vectors. Equivalently, the only linear combination such that

$$c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$$

is the trivial solution $c_1 = c_2 = \ldots = c_n = 0$.

This implies that S is a minimal generating set for its span. That is, we cannot remove any of the vectors without reducing the span.

Remark 55. Equivalently, one can define linear independence as the property that representation by linear combinations is unique. That is, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent and

$$\mathbf{x} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = k_1\mathbf{v}_1 + \ldots + k_n\mathbf{v}_n$$

then $c_1 = k_1, \ldots, c_n = k_n$.

i.e. we can't get the same vector with a different linear combination.

Note that this does imply that a linearly dependent set of vectors is exactly a set where linear combination representations are *not* unique. This is consistent with the idea of trying to find a nontrivial linear combination representation of $\vec{0}$.

10.1 Subspaces, Column Space, Null Space, Linear Transformations

Definition 56. A linear transformation is a function between vector spaces that preserves linear combinations

$$T\colon V \to W, \quad T\left(c_1\mathbf{v}_1 + c_2\mathbf{v}_2\right) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$$

Theorem 57. A function T between vector spaces is a linear transformation if and only if both $T(\vec{0}) = \vec{0}$ and

$$T(\mathbf{v} + k\mathbf{w}) = T(\mathbf{v}) + kT(\mathbf{w})$$

Definition 58. A **subspace** W is a nonempty subset of a vector space V that is closed under linear combinations. i.e. if $\mathbf{u}, \mathbf{v}$ are in W, then

$$c_1\mathbf{u} + c_2\mathbf{v} \in W$$

for all $c_1, c_2 \in \mathbb{R}$.

Theorem 59. A nonempty subset of a real vector space $W \subset V$ is a subspace if and only if both $\vec{0} \in W$ and

$$\mathbf{w}_1 + k\mathbf{w}_2 \in W$$

for all $k \in \mathbb{R}$ and $\mathbf{w}_1, \mathbf{w}_2 \in W$.

Theorem 60. The span of a set of vectors is a subspace. Additionally, any subspace can be represented as the span of a set of vectors.

Further, the span of a set of vectors $W = \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ (called the subspace generated by $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$) is the smallest subspace that contains the vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.

Definition 61. We often use the notation T_A to denote the linear transformation $T_A(\mathbf{x}) = A\mathbf{x}$. Some texts make a distinction and say Nul A and $\ker(T_A)$, but I personally make no distinction. $\ker(A)$ is perfectly fine, in my opinion.

Definition 62. The **column space** of a matrix A, Col A, is the span of the columns. This is just the image of T_A. Thus, you can think of it as every vector $\mathbf{b}$ such that $A\mathbf{x} = \mathbf{b}$ is consistent.

Theorem 63. A transformation T_A is surjective/onto if and only if the column space of $A \in \mathbb{R}^{m\times n}$ is $\mathbb{R}^m$.

Definition 64. The **null space** or **kernel** of a matrix A, Nul A, $\ker(A)$, or $\ker(T_A)$ is the set of all solutions to the homogeneous equation $A\mathbf{x} = \vec{0}$. This is just the set of all preimages of $\vec{0}$.

Theorem 65. A transformation is injective or one-to-one if and only if $\ker(A) = \left\{\vec{0}\right\}$. That is, if the only preimage of $\vec{0}$ is $\vec{0}$.

Definition 66. A **basis** of a subspace W is a set of vectors $\{v_1, \ldots, v_k\}$ in W such that

1. $\text{span}\{v_1, \ldots, v_k\} = W$
2. $\{v_1, \ldots, v_k\}$ is linearly independent

i.e. a linearly independent spanning/generating set

Remark 67. If you think of linear independence as the property of having unique representations by linear combinations (i.e. $c_1v_1 + \ldots + c_nv_n = k_1v_1 + \ldots + k_nv_n$ if and only if $c_i = k_i$ for all i), then a basis for W is a set of vectors such that every vector in W can be represented *uniquely* as a linear combination of the vectors.

Definition 68. The **dimension** of a subspace W is the number of elements in a basis for W.

Theorem 69. Suppose V is a vector space of dimension n and S is any set of vectors.

- If S has more than n vectors, it is not linearly independent.
- If S has less than n vectors, then it does not span V.
- If S has exactly n vectors, then it spans V if and only if it is a linearly independent set.
- As a corollary of the above, any linearly independent set with $\dim(W)$ vectors in W
  1. spans W
  2. is a basis for W

Definition 70. The **rank** of a matrix $\text{rank}(A)$ is the dimension of its column space/image. The **nullity** $\text{nullity}(A)$ is the dimension of its null space.

Theorem 71. The rank of a matrix is the number of pivots, and the nullity is the number of free variables.

Theorem 72 (Rank-Nullity). $A \in \mathbb{R}^{m\times n} \implies \text{rank}(A) + \text{nullity}(A) = n$

Note: this is equivalent to

$$[\#\text{ of pivot cols}] + [\#\text{ of free variables}] = [\#\text{ of total cols}]$$

10.2 Eigenvalues and Eigenvectors

We may use “ew” for “eigenvalue” and “ev” for eigenvector.

Definition 73. An eigenvector (ev) of a linear operator T is a vector $\mathbf{v} \neq \vec{0}$ such that

$$T(\mathbf{v}) = \lambda\mathbf{v}$$

where λ is the eigenvalue (ew) of $\mathbf{v}$.

Definition 74. We call $E_\lambda(T)$ the "eigenspace" of T of eigenvalue λ. It's the subspace containing all eigenvector of eigenvalue λ.

$$E_\lambda(T) = \ker(T - \lambda I)$$

Note: $E_0(T) = \ker(T)$ means that nontrivial kernel elements are eigenvectors with eigenvalue 0 (which makes sense since $T\mathbf{v} = \vec{0} = 0\mathbf{v}$).

Theorem 75. λ is an eigenvalue of T if and only if

$$\det(T - \lambda I) = 0$$

(otherwise the kernel is trivial, so there are no nonzero solutions to $T\mathbf{v} = \lambda\mathbf{v}$).

Remark 76. In general, in a linear algebra course we're kind of anal about what is or isn't an eigenvector. It has to be a part of the "field" the matrix is over. For example,

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

is a 90° rotation matrix. Its characteristic polynomial is $\lambda^2 + 1 = 0$ which has roots $\lambda = \pm i$.

The non-mathematicians say $\pm i$ are the eigenvalues, and the eigenvectors are $\begin{bmatrix} \pm i \\ 1 \end{bmatrix}$.

But mathematicians say "no! if A is a matrix over $\mathbb{R}$ then its eigenvalues can only be real! thus A has no eigenvalues (over $\mathbb{R}$)!"

So while this matrix does have two linearly independent eigenvectors, because they're in $\mathbb{C}^2$ (even though they form a basis), we can't say A is diagonalizable. If you go on to do more advanced linear algebra, just be careful with what you call an eigenvalue.

Definition 77. The "characteristic polynomial" of T is

$$\det(\lambda I - T) = 0$$

($\det(T - \lambda I)$ works too, but isn't always monic)

Definition 78. We say T is diagonalizable if and only if T has an eigenbasis. That is, if there is a basis of eigenvectors. In the case of matrices, we can say

$$A = PDP^{-1}$$

where P is an invertible matrix of eigenvectors (this is only possible if A is diagonalizable) and D is a diagonal matrix of corresponding eigenvalues.

Theorem 79. If $\mathbf{v}_1$ is an ev with ew λ_1 and $\mathbf{v}_2$ is an ev with ew λ_2 and $\lambda_1 \neq \lambda_2$, then $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent.

Corollary 80. If T is an operator on $\mathbb{R}^n$ with n distinct eigenvalues, then T is diagonalizable.

Theorem 81. $\lambda = 0$ is an ew of T if and only if T is not invertible (it has a nontrivial kernel).

Span / Linear (in)dependence

11.1 Span

We can imagine a vector as a point or an arrow. For example, $(1,1)$ is an arrow that stretches from the origin to the point $(1,1)$, which is at a $45°$ degree angle from the x-axis. We can also say that it lies on the line $y = x$. So what is the connection between the vector $(1,1)$ and the line $y = x$? The answer is *span*.

Definition 82. The **span** of a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ in $\mathbb{R}^n$ is the set of all linear combinations of those vectors.

$$\text{span}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} := \{c_1\mathbf{v}_1 + \ldots + c_m\mathbf{v}_m : c_i \in \mathbb{R}\} \tag{11}$$

To take an example, the span of $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$ is the set of all vectors of the form $c_1\begin{bmatrix}1\\0\\0\end{bmatrix} + c_2\begin{bmatrix}0\\0\\1\end{bmatrix}$. The span of $(1,1)$ is going to be all vectors of the form $c(1,1)$.

Take a set of vectors $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq \mathbb{R}^n$. If the span of B gives the entire space, that is $\text{span}\,B = \mathbb{R}^n$, then we say that B **spans** $\mathbb{R}^n$.

The span of a set of vectors generates what we call a *subspace*. For now, you can just think of that as meaning a point, line, or a plane in space (that includes the origin). This is a massive oversimplification but it's good enough for now.

For example, $\text{span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$ describes a plane in $\mathbb{R}^3$ (specifically, the xz-plane).

The span of the zero vector is just the origin. The span of a single nonzero vector is just a line through the origin. This isn't completely deterministic, though. You can take the span of a million vectors and it could still just be a line (if every vector is a scalar multiple of each other).

Ex. $\text{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}, 2\begin{bmatrix}1\\1\end{bmatrix}, \ldots, 10^6\begin{bmatrix}1\\1\end{bmatrix}\right\} = \text{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$ (the line $y = x$).

11.2 Linear (in)dependence

So the span of some vectors is the set of every linear combination. Let's reexamine the example

$$\text{span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\} = \left\{c_1\begin{bmatrix}1\\0\\0\end{bmatrix} + c_2\begin{bmatrix}0\\0\\1\end{bmatrix} : c_1, c_2 \in \mathbb{R}\right\} = \left\{\begin{bmatrix}a\\0\\b\end{bmatrix} : a, b \in \mathbb{R}\right\}$$

in $\mathbb{R}^3$. Notice that the span of these vectors is all vectors with a zero in the second component.

What if we tack on another vector, $\begin{bmatrix}1\\0\\1\end{bmatrix}$?

$$c_1\begin{bmatrix}1\\0\\0\end{bmatrix} + c_2\begin{bmatrix}0\\0\\1\end{bmatrix} + c_3\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}c_1 + c_3\\0\\c_2 + c_3\end{bmatrix}$$

But that doesn't really get us anything new. That is, the span is unchanged after adding the vector. You can think of it as $\begin{bmatrix}1\\0\\1\end{bmatrix}$ doesn't give us any new information that $\begin{bmatrix}1\\0\\0\end{bmatrix}$ and $\begin{bmatrix}0\\0\\1\end{bmatrix}$ already did. The way we say that mathematically is that $\begin{bmatrix}1\\0\\1\end{bmatrix}$ is linearly dependent with $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$. This happened specifically because

$$\begin{bmatrix}1\\0\\1\end{bmatrix} = 1\begin{bmatrix}1\\0\\0\end{bmatrix} + 1\begin{bmatrix}0\\0\\1\end{bmatrix}$$

That is, $\begin{bmatrix}1\\0\\1\end{bmatrix}$ is a linear combination of $\begin{bmatrix}1\\0\\0\end{bmatrix}$ and $\begin{bmatrix}0\\0\\1\end{bmatrix}$. Or, it is in the span of those vectors.

Definition 83. A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is **linearly dependent** if one of the vectors (say $\mathbf{v}_i$) can be written as a linear combination of the other vectors

$$\mathbf{v}_i = c_1\mathbf{v}_1 + \ldots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \ldots + c_n\mathbf{v}_n$$

i.e. at least one vector $\mathbf{v}_i$ is in the span of the other vectors. Note that this means we can remove $\mathbf{v}_i$ from the set of vectors, and the span will be unchanged. In other words, the set isn't a "minimal generating set", since we can get the same span with fewer vectors.

Equivalently, there is a linear combination such that

$$c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$$

where at least one $c_i \neq 0$.

Here's the thing about linear algebra: just because two vectors are distinct, doesn't mean we need both of them. In a way, the vectors $(1,1)$ and $(2,2)$ are nearly the same because one is a scalar multiple of the other. Students are often used to using every distinct object available, but the way we sift through vectors, and decide what is important, is by considering linear independence. Linear independence is truly one of the most important concepts in linear algebra! Unsurprisingly, a set of vectors are linearly independent when they are not linearly dependent. Here is the formal definition:

Definition 84. A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is **linearly independent** if none of the vectors can be written as a linear combination of the other vectors. Equivalently, the only linear combination such that

$$c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$$

is the trivial solution $c_1 = c_2 = \ldots = c_n = 0$.

This implies that S is a minimal generating set for its span. That is, we cannot remove any of the vectors without reducing the span.

Remark. Equivalently, one can define linear independence as the property that representation by linear combinations is unique. That is, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent and

$$\mathbf{x} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = k_1\mathbf{v}_1 + \ldots + k_n\mathbf{v}_n$$

then $c_1 = k_1, \ldots, c_n = k_n$.

i.e. we can't get the same vector with a different linear combination.

Note that this does imply that a linearly dependent set of vectors is exactly a set where linear combination representations are *not* unique.

11.2.1 Showing a set of vectors is linearly (in)dependent

To show a set of vectors is linearly dependent, you can go about it two main ways:

1. Write one of the vectors as a linear combination of the others.

2. Find a nontrivial linear combination that yields zero.

To show vectors are linearly independent, you almost always start with the equation

$$c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$$

and attempt to show that all the constants must be zero $c_1 = c_2 = \ldots = c_n = 0$. In the case where you have very few vectors (two to three, generally, max), you can also show that no vector is a linear combination of the others.

You can also start with $c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$ to show that something is linearly dependent, and attempt to simplify until you can find a nonzero solution for $c_1, \ldots, c_n$.

Injective / Surjective Handout

What the heck do these terms mean?

12.1 Images and Preimages

Recall that if

$$T(\mathbf{x}) = \mathbf{y}$$

then we say $\mathbf{y}$ is the image of $\mathbf{x}$, and $\mathbf{x}$ is a preimage of $\mathbf{y}$.

An image is unique (as the output for a function is unique), but a preimage is not necessarily. A preimage also need not exist for every element in the codomain.

12.2 Linear Transformations

A linear transformation is a function that *preserves* linear combinations

$$T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \ldots + c_nT(\mathbf{v}_n)$$

12.3 Kernel and Range

The kernel is all preimages of zero. That is, stuff that gets sent to zero.

$$T(\mathbf{x}) = \vec{0} \iff \mathbf{x} \in \ker(T)$$

The range of a transformation is the set of all possible outputs (that T can reach). i.e. all the vectors in the codomain that have a preimage.

12.4 Injective

Injective (often called being one-to-one) is the property that *preimages are unique.* That is, there's only one input for every output. In symbols

$$T(\mathbf{x}) = T(\mathbf{y}) \implies \mathbf{x} = \mathbf{y}$$

But using our properties of linear transformations, this is equivalent to

$$T(\mathbf{x} - \mathbf{y}) = \vec{0} \implies \mathbf{x} - \mathbf{y} = 0$$

or $T(\mathbf{z}) = \vec{0} \implies \mathbf{z} = 0$. That is, the only preimage of $\vec{0}$ is $\vec{0}$. This is the same as saying the kernel is "trivial" or

$$\ker(T) = \left\{\vec{0}\right\}$$

How do we determine if something like $A\mathbf{x} = \mathbf{b}$ has a unique solution, though? Well, the basic idea is that if we solve

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

then we want to have only the unique solution of $\vec{0}$. Thus, we need there to be **no free variables**. That is, if the preimage of $\vec{0}$ is unique, then all preimages are unique!

This gives us a criterion to determine if $A\mathbf{x}$ is injective:

Theorem 85. $A\mathbf{x}$ is injective (one-to-one) if and only if

- every column of A is a pivot column
- there are no free variables
- the columns of A are linearly independent
- every column has a pivot
- the only solution to $A\mathbf{x} = 0$ is $x = 0$
- every *consistent* system $A\mathbf{x} = \mathbf{b}$ has a unique solution (a solution does not need to exist, though)
- There is no nontrivial solution to $A\mathbf{x} = \vec{0}$
- $\ker(A) = \{0\}$

All of these are equivalent ways to say the same thing.

Notice that injectivity is about the columns not rows. You can get a row of zeros and still be injective!

12.4.1 Showing it's not injective

Reversing some of the statements we can get some easy checks to show something is not injective.

- If there are more columns than rows, we must have a free variable so it can't be injective
- If you can find a nonzero preimage of $\vec{0}$ (a nonzero solution to $A\mathbf{x} = \vec{0}$), it's not injective

One way to interpret having "too many columns" is that the domain of the transformation is much larger than the codomain. Thus, if you try to squeeze, say, a seven dimensional space into a three dimensional space, then there will inevitably be things that "fold onto themselves", which means multiple inputs map to the same output.

12.5 Surjective

To be surjective (often called onto) means that every vector in the codomain has a preimage. That is, the range is the whole codomain. We can also think of it as

$$T(\mathbf{x}) = \mathbf{b}$$

always has a solution for all $\mathbf{b}$.

Let's think about this in terms of matrices. If we want

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

to be consistent, then you might make the observation that there can't be a row of zeros. If there is a row of zeros, then we might get $0 = 1$. If there isn't a row of zeros, then we'll always get a solution. We have the criterion

Theorem 86. $A\mathbf{x}$ is surjective (onto) if and only if

- every row of A has a pivot
- the rows of A are linearly independent
- there are no rows of zeros in the EF or REF
- $A\mathbf{x} = \mathbf{b}$ is consistent for all $\mathbf{b}$

Similar to how injectivity is about the columns, surjectivity is about the rows. You can have free variables and still be surjective (though students don't typically make this mistake as much as thinking a row of zeros means not injective).

To show something is surjective, you can do any of these things, but the most basic (but not necessarily easy way) is to find a generic preimage of anything in the codomain. For example, if you can find a preimage of both $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$, then it must be surjective because every vector is a linear combination of those, so you can get any preimage as a linear combination of those preimages.

But a more surefire method is to start row reducing until you can show every row will have a pivot.

Similar to how having too many columns means its not injective, too many rows means it can't be surjective. We can have a similar geometric image that if we try to spread out a smaller dimensional space into a larger one, it won't work. For example, stretching $\mathbb{R}^2$ into $\mathbb{R}^3$ is impossible. You can, at best, get a plane.

12.6 Square Matrices

Notice that injectivity is equivalent to having a pivot in every column, and surjectivity is equivalent to having a pivot in every row.

Thus, if A is a square matrix, having a pivot in every column is the same as having a pivot in every row. Therefore,

Theorem 87. If A is a square matrix, then $A\mathbf{x}$ is surjective if and only if $A\mathbf{x}$ is injective.

This is really nice because checking injectivity is generally easier than checking surjectivity.

Remark 88. Personally, I dislike justifying the equivalence of injectivity and surjectivity this way, but it is a valid method.

That does mean, though, that a row of zeros is enough to say it's not injective or surjective.

Non-Worksheet Problems

These were going to be in worksheets or were on worksheets in a previous quarter (but due to the different way the course was taught, they weren't really appropriate).

1. Consider the matrix $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 5 & 3 \\ 0 & 0 & 2 & 4 \\ -2 & 1 & -9 & -4 \end{bmatrix}$

   (a) Perform the following row operations:

   - $R_1 \leftrightarrow R_2$
   - $R_3 \to -R_1 + 2R_2 + R_4$
   - $R_4 \to \frac{1}{2}R_3$

   (b) Compute the following matrix multiplication

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

   and compare the result to part a.

   (c) Explain the connection between the matrix

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

   and the row operations from part a.

   - $R_1 \to R_2$

- $R_2 \to R_1$
- $R_3 \to -R_1 + 2R_2 + R_4$
- $R_4 \to \frac{1}{2}R_3$

Hint: Compare the operations to the rows of B

$$\begin{aligned}
\begin{bmatrix}0 & 1 & 0 & 0\end{bmatrix} &\iff & R_1 \to R_2 \\
\begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix} &\iff & R_2 \to R_1 \\
\begin{bmatrix}-1 & 2 & 0 & 1\end{bmatrix} &\iff & R_3 \to -R_1 + 2R_2 + R_4 \\
\begin{bmatrix}0 & 0 & \frac{1}{2} & 0\end{bmatrix} &\iff & R_4 \to \frac{1}{2}R_3
\end{aligned}$$

(d) Write a matrix C that performs the following operations

- Multiply row 1 by four
  $R_1 \to 4R_1$
- Swap rows 2 and 4
  $R_2 \leftrightarrow R_4$
- Leave row 3 the same.
  $R_3 \to R_3$

2. Determine if the given set is a subspace of P_n for an appropriate value of n. Justify your answers.

(a) All polynomials of the form $p(t) = a + t^2$, where a is in $\mathbb{R}$.

(b) All polynomials of degree at most 3, with integers as coefficients.

(c) Let H be the set of all vectors of the form

$$\begin{bmatrix} s \\ 3s \\ 2s \end{bmatrix}$$

Find a vector $\mathbf{v}$ in $\mathbb{R}^3$ such that $H = \text{Span}\{\mathbf{v}\}$. Why does this show that H is a subspace of $\mathbb{R}^3$?

3. The set of all continuous real-valued functions defined on a closed interval $[a, b]$ in $\mathbb{R}$ is denoted by $C[a, b]$. This set is a subspace of the vector space of all real-valued functions defined on $[a, b]$.

   Show that $\{f \in C[a, b] : f(a) = f(b)\}$ is a subspace of $C[a, b]$.

4. For fixed positive integers m and n, the set $M_{m\times n}$ of all $m \times n$ matrices is a vector space, under the usual operations of addition of matrices and multiplication by real scalars.

   (a) Determine if the set H of all matrices of the form

   $$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$$

   is a subspace of $M_{2\times 2}$.

   (b) Let F be a fixed 3×2 matrix, and let H be the set of all matrices A in $M_{2\times 4}$ with the property that $FA = 0$ (the zero matrix in $M_{3\times 4}$). Determine if H is a subspace of $M_{2\times 4}$.

5. Define $T : P_2 \to \mathbb{R}^2$ by $T(p) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$. For instance, if $p(t) = 3 + 5t + 7t^2$, then $T(p) = \begin{bmatrix} 3 \\ 15 \end{bmatrix}$.

   Show that T is a linear transformation. [Hint: For arbitrary polynomials p, q in P_2, compute $T(p + q)$ and $T(cp)$.]

6. Let $M_{2\times 2}$ be the vector space of all 2×2 matrices, and define $T : M_{2\times 2} \to M_{2\times 2}$ by $T(A) = A + A^T$, where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

   Show that T is a linear transformation.

7. Define $T : P_2 \to \mathbb{R}^2$ by $T(p) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$. For instance, if $p(t) = 3 + 5t + 7t^2$, then $T(p) = \begin{bmatrix} 3 \\ 15 \end{bmatrix}$.

Find a polynomial p in P_2 that spans the kernel of T, and describe the range of T.

8. Define a linear transformation $T : P_2 \to \mathbb{R}^2$ by $T(p) = \begin{bmatrix} p(0) \\ p(0) \end{bmatrix}$. Find polynomials P_1 and P_2 in P_2 that span the kernel of T, and describe the range of T.

9. Let $M_{2\times2}$ be the vector space of all 2×2 matrices, and define $T : M_{2\times2} \to M_{2\times2}$ by $T(A) = A + A^T$, where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

   Find the kernel and range of T.

10. The set $\mathcal{B} = \{1 + t^2, t + t^2, 1 + 2t + t^2\}$ is a basis for P_2. Express $p(t) = 1 + 4t + 7t^2$ as a linear combination of the vectors of $\mathcal{B}$.

13.1 RIP Determinants

11. Find the determinants by row reduction to echelon form.

$$(a)\quad \begin{vmatrix} 1 & 3 & 2 & -4 \\ 0 & 1 & 2 & -5 \\ 2 & 7 & 6 & -3 \\ -3 & -10 & -7 & 2 \end{vmatrix} \qquad\qquad (b)\quad \begin{vmatrix} 1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 4 \\ -1 & 0 & 5 & 3 \\ 3 & -3 & -2 & 3 \end{vmatrix}$$

12. Compute the determinants by cofactor expansions. At each step, choose a row or column that involves the least amount of computation.

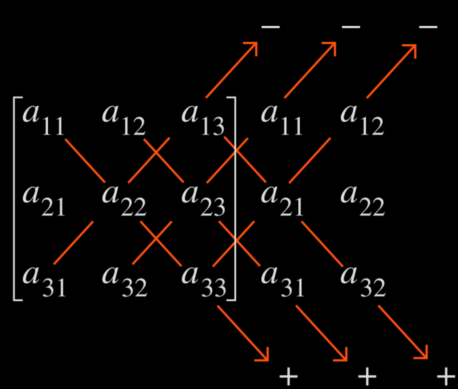
$$(a)\quad \begin{vmatrix} 4 & 0 & -7 & 3 & -5 \\ 0 & 2 & 0 & 0 & 0 \\ 7 & 3 & -6 & 4 & -8 \\ 5 & 0 & 5 & 2 & -3 \\ 0 & 0 & 9 & -1 & 2 \end{vmatrix} \qquad\qquad (b)\quad \begin{vmatrix} 6 & 3 & 2 & 4 & 0 \\ 9 & 0 & -4 & 1 & 0 \\ 8 & -5 & 6 & 7 & 1 \\ 2 & 0 & 0 & 0 & 0 \\ 4 & 2 & 3 & 2 & 0 \end{vmatrix}$$

13. If two rows/columns of a matrix are identical, then the determinant is zero. Show that if $x_1 \neq x_2$, then

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

defines a line that passes through (x_1, y_1) and (x_2, y_2)

14. The expansion of a 3×3 determinant can be remembered by the following device. Write a second copy of the first two columns to the right of the matrix, and compute the determinant by multiplying entries on six diagonals:



Add the downward diagonal products and subtract the upward products. Use this method to compute the determinants in Exercises 15–18. **Warning: This trick does not generalize in any reasonable way to 4×4 or larger matrices.**

$$(a) \quad \begin{vmatrix} 1 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -2 \end{vmatrix} \qquad\qquad (b) \quad \begin{vmatrix} 1 & 3 & 4 \\ 2 & 3 & 1 \\ 3 & 3 & 2 \end{vmatrix}$$