TOPICS

Notation: This class introduces the following mathematical notation that might be new:

1. $\mathbb{R}$: The set of all real numbers. $\mathbb{R}$ includes numbers like 0, 1, π, $-e^{-123}$, and every other real number you can think of.

2. $\in$: "In." For example "$\pi \in \mathbb{R}$" means "π is in $\mathbb{R}$."

3. Set notation: We use curly brackets $\{\}$ to denote sets of things like numbers, vectors, etc. For example, $\{0, 1\}$ is the set that contains just the number 0 and the number 1. We can also use a colon (or vertical line) to add more description to the set: for example $\{x \in \mathbb{R} : x > \pi\}$ or $\{x \in \mathbb{R} | x > \pi\}$ both mean the set of all real numbers that are greater than π.

4. $\subseteq$:"Subset." Used when one set is contained in another set. Specifically, "$A \subseteq B$" means "everything that's in A is also in B." For example $\{0, 1, \pi, -e^{-123}\} \subseteq \mathbb{R}$.

5. $\mathbb{R}^2$: The plane. This is the set of all vector that look like $\begin{bmatrix} a \\ b \end{bmatrix}$ where a and b are both in $\mathbb{R}$. For example, $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathbb{R}^2$.

6. $\mathbb{R}^n$: n-dimensional vector space, where n is any positive whole number. This comprises vectors of the form $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ where $a_1, \ldots, a_n \in \mathbb{R}$. For example, $\begin{bmatrix} 0 \\ 1 \\ \pi \\ -e^{-123} \end{bmatrix} \in \mathbb{R}^4$.

Definitions: *1: system of linear equations, 2: homogeneous, 3: solution set, 4: dot product & magnitude, 5: matrix, 6: row-operations, 7: row-echelon & reduced row-echelon form, 8: basic & free variables, 9: rank, 10: transpose, 11: inverse, 12: determinant, 13: elementary matrix, 14: linear combination, 15 span, 16 linear independence & linear* *independence, 17: vector space, 18 basis, 19: null space & nullity, 20: column & row space, 21: image, 22: orthogonal & orthonormal basis, 23: Fourier expansion, 24: orthogonal matrix*

Definition 1. Linear equation: A **linear equation** is an equation involving some variables (like x_1, x_2, and x_3 for example) where the variables can be multiplied by numbers in $\mathbb{R}$ and can be added to each other, but no other operations (like multiplying the variables by *each other*, raising a variable to a power, exponentiation, logarithms, trig functions, etc.) are allowed.

For example, $5x - 7y + 11z = 13$, $x_1 - x_2 + 1000x_3 - 999x_4 + x_5 = 0$, and $a - b + c = 10$ are linear equations, but $xy - z = 9$, $x^2 - 2^y = 5$, $\ln(x^y) - \cos(x + 5z) = 0$ are not linear equations.

A **system of linear equations** is a set of multiple linear equations.

Definition 2. Homogeneous: A system of linear equations

$$\begin{aligned} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= b_1 \\ &\vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &= b_m \end{aligned}$$

is **homogeneous** if $b_i = 0$ for all $1 \leq i \leq m$.

For example,

$$\begin{aligned} -x + 13y &= 0 \\ 5x - 3y &= 0 \end{aligned}$$

is a homogenous system of linear equations, while

$$-x + 13y = 10$$
$$5x - 3y = -2$$

is a non-homogeneous system of linear equations.

Definition 3. Solution set: Given a system of linear equations

$$a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1$$
$$\vdots$$
$$a_{m,1}x_1 + \cdots + a_{m,n}x_n = b_m,$$

the **solution set** to the system is the set of all solutions to the system (written as vectors $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$). It is possible for a solution set to be empty (i.e. the system has *no* solutions), have 1 vector, or have ∞ vectors.

Definition 4. Dot Product and Magnitude Given two vectors $\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ and $\vec{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ in $\mathbb{R}^n$, the **dot product**, denoted $\vec{v} \cdot \vec{u}$, is

$$v_1u_1 + \cdots v_nu_n.$$

The **magnitude** (aka **length**) of $\vec{v}$, denoted $\|\vec{v}\|$, is $\sqrt{\vec{v}\cdot\vec{v}}$. In other words,

$$\|\vec{v}\| = \sqrt{v_1^2 + \cdots v_n^2}.$$

The dot product can be used to calculate the cosine of the angle θ between two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$, with the following formula:

$$\cos(\theta) = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|}.$$

In particular, this means that $\vec{u}$ and $\vec{v}$ are perpendicular if and only if $\vec{u}\cdot\vec{v} = 0$.

Definition 5. Matrix: An $m \times n$ matrix is an arrangement of real numbers into a rectangle consisting of m rows and n columns. For example, $\begin{bmatrix} 2 & 3 & 5 \\ 7 & 11 & 13 \end{bmatrix}$ is a 2×3 matrix. The (i, j)-entry of a matrix is the number in the matrix's i^{th} row and j^{th} column. If the matrix is given a name, like M, then this entry can also be denoted as $M_{i,j}$.

A matrix is **square** if $m = n$, which means there are as many rows as there are columns.

Special kinds of matrices are the **zero matrix**, which is a matrix whose every entry is 0, and the $n \times n$ **identity matrix**, denoted I_n, whose (i, j)-entry is 1 if $i = j$ and 0 if $i \neq j$. For example,

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is the 3×3 identity matrix.

Matrices can be **added**, **scalar multiplied**, and **multiplied**.

1. **Addition**: Two matrices A and B can be added if they have they exact same rows

and columns as each other. In this case, $(A+B)_{i,j} = A_{i,j} + B_{i,j}$. For example,

$$\begin{bmatrix} 1 & 0 & 3 & 9 \\ -10 & 15 & 14 & 2 \end{bmatrix} + \begin{bmatrix} 6 & 7 & 1 & 9 \\ 3 & -6 & 4 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 7 & 4 & 18 \\ -7 & 9 & 18 & 6 \end{bmatrix}.$$

2. **Scalar Multiplication**: Any matrix A can be multiplied by any scalar $s \in \mathbb{R}$. In this case, $(sA)_{i,j} = s * A_{i,j}$. For example,

$$3 * \begin{bmatrix} 1 & 0 & 3 & 9 \\ -10 & 15 & 14 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 9 & 27 \\ -30 & 45 & 42 & 6 \end{bmatrix}.$$

3. **Multiplication**: If A is an $m \times n$ matrix and B is an $r \times s$ matrix then the product AB can only exist if $n = r$ (in other words, if the number of *columns* of A is the same as the number of *rows* of B). If AB exists, then it will be an $m \times s$ matrix and $(AB)_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$.

   For example, take

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

   and

$$B = \begin{bmatrix} 0 \\ -5 \\ 3 \\ 15 \end{bmatrix}.$$

   Then A is 1×4 and B is 4×1, so AB exists and is a 1×1 matrix. In this case, multiplying a row matrix to a column matrix is like taking the dot product: $AB = \begin{bmatrix} 1 * 0 + 2 * (-5) + 3 * 3 + 4 * 15 \end{bmatrix} = \begin{bmatrix} 59 \end{bmatrix}$.

For a more complex example, we can take

$$A = \begin{bmatrix} 1 & 0 & 3 & 9 \\ -10 & 15 & 14 & 2 \end{bmatrix}$$

and

$$B = \begin{bmatrix} 4 & 8 & 0 \\ 1 & 6 & -2 \\ 0 & 0 & 0 \\ 1 & 5 & 9 \end{bmatrix}.$$

In this case, A is 2×4 and B is 4×3, so AB exists and will be a 2×3 matrix. Notice in this particular case BA does ***NOT*** exist.

We can calculate the (i, j)-entry of AB by multiplying the i^{th} row of A to the j^{th} column of B. So for example, the $(1, 1)$-entry of AB is

$$\begin{bmatrix} 1 & 0 & 3 & 9 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 0 \\ 1 \end{bmatrix} = 1 * 4 + 0 * 1 + 3 * 0 + 9 * 1 = 13.$$

Or the $(2, 3)$-entry of AB is

$$\begin{bmatrix} -10 & 15 & 14 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \\ 0 \\ 9 \end{bmatrix} = -10 * 0 + 15 * (-2) + 14 * 0 + 2 * 9 = -12.$$

In total,

$$\begin{bmatrix} 1 & 0 & 3 & 9 \\ -10 & 15 & 14 & 2 \end{bmatrix} \begin{bmatrix} 4 & 8 & 0 \\ 1 & 6 & -2 \\ 0 & 0 & 0 \\ 1 & 5 & 9 \end{bmatrix} = \begin{bmatrix} 13 & 53 & 81 \\ -23 & 20 & -12 \end{bmatrix}.$$

CAUTION: It not guaranteed that AB will be equal to BA, even if both products exist.

Definition 6. Row Operations: Given a matrix M, there are three kinds of row operations we can perform on M:

1. Swap the positions of two rows in a matrix. This is denoted $r_i \leftrightarrow r_j$. For example, if

$$M = \begin{bmatrix} 0 & 1 & 5 & 6 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

then the row operation $r_1 \leftrightarrow r_2$ swaps the positions of rows 1 and 2, giving us

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

Multiply a row by a nonzero number. This is denoted $r_i \to sr_i$. For example, if

$$M = \begin{bmatrix} 0 & 1 & 5 & 6 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

then the row operation $r_2 \to 3r_2$ multiplies the second row by 3, giving us

$$\begin{bmatrix} 0 & 1 & 5 & 6 \\ 3 & 6 & 9 & 12 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Add a multiple of one row to another row. This is denoted $r_i \to r_i + tr_j$. For example, if

$$M = \begin{bmatrix} 0 & 1 & 5 & 6 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

then the row operation $r_1 \to r_1 + 2r_2$ adds two times the second row to the first row, giving us

$$\begin{bmatrix} 2 & 5 & 11 & 14 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Definition 7. Row-echelon and Reduced row-echelon form: A matrix is in row-echelon form if the following three criteria are followed:

1. Any row that contains a nonzero entry is above any row that contains only zeroes.

2. All the **leading entries** are 1 (a leading entry is the leftmost nonzero entry of a row).

3. Every leading each of every row is in a column to the right of the leading entries of the rows above it.

Furthermore, a matrix is in **reduced** row-echelon form if

1. Every entry above each leading entry is 0.

For example, the following are in **row-echelon form** but <u>not</u> reduced row-echelon form:

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

The following are in **reduced** row-echelon form:

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

And the following are **not** in row-echelon form at all:

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

An algorithm for turning any matrix into reduced row-echelon form is **Gauss-Jordan elimination**, which is the following steps:

1. Concentrate on the leftmost nonzero column of the matrix. Make sure there is a nonzero number in the top row of this column (use a row operation to make sure there is if there isn't). This entry of the matrix is the current **pivot position**.

2. Use row operations to make every number underneath the pivot position 0.

3. Move on to the next row and repeat steps 1 and 2 until there are no columns left to modify.

4. At this point every number underneath each leading entry should be 0. Now use the

scalar multiplication row operation to make every leading coefficient 1.

5. At this point your matrix is in row-echelon form. To make it reduced row-echelon, use the $r_i \to r_i + tr_j$ row operation to make every entry above each leading term 0. Now your matrix will be in reduced row- echelon form.

A common way of solving systems of linear equations is to transform the system of equations into a type of matrix called an augmented matrix by taking all of the coefficients. Then use row operations to transform this augmented matrix into reduced row-echelon form.

Example 1. For example, the system of equations

$$x + 3y - 2z = -18$$
$$2x + 9y + 11z = 96$$
$$3x + 18y + 41 = 360$$

can be transformed into the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 3 & -2 & -18 \\ 2 & 9 & 11 & 96 \\ 3 & 18 & 41 & 360 \end{array}\right]$$

We can follow along with Gauss-Jordan elimination to turn this matrix into reduced row-echelon form.

1. We will start at the leftmost column, where the top entry is 1.

2. We will use the 1 to make the 2 and 3 underneath it into 0 by using the row operations

$r_2 \to r_2 - 2r_1$ and $r_3 \to r_3 - 3r_1$. Then our augmented matrix becomes

$$\left[\begin{array}{ccc|c} 1 & 3 & -2 & -18 \\ 0 & 3 & 15 & 132 \\ 0 & 9 & 47 & 414 \end{array}\right].$$

3. Now we will repeat Step 1 by concentrating on the next row, where the leading entry is 3.

4. Repeat Step 2 by using the 3 to make the 9 underneath it 0 via the row operation $r_3 \to r_3 - 3r_2$. We then have the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 3 & -2 & -18 \\ 0 & 3 & 15 & 132 \\ 0 & 0 & 2 & 18 \end{array}\right].$$

5. Now every number beneath every leading entry is 0. We will transform each leading term into 1 by dividing by the row operations $r_2 \to r_2/3$ and $r_3 \to r_3/2$. Then we get the row-echelon matrix

$$\left[\begin{array}{ccc|c} 1 & 3 & -2 & -18 \\ 0 & 1 & 5 & 44 \\ 0 & 0 & 1 & 9 \end{array}\right].$$

6. Now we can turn the matrix into **reduced** row-echelon form by the row operations $r_2 \to r_2 - 5r_3$, $r_1 \to r_1 + 2r_3$, and $r_1 \to r_1 - 3r_2$ in that order. Then we get the reduced row-echelon matrix

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

Then we have our solution. Recall the first column represents x, the second column represents

y, and the third column represents z, so our solution is

$$x = 3$$
$$y = -1$$
$$z = 9.$$

Definition 8. Basic and Free Variables: If you turn a system of equations into an augmented matrix and use row-operations to turn the augmented matrix into row-echelon form, then any variable whose associated column has a leading term in it is a **basic variable** of the system of equations. Any other variable is a **free variable**.

The number of free variables is equal to the dimension of the solution set of the system of equations.

Example 2. Suppose you have the variables x_1, x_2, x_3, x_4, x_5, and x_6, and you have a system of equations that you have successfully transformed into the following reduced row-echelon form

$$\left[\begin{array}{cccccc|c} 1 & 0 & 2 & 0 & -3 & 4 & 14 \\ 0 & 1 & 12 & 0 & 6 & 7 & 8 \\ 0 & 0 & 0 & 1 & 4 & -9 & 7 \end{array}\right].$$

To interpret this matrix for your solution set, recall that each column corresponds to one of your variables. In this case, the **first**, **second**, and **fourth** columns have leading terms in them, so x_1, x_2, and x_4 are the **basic variables**.

That leaves x_3, x_5, and x_6 as the **free variables**. Right away, the three free variables tell you that the solution set is three-dimensional.

Set your free variables equal to its own **parameter** (a letter that represents a number from $\mathbb{R}$), like r, s, and t. So

$$x_3 = r$$
$$x_5 = s$$
$$x_6 = t.$$

Then use the rows of the augmented matrix to give you equations for x_1, x_2, and x_4 in terms of the parameters r, s, and t. For example, the first row of the matrix tells you

$$x_1 + 2x_3 - 3x_5 + 4x_6 = 14,$$

so

$$x_1 = 14 - 2r + 3s - 4t.$$

Continuing this way for x_2 and x_4, we get the following complete parametrization of the solution set:

$$\begin{aligned} x_1 &= 14 - 2r + 3s - 4t \\ x_2 &= 8 - 12r - 6s - 7t \\ x_3 &= r \\ x_4 &= 7 - 4s + 9t \\ x_5 &= s \\ x_6 &= t. \end{aligned}$$

If a system does not have any free variables, then the solution set is 0-dimensional... meaning that there is just one solution.

Definition 9. Rank: The rank of a matrix is the number of leading terms present after using row operations to get the matrix into row-echelon form. For example, if

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

then performing the row operation $r_3 \to r_3 - 2r_1$ gives us the matrix

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

which is in row-echelon form and has two leading entries (the 1 in the first row and the 1 in the second row). Therefore, the rank of M is 2, or $\text{rank}(M) = 2$.

Definition 10. Transpose: Given an $m \times n$ matrix A, the transpose of A– denoted A^T– is an $n \times m$ matrix such that $(A^T)_{i,j} = A_{j,i}$, so the (i, j)-entry of A^T is the (j, i)-entry of A. For example, if

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

then

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

So every row of A becomes a column of A^T and vice versa.

Transposes satisfy the property that $(AB)^T = B^T A^T$.

Definition 11. Inverse: Given a square $n \times n$ matrix A, the inverse of A is a matrix B

such that

$$AB = BA = I_n.$$

When an inverse for A exists, we may write it as A^{-1}. If A^{-1} exists, then A is called invertible.

To find the inverse of A, you can follow this algorithm:

1. Set up an augmented matrix as follows:

$$[A \mid I_n].$$

2. Use the Gauss-Jordan elimination algorithm to use row operations to turn A into reduced row-echelon form. Mimic every row operation you're doing on the left side onto the right side of the augmented matrix as well.

3. If A is invertible, its reduced row-echelon form will be the identity matrix I_n. If A's reduced row-echelon form is not the identity, then A is not invertible and you can stop.

4. If A is invertible, you should end up with an augmented matrix of the form

$$[I_n \mid B]$$

with the identity matrix on the left side instead of the right side. Then $B = A^{-1}$.

Example 3. Let's use the above algorithm to calculate the inverse of the matrix

$$A = \begin{bmatrix} 1 & 3 & 4 & 5 \\ 2 & 7 & 8 & 10 \\ 1 & 4 & 5 & 5 \\ 3 & 11 & 13 & 16 \end{bmatrix}.$$

First let's set up the augmented matrix

$$\left[\begin{array}{cccc|cccc} 1 & 3 & 4 & 5 & 1 & 0 & 0 & 0 \\ 2 & 7 & 8 & 10 & 0 & 1 & 0 & 0 \\ 1 & 4 & 5 & 5 & 0 & 0 & 1 & 0 \\ 3 & 11 & 13 & 16 & 0 & 0 & 0 & 1 \end{array}\right]$$

and start with the three row operations

- $r_2 \to r_2 - 2r_1$
- $r_3 \to r_3 - r_1$
- $r_4 \to r_4 - 3r_1$.

Then we get

$$\left[\begin{array}{cccc|cccc} 1 & 3 & 4 & 5 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 1 & -3 & 0 & 0 & 1 \end{array}\right].$$

Now we can do the next two row operations

- $r_3 \to r_3 - r_2$
- $r_4 \to r_4 - 2r_2$

and get the following matrix:

$$\left[\begin{array}{cccc|cccc} 1 & 3 & 4 & 5 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & -2 & 0 & 1 \end{array}\right].$$

We can turn the left matrix into row-echelon form from the row operation

- $r_4 \to r_4 - r_3$

and get

$$\left[\begin{array}{cccc|cccc} 1 & 3 & 4 & 5 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & 1 \end{array}\right].$$

Finally, we can get reduced row-echelon form from the three row operations

1. $r_1 \to r_1 - 5r_4$
2. $r_1 \to r_1 - 4r_3$
3. $r_1 \to r_1 - 3r_2$

and get

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

Now that the left matrix has successfully been transformed into the identity matrix via row operations, we know that the matrix on the right is the inverse of A, so

$$\begin{bmatrix} 1 & 3 & 4 & 5 \\ 2 & 7 & 8 & 10 \\ 1 & 4 & 5 & 5 \\ 3 & 11 & 13 & 16 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & 6 & 1 & -5 \\ -2 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{bmatrix}.$$

Inverses are useful because they give us another way of solving systems of linear equations (if there is a unique solution). If we have a system of equations of n equations and variables

$$a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1$$
$$\vdots$$
$$a_{n,1}x_1 + \cdots + a_{n,n}x_n = b_n$$

we can create the matrix equation

$$\begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix},$$

or $A\vec{x} = \vec{b}$ for short. If A is invertible, we can find the solution $\vec{x}$ by multiplying both sides of the equation by A^{-1}: we get $A^{-1}A\vec{x} = A^{-1}\vec{b}$, so $\vec{x} = A^{-1}\vec{b}$.

FACT: The inverse of a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Example 4. Consider the system of equations

$$5x + 3y = 2$$
$$3x + 2y = 3.$$

We get the matrix equation

$$\begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

The matrix $\begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$ has inverse $\begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}$, so multiplying both sides by the inverse yields

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -5 \\ 9 \end{bmatrix}$$

as the solution.

This brings us to an important theorem: Theorem 2.46 in the textbook.

Theorem 2.46. Let A be an $n \times n$ square matrix, and let $\vec{x}$ and $\vec{b}$ be $n \times 1$ vectors. Then the following statements are equivalent:

1. $\operatorname{rank}(A) = n$.
2. A can be transformed into I_n by elementary row operations.
3. A is invertible.
4. There exists an $n \times n$ matrix C such that $AC = CA = I_n$.
5. The system $A\vec{x} = \vec{b}$ has exactly one solution.
6. The homogeneous system $A\vec{x} = \vec{0}$ has $\vec{x} = \vec{0}$ as the only solution.

Inverses also satisfy the properties $(AB)^{-1} = B^{-1}A^{-1}$ and $(A^{-1})^T = (A^T)^{-1}$.

Definition 12. Determinant: The determinant of a square $n \times n$ matrix A– denoted $\det(A)$ or $|A|$– can be calculated recursively. To begin, the determinant of a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $ad - bc$.

Then, we can define the (i, j)-minor of a matrix A– denoted $\operatorname{minor}(A)_{i,j}$– as the determinant of the submatrix we get by deleting the i^{th} row and j^{th} column of A. Furthermore, we can then define the (i,j)-cofactor of A– denoted $\text{cof}(A)_{i,j}$– as $(-1)^{i+j}\text{minor}(A)_{i,j}$. Then, we can finally define

$$\det(A) = \sum_{j=1}^{n} a_{1,j}\text{cof}(A)_{1,j}.$$

So knowing how to calculate the determinant of an $n \times n$ matrix requires knowing how to calculate the determinant of an $(n-1) \times (n-1)$ matrix. But since we know how to calculate the determinant of a 2×2 matrix, we have a place to start!

Example 5. Let us use the definition to calculate the determinant of the matrix

$$A = \begin{bmatrix} 5 & 7 & 4 \\ 1 & 1 & 3 \\ 0 & 2 & 4 \end{bmatrix}.$$

According to the definition,

$$\begin{aligned} \det(A) = \sum_{j=1}^{3} a_{1,j}\text{cof}(A)_{1,j} &= 5\text{cof}(A)_{1,1} + 7\text{cof}(A)_{1,2} + 4\text{cof}(A)_{1,3} \\ &= 5\text{minor}(A)_{1,1} - 7\text{minor}(A)_{1,2} + 4\text{minor}(A)_{1,3}. \end{aligned}$$

Now we need to calculate $\text{minor}(A)_{1,1}$, $\text{minor}(A)_{1,2}$, and $\text{minor}(A)_{1,3}$.

First, $\text{minor}(A)_{1,1}$ is the determinant of the matrix we get from ignoring the first row and first column of A, so it is the determinant of $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$. This determinant is $1 * 4 - 3 * 2 = -2$.

Second, $\text{minor}(A)_{1,2}$ is the determinant of the matrix we get from ignoring the first row and second column of A, so it is the determinant of $\begin{bmatrix} 1 & 3 \\ 0 & 4 \end{bmatrix}$. This determinant is $1 * 4 - 3 * 0 = 4$.

Finally, $\text{minor}(A)_{1,3}$ is the determinant of $\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$, which is $1 * 2 - 1 * 0 = 2$.

So putting everything together,

$$\det(A) = 5\text{minor}(A)_{1,1} - 7\text{minor}(A)_{1,2} + 4\text{minor}(A)_{1,3} = 5 * (-2) - 7 * (4) + 4 * (2) = -30.$$

The determinant has many useful properties.

- $\det(AB) = \det(A)\det(B)$.
- $\det(I_n) = 1$ where I_n is the $n \times n$ identity matrix.
- $\det(A^k) = \det(A)^k$.
- $\det(A^T) = \det(A)$.
- The determinant of a matrix also gives you information about its invertibility. A matrix A has an inverse if and only if $\det(A) \neq 0$. (In fact, according to the third bullet point, $\det(A^{-1}) = \det(A)^{-1}$, which exists if and only if $\det(A) \neq 0$.)

Definition 13. Elementary matrix: An elementary matrix is a square $n \times n$ matrix E such that multiplying any $n \times m$ matrix A by E *on its left* has the effect of performing a row operation on A.

Example 6. For example the matrix

$$E = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is the elementary matrix corresponding to $r_1 \leftrightarrow r_2$ for a 3×3 matrix. We can observe its effect on $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ below:

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

In general, you can always build an elementary matrix for a given row operation by performing that row operation on the identity matrix. For example, if we want to identify the elementary matrix that corresponds to the row operation $r_3 \to r_3 - 3r_1$ for 3×3 matrices, we can take the identity matrix $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and perform $r_3 \to r_3 - 3r_1$. We end up with

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

as the elementary matrix.

By calculating the determinant of each elementary matrix and using the property that $\det(EA) = \det(E)\det(A)$, we can determine the effect of each row operation on the determinant of A.

1. $r_i \leftrightarrow r_j$: Multiplies $\det(A)$ by -1.

2. $r_i \to sr_i$: Multiplies $\det(A)$ by s.

3. $r_i \to r_i + tr_j$: No effect on $\det(A)$.

Definition 14. Linear combination: A vector $\vec{v}$ is a linear combination of the vectors $\vec{u}_1, \ldots, \vec{u}_n$ if there are scalars $s_1, \ldots, s_n$ such that $\vec{v} = s_1\vec{u}_1 + \cdots + s_n\vec{u}_n$.

Example 7. The vector $\begin{bmatrix} 2 \\ 2 \\ 17 \end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}$, $\begin{bmatrix} 9 \\ 7 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 4 \\ 1 \\ 4 \end{bmatrix}$ because

$$\begin{bmatrix} 2 \\ 2 \\ 17 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} - \begin{bmatrix} 9 \\ 7 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 4 \\ 1 \\ 4 \end{bmatrix}.$$

Definition 15. Span: The span of a set of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is the set of all linear combinations of $\vec{v}_1, \ldots, \vec{v}_n$.

Example 8. Is $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \in \text{span}\left\{\begin{bmatrix} 3 \\ 9 \\ 12 \end{bmatrix}, \begin{bmatrix} 6 \\ 7 \\ 13 \end{bmatrix}, \begin{bmatrix} 9 \\ 0 \\ 9 \end{bmatrix}\right\}$? To answer this, we need to know if $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ is a linear combination of $\begin{bmatrix} 3 \\ 9 \\ 12 \end{bmatrix}$, $\begin{bmatrix} 6 \\ 7 \\ 13 \end{bmatrix}$, and $\begin{bmatrix} 9 \\ 0 \\ 9 \end{bmatrix}$

. We can set up the equation

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = a\begin{bmatrix} 3 \\ 9 \\ 12 \end{bmatrix} + b\begin{bmatrix} 6 \\ 7 \\ 13 \end{bmatrix} + c\begin{bmatrix} 9 \\ 0 \\ 9 \end{bmatrix}$$

which can be turned into the system of linear equations

$$\begin{aligned} 3a + 6b + 9c &= 1 \\ 9a + 7b + 0c &= 1 \\ 12a + 13b + 9c &= 0 \end{aligned}$$

and the augmented matrix

$$\left[\begin{array}{ccc|c} 3 & 6 & 9 & 1 \\ 9 & 7 & 0 & 1 \\ 12 & 13 & 9 & 0 \end{array}\right].$$

We can begin getting this augmented matrix into reduced row-echelon form, but just after the three row operations

1. $r_2 \to r_2 - 3r_1$
2. $r_3 \to r_3 - 4r_1$
3. $r_3 \to r_3 - r_2$

we get the augmented matrix

$$\left[\begin{array}{ccc|c} 3 & 6 & 9 & 1 \\ 0 & -11 & -27 & -2 \\ 0 & 0 & 0 & -2 \end{array}\right].$$

The last row of this matrix says $0a + 0b + 0c = -2$, which cannot possibly be true. Thus there is no solution to this system of equations, so we can conclude that

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \notin \operatorname{span}\left\{ \begin{bmatrix} 3 \\ 9 \\ 12 \end{bmatrix}, \begin{bmatrix} 6 \\ 7 \\ 13 \end{bmatrix}, \begin{bmatrix} 9 \\ 0 \\ 9 \end{bmatrix} \right\}.$$

Definition 16. Linear independence and dependence: A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is linearly independent if the linear combination $s_1\vec{v}_1 + \cdots + s_n\vec{v}_n$ results in $\vec{0}$ ***only*** if $s_i = 0$ for all $1 \leq i \leq n$.

On the contrary, the set is linearly dependent if $s_1\vec{v}_1 + \cdots + s_n\vec{v}_n$ can be $\vec{0}$ while allowing some $s_i \neq 0$.

Example 9. The set $\left\{ \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}, \begin{bmatrix} 9 \\ 7 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 4 \end{bmatrix} \right\}$ is linearly independent. We can verify this by setting up the linear combination

$$s_1 \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} + s_2 \begin{bmatrix} 9 \\ 7 \\ 0 \end{bmatrix} + s_3 \begin{bmatrix} 4 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and verifying that there is ONE solution for s_1, s_2, and s_3: specifically that $s_1 = 0$, $s_2 = 0$, and $s_3 = 0$ is the only solution.

We can make the augmented matrix

$$\left[\begin{array}{ccc|c} 3 & 9 & 4 & 0 \\ 6 & 7 & 1 & 0 \\ 9 & 0 & 4 & 0 \end{array}\right].$$

We can turn this matrix into reduced row-echelon form from the following ten row operations:

1. $r_2 \to r_2 - 2r_1$
2. $r_3 \to r_3 - 3r_1$
3. $r_3 \to -11r_3$
4. $r_3 \to r_3 + 27r_2$

5. $r_1 \to r_1/3$

6. $r_2 \to -r_2/11$

7. $r_3 \to -r_3/101$

8. $r_2 \to r_2 - 7/4r_3$

9. $r_1 \to r_1 - 4/3r_3$

10. $r_1 \to r_1 - 3r_2$

and end up with

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

Therefore the only solution to this system of linear equations is $s_1 = 0$, $s_2 = 0$, and $s_3 = 0$. This means that $\left\{ \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}, \begin{bmatrix} 9 \\ 7 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 4 \end{bmatrix} \right\}$ is linearly independent.

Another useful way to think about linear dependence is that $\{\vec{v}_1 \dots \vec{v}_n\}$ is linearly dependent if one of the vectors $\vec{v}_i$ can be written as a linear combination as some of the other vectors. And the set is linearly independent if none of the vectors in $\{\vec{v}_1, \dots, \vec{v}_n\}$ can be written as a linear combination of any of the others.

A more computational way you can determine linear dependence and independence of a set of n vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ is by computing the rank of the matrix you get by fusing together $\vec{v}_1, \dots, \vec{v}_n$ as columns. Specifically

$$\text{rank}[\vec{v}_1 \cdots \vec{v}_n] = n \text{ if and only if } \{\vec{v}_1, \dots, \vec{v}_n\} \text{ is linearly independent}$$

and

$$\text{rank}[\vec{v}_1 \cdots \vec{v}_n] < n \text{ if and only if } \{\vec{v}_1, \dots, \vec{v}_n\} \text{ is linearly dependent.}$$

Definition 17. Vector Space: A vector space is a set of vectors V that satisfies the following three criteria:

1. $V \neq \varnothing$. (V is not the empty set, meaning V must actually contain something.)
2. V is *closed under addition*: if $\vec{v}$ and $\vec{u}$ are in V, then $\vec{v} + \vec{u} \in V$.
3. V is *closed under scalar multiplication*: if $\vec{v} \in V$ and $s \in \mathbb{R}$, then $s\vec{v} \in V$.

We have seen multiple examples of vector spaces so far:

- $\mathbb{R}^n$ is a vector space for all n.
- For any set of vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$, $\text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$ is a vector space.
- The solution set of any **homogeneous** system of equations is a vector space.

NOTE: The solution set of a non-homogeneous system of equations is **never** a vector space.

Definition 18. Basis: Given a vector space V, a basis for V is a set of vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ that satisfies the following two criteria:

1. $\{\vec{v}_1, \dots, \vec{v}_n\}$ is linearly independent.
2. $\text{span}\{\vec{v}_1, \dots, \vec{v}_n\} = V$.

For a vector space V, every basis for V will always have the same number of vectors. The **dimension** of V is the number of vectors in a basis for V.

The vector space $\mathbb{R}^n$ has the **standard basis** $\{\vec{e}_1, \dots, \vec{e}_n\}$ where $\vec{e}_i$ has a 1 in the i^{th} entry and 0 everywhere else. For example, the standard basis for $\mathbb{R}^2$ is $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ and the standard basis for $\mathbb{R}^3$ is $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Example 10. The set $\left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 10 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^2$. We can check that the set is linearly independent because

$$\text{rank} \begin{bmatrix} 2 & -1 \\ 3 & 10 \end{bmatrix} = 2$$

and we can check that the set spans $\mathbb{R}^2$ by setting up this system with an arbitrary vector $\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2$:

$$\begin{bmatrix} a \\ b \end{bmatrix} = r \begin{bmatrix} 2 \\ 3 \end{bmatrix} + s \begin{bmatrix} -1 \\ 10 \end{bmatrix}$$

and determining if there is a solution for r and s. Note that this turns into the matrix equation

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 3 & 10 \end{bmatrix} \begin{bmatrix} r \\ s \end{bmatrix}.$$

Since we already know that the rank of that 2×2 matrix is 2, Theorem 2.46 tells us that the matrix equation has a unique solution. Since there is a solution, that means that any vector in $\mathbb{R}^2$ is a linear combination of $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 10 \end{bmatrix}$, and so span $\left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 10 \end{bmatrix} \right\} = \mathbb{R}^2$.

Becuase it meets the two necessary criteria, $\left\{ \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 10 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^2$.

Definition 19. Nullspace and Nullity: Given an $m \times n$ matrix A, the set of all $n \times 1$ vectors $\vec{x}$ that satisfy the equation

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

is called the **nullspace** of A, denoted null(A). In set notation,

$$\text{null}(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}.$$

The nullspace of A is a vector space, and its dimension is called the **nullity** of A.

Example 11. Let's find a basis for the nullspace of

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

The nullspace is the set of solutions to

$$\begin{bmatrix} 1 & 0 & 0 & -6 & 2 \\ 0 & 2 & 0 & -10 & 2 \\ 0 & 0 & 3 & 0 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

for which we can set up the augmented matrix

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & -6 & 2 & 0 \\ 0 & 2 & 0 & -10 & 2 & 0 \\ 0 & 0 & 3 & 0 & 9 & 0 \end{array}\right].$$

This matrix is nearly in reduced row-echelon form already: all we need to do is $r_2 \to r_2/2$ and $r_3 \to r_3/3$ and we get

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

So x_1, x_2, and x_3 are our basic variables and x_4 and x_5 are our free variables. Because we have two free variables, we already know that the solution set is two-dimensional. We can parametrize with $x_4 = s$ and $x_5 = t$, and then the complete paramatrization for the solution set is

$$x_1 = 6s - 2t$$
$$x_2 = 5s - t$$
$$x_3 = -3t$$
$$x_4 = s$$
$$x_5 = t.$$

The above is a parametrization for null(A). We can construct a basis for null(A) by turning to a different two-dimensional vector space, $\mathbb{R}^2$, and stealing a basis from that. The easiest thing to do is to steal the standard basis $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ and copy $\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

When $\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 6 \\ 5 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

and when $\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ we get

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ -3 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore

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

is a basis for null(A).

Definition 20. Column & Row Space: Given a matrix A, the **column space** of A, denoted col(A), is the span of the columns of A, each treated like its own vector. The **row space** of A, denoted row(A), is equal to col(A^T).

For any matrix A, $\dim(\text{col}(A)) = \dim(\text{row}(A)) = \text{rank}(A)$.

Theorem (Rank-Nullity Theorem). For any $m \times n$ matrix A,

$$\text{rank}(A) + \text{nullity}(A) = n.$$

Definition 21. Image: For an $m \times n$ matrix A, the **image** of A, denoted img(A) is the set of $m \times 1$ vectors $\vec{v} \in \mathbb{R}^m$ that are multiples of A: that is, there is some $n \times 1$ vector $\vec{u} \in \mathbb{R}^n$ where $\vec{v} = A\vec{u}$. In set notation,

$$\text{img}(A) = \{\vec{v} \in \mathbb{R}^m : \vec{v} = A\vec{u} \text{ for some } \vec{u} \in \mathbb{R}^n\}.$$

The image of a matrix A is always the same thing as the column space of A. Therefore $\dim(\text{img}(A)) = \text{rank}(A)$, just like the column space and row space.

Definition 22. Orthogonal & Orthonormal Basis: Let V be a vector space. An **orthogonal basis** for V is a basis $B = \{\vec{b}_1, \ldots, \vec{b}_n\}$ such that $\vec{b}_i \cdot \vec{b}_j = 0$ for all $i \neq j$. In other words, all the vectors in the basis are perpendicular (aka orthogonal) to each other.

Furthermore, B is an **orthonormal basis** if the vectors of B satisfy the additional condition $\|b_i\| = 1$ for all i. In other words, B is orthonormal if all the vectors of B are perpendicular to each other and they all have a magnitude of 1.

Definition 23. Fourier expansion: Given an orthogonal basis $B = \{\vec{b}_1, \ldots, \vec{b}_n\}$ for a vector space V, the **Fourier expansion** is a formula used to write any vector $\vec{v} \in V$ as a linear combination of the basis vectors in B, as follows:

$$\vec{v} = \left(\frac{\vec{v} \cdot \vec{b}_1}{\|\vec{b}_1\|^2}\right)\vec{b}_1 + \left(\frac{\vec{v} \cdot \vec{b}_2}{\|\vec{b}_2\|^2}\right)\vec{b}_2 + \cdots + \left(\frac{\vec{v} \cdot \vec{b}_n}{\|\vec{b}_n\|^2}\right)\vec{b}_n.$$

Furthermore, when B is orthonormal, we know that $\|\vec{b}_i\| = 1$, so the Fourier exansion simplifies to

$$\vec{v} = (\vec{v} \cdot \vec{b}_1)\vec{b}_1 + (\vec{v} \cdot \vec{b}_2)\vec{b}_2 + \cdots + (\vec{v} \cdot \vec{b}_n)\vec{b}_n.$$

Example 12. The set $B = \left\{ \begin{bmatrix} 5 \\ 6 \end{bmatrix}, \begin{bmatrix} -12 \\ 10 \end{bmatrix} \right\}$ is an orthogonal basis for $\mathbb{R}^2$. If we want to write $\begin{bmatrix} 7 \\ 9 \end{bmatrix}$ as a linear combination of the basis vectors, we can use the Fourier expansion

$$\begin{bmatrix} 7 \\ 9 \end{bmatrix} = \left(\frac{\begin{bmatrix} 7 \\ 9 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 6 \end{bmatrix}}{\left\| \begin{bmatrix} 5 \\ 6 \end{bmatrix} \right\|^2} \right) \begin{bmatrix} 5 \\ 6 \end{bmatrix} + \left(\frac{\begin{bmatrix} 7 \\ 9 \end{bmatrix} \cdot \begin{bmatrix} -12 \\ 10 \end{bmatrix}}{\left\| \begin{bmatrix} -12 \\ 10 \end{bmatrix} \right\|^2} \right) \begin{bmatrix} -12 \\ 10 \end{bmatrix} = \frac{89}{61} \begin{bmatrix} 5 \\ 6 \end{bmatrix} + \frac{6}{244} \begin{bmatrix} -12 \\ 10 \end{bmatrix}.$$

Definition 24. Orthogonal Matrix: A square $n \times n$ matrix Q is **orthogonal** if the columns of Q form an orthonormal basis of $\mathbb{R}^n$, when treated like their own vectors. (It would make more sense to call Q an orthonormal matrix, and some people do, but most people don't).

If Q is an orthogonal matrix, then as a result the rows of Q also form an orthogonal matrix of $\mathbb{R}^n$.

Furthermore, Q is an orthogonal matrix if and only if $Q^{-1} = Q^T$.