Preliminaries: Sets

Notation: Let's set some standard notation.

We let $\mathbb{Z}$ be the set of integers: $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

We let $\mathbb{Q}$ be the set of rational numbers: $\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z} \text{ and } q \neq 0 \right\}$.

We let $\mathbb{R}$ be the set of real numbers.

More Notation:

The symbol $\in$ means "is an element of."

The symbol $\notin$ means "is not an element of."

The symbol $\subset$ means "is a subset of."

In the context of a set, the symbol $\mid$ means "such that."

Example 1: Let $A = \{2, 3, \text{Drake}\}$.

Notice that $2 \in A$ and $\text{Drake} \in A$, but $4 \notin A$.

Notice that $\{2, \text{Drake}\} \subset A$.

Example 2: Evaluate the following statements as True or False.

(a) The statement $\sqrt{2} \in \mathbb{R}$ is: True

(b) The statement $\sqrt{2} \in \mathbb{Z}$ is: False

(c) The statement $\text{Yeezy} \in \mathbb{Z}$ is: False

(d) The statement $\mathbb{Z} \subset \mathbb{R}$ is: True.

Example 3:

(a) $\left\{ x \in \mathbb{Z} \mid x^2 = 25 \right\} = \{5, -5\}$

(b) $\left\{ x \in \mathbb{Z} \mid x^2 = -1 \right\} = \emptyset$

(c) $\left\{ x^2 + 3 \mid x \in \{-1, 1, 2\} \right\} = \{4, 7\}$

Preliminaries: Functions

Informally speaking, a *function* is an input-output rule, with the requirement that: For every input, there exists exactly one output.

- ◦ Domain of a function: The set of inputs.
- ◦ Codomain of a function: The set of possible outputs.
- ◦ Range of a function: The set of actual outputs.

Note: The range is a subset of the codomain.

Notation: We write $f\colon A \to B$ to indicate that A is the domain of f, and B is the codomain of f.

Example 1: Consider the following "birth year" function:

$$B\colon \{\text{Jesse, T-Swizzle}\} \to \mathbb{Z}$$
$$B(\text{person}) = \text{year that "person" was born.}$$

- ◦ The domain is: {Jesse, T-Swizzle}
- ◦ The codomain is: $\mathbb{Z}$
- ◦ The range is: $\{1990, 1989\}$

Example 2: Consider the function

$$h\colon \mathbb{Z} \to \mathbb{R}$$
$$h(x) = x^2.$$

- ◦ The domain is: $\mathbb{Z}$
- ◦ The codomain is: $\mathbb{R}$
- ◦ The range is: $\{0, 1, 4, 9, 16, \ldots\}$

Example 3: The function

$$F\colon \mathbb{R} \to \mathbb{R}$$
$$F(x) = \ln(x)$$

is **not** well-defined! Functions need to be able to have exactly one output for *every* element in the domain. But here, $F(x) = \ln(x)$ is undefined for $x \leq 0$.

However, the function

$$G\colon (0, \infty) \to \mathbb{R}$$
$$G(x) = \ln(x)$$

is well-defined.

Vectors

Def: A **scalar** is an element of $\mathbb{R}$. i.e.: A scalar is literally just a real number.

A **vector** is an element of $\mathbb{R}^n$.

$$\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \;\middle|\; x_1, \ldots, x_n \in \mathbb{R} \right\}.$$

i.e.: A vector is an ordered n-tuple of real numbers.

Operations:

- ∘ Addition (of vectors with vectors)
- ∘ Subtraction (of vectors with vectors)
- ∘ Scaling (of a scalar with a vector)

Remarks:

∘ Each of these three operations has a geometric interpretation.

∘ It does not make sense (for example) to add a scalar to a vector. Vectors can be added/subtracted to other vectors.

Def: A **linear combination** of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ is any vector of the form

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k,$$

where $c_1, \ldots, c_k \in \mathbb{R}$ are scalars.

Example: Let $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^{17}$ be three vectors in $\mathbb{R}^{17}$. Here are some linear combinations of $\mathbf{a}, \mathbf{b}, \mathbf{c}$:

$$2\mathbf{a} + \mathbf{b} + 3\mathbf{c}, \qquad \mathbf{0} = 0\mathbf{a} + 0\mathbf{b} + 0\mathbf{c}, \qquad \mathbf{a} - \mathbf{b}.$$

Span & Linear Independence

The basic notions in linear algebra arise from the twin concepts of "span" and "linear independence." Today, we'll just define span.

Def: The **span** of a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_k$. That is:

$$\text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \{c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \mid c_1, \ldots, c_k \in \mathbb{R}\}.$$

Span

Rough Idea: The span of a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is the "smallest" "subspace" of $\mathbb{R}^n$ containing $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

This is not very precise as stated (e.g., what is meant by "subspace"?). Here is the precise definition:

Def: The **span** of a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_k$. That is:

$$\text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \{c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \mid c_1, \ldots, c_k \in \mathbb{R}\}.$$

Linear Independence

The definition in the textbook is:

Def: Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a set of at least two vectors.

The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is **linearly independent** if none of the vectors in the set is a linear combination of the others.

The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is **linearly dependent** if at least one of the vectors in the set is a linear combination of the others.

Def: Let $\{\mathbf{v}_1\}$ be a set of just one vector.

It is **linearly independent** if $\mathbf{v}_1 \neq \mathbf{0}$. It is **linearly dependent** if $\mathbf{v}_1 = \mathbf{0}$.

There is also an **equivalent** definition, which is somewhat more standard:

Def: Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a set of (any number of) vectors.

The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is **linearly independent** if the only linear combination $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ equal to the zero vector is the one with $c_1 = \cdots = c_k = 0$.

The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is **linearly dependent** if there is a linear combination $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ equal to the zero vector, where **not all** the scalars $c_1, \ldots, c_k$ are zero.

Point: Linear independence of $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ means:

$$\text{If } c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}, \text{ then } c_1 = \cdots = c_k = 0.$$

This way of phrasing linear independence is very useful for proofs.

Linear Independence: Intuition

Why is "linear independence" a concept one would want to define? What does it mean intuitively? The following examples may help explain.

Example 1: The set $\text{span}(\mathbf{v})$ is one of the following:
(i) A line through $\mathbf{0}$.
(ii) The origin itself.
Further: The first case (i) holds if and only if $\{\mathbf{v}\}$ is linearly independent. Otherwise, the other case holds.

Example 2: The set $\text{span}(\mathbf{v}_1, \mathbf{v}_2)$ is one of the following:
(i) A plane through $\mathbf{0}$.
(ii) A line through $\mathbf{0}$.
(iii) The origin itself.
Further: The first case (i) holds if and only if $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent. Otherwise, one of the other cases holds.

Example 3: The set $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ is one of the following:
(i) A "3-dimensional space" through $\mathbf{0}$.
(ii) A plane through $\mathbf{0}$.
(iii) A line through $\mathbf{0}$.
(iv) The origin itself.
Further: The first case (i) holds if and only if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent. Otherwise, one of the other cases holds.

Q: Do you see the pattern here? What are the possibilities for the span of four vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$? Seven vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_7\}$?

Q: Looking at Example 3, what happens if the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are in $\mathbb{R}^2$? Can possibility (i) occur in that case? What does this tell you about sets of three vectors in $\mathbb{R}^2$?

Lines and Planes: Parametric Descriptions

Parametric form of a Line in $\mathbb{R}^n$: The parametric form of a line in $\mathbb{R}^n$ passing through $\mathbf{p}$ with direction $\mathbf{v}$ is

$$\begin{aligned} L &= \{\mathbf{p} + t\mathbf{v} \mid t \in \mathbb{R}\} \\ &= \left\{ \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix} + t \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \;\middle|\; t \in \mathbb{R} \right\} \end{aligned}$$

We call t the parameter.

Equivalently:

$$\begin{cases} x_1 = p_1 + tv_1 \\ \vdots \\ x_n = p_n + tv_n. \end{cases}$$

Parametric form of a Plane in $\mathbb{R}^n$: The parametric form of a plane in $\mathbb{R}^n$ passing through $\mathbf{p}$ and parallel to $\text{span}(\mathbf{v}, \mathbf{w})$ is

$$\begin{aligned} P &= \{\mathbf{p} + s\mathbf{v} + t\mathbf{w} \mid s, t \in \mathbb{R}\} \\ &= \left\{ \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix} + s \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} + t \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \;\middle|\; s, t \in \mathbb{R} \right\} \end{aligned}$$

We call s and t the parameters.

Q: Do you see the pattern here? How might one describe the parametric form of a 3-dim space in $\mathbb{R}^n$ passing through $\mathbf{p}$ and parallel to $\text{span}(\mathbf{u}, \mathbf{v}, \mathbf{w})$?

Level Set form of a Plane in $\mathbb{R}^3$: The equation for a plane in $\mathbb{R}^3$ passing through $\mathbf{x_0} = (x_0, y_0, z_0)$ with normal vector $\mathbf{n} = (n_1, n_2, n_3)$ is:

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{x}_0) = 0. \tag{$*$}$$

(Q: Why does this describe a plane?)

In other words:

$$n_1(x - x_0) + n_2(y - y_0) + n_3(z - z_0) = 0.$$

(Q: How is this the same as the formula $(*)$?)

Dot Products: Algebra

Def: The **dot product** of two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ is

$$\mathbf{v} \cdot \mathbf{w} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = v_1 w_1 + \cdots + v_n w_n.$$

The **length** (or **norm**) of a vector $\mathbf{v} \in \mathbb{R}^n$ is:

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

(Q: Why is this a reasonable definition of length?)

Observation: Notice that

$$\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2.$$

Inequalities: For any non-zero vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, we have:

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\| \qquad \text{(Triangle Inequality)}$$

$$|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\|\|\mathbf{w}\|. \qquad \text{(Cauchy-Schwarz Inequality)}$$

In both, equality holds if and only if $\mathbf{w} = c\mathbf{v}$ for some non-zero scalar c.

Dot Products: Geometry

Prop: Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ be non-zero vectors. Then:

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\| \cos\theta,$$

where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$.

Therefore, $\mathbf{v}$ and $\mathbf{w}$ are perpendicular if and only if $\mathbf{v} \cdot \mathbf{w} = 0$.

Def: We say that two vectors $\mathbf{v}, \mathbf{w}$ are **orthogonal** if $\mathbf{v} \cdot \mathbf{w} = 0$.

Pythagorean Theorem: If $\mathbf{v}$ and $\mathbf{w}$ are orthogonal, then

$$\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2.$$

(Q: How exactly is this the "Pythagorean Theorem" about right triangles?)

Cross Products (Optional but Useful)

Def: The **cross product** of two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ is

$$\mathbf{v} \times \mathbf{w} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \times \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{bmatrix}.$$

Note: Dot products make sense in $\mathbb{R}^n$ for any dimension n.
But: Cross products only make sense in $\mathbb{R}^3$.

Prop: Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. Then:

$$\mathbf{v} \cdot (\mathbf{v} \times \mathbf{w}) = 0$$
$$\mathbf{w} \cdot (\mathbf{v} \times \mathbf{w}) = 0.$$

In other words: $\mathbf{v} \times \mathbf{w}$ is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$.

Prop: Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ be non-zero vectors. Then

$$\|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\|\|\mathbf{w}\| \sin\theta,$$

where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$.

Prop: Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. Then:

$$\text{Area(Parallelogram formed by } \mathbf{v} \text{ and } \mathbf{w}) = \|\mathbf{v} \times \mathbf{w}\|.$$

Linear Systems as Matrix-Vector Products

A **linear system** of m equations in n unknowns is of the form:

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

We can write a linear system as a single vector equation:

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

The **coefficient matrix** of the system is the $m \times n$ matrix

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

The **matrix-vector product** of the $m \times n$ matrix A with the vector $\mathbf{x} \in \mathbb{R}^n$ is the vector $A\mathbf{x} \in \mathbb{R}^m$ given by:

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

We can now write the system $(*)$ as:

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

Homogeneous vs Inhomogeneous

Def: A linear system of the form $A\mathbf{x} = \mathbf{0}$ is called **homogeneous**.
A linear system of the form $A\mathbf{x} = \mathbf{b}$ with $\mathbf{b} \neq \mathbf{0}$ is called **inhomogeneous**.

Fact: Every homogeneous system $A\mathbf{x} = \mathbf{0}$ has at least one solution (why?).
$\therefore$ For homogeneous systems: only cases (B) and (C) of Prop 6.2 can occur.

Reduced Row Echelon Form; Solutions of Systems

Row Operations:

(1) Multiply/divide a row by a non-zero scalar.
(2) Add/subtract a scalar multiple of one row from another row.
(3) Exchange two rows.

Facts:

(a) Row operations do not change the set of solutions of a linear system.
(b) Using row operations, every matrix can be put in **reduced row echelon form**.

Def: A matrix is in **reduced row echelon form** if:

(1) The first non-zero entry in each row is 1. (These 1's are called **pivots**.)
(2) Each pivot is further to the right than the pivot of the row above it.
(3) In the column of a pivot, all other entries are zero.
(4) Rows containing all zeros are at the very bottom.

Def: Given a linear system of equations (whose augmented matrix is) in reduced row echelon form.

The variables whose corresponding column contains a pivot are called **pivot variables**. The other variables are called **free variables**.

Note: For an $m \times n$ matrix (i.e., m rows and n columns), we have:

$$(\#\text{ of pivot variables}) + (\#\text{ of free variables}) = n.$$

This basic fact is surprisingly important!

Prop 6.2: For a linear system of equations (whose augmented matrix is) in reduced row echelon form, there are three possibilities:

(A) **No solutions.** One of the equations is $0 = 1$.
(B) **Exactly one solution.** There's no $0 = 1$, and no free variables.
(C) **Infinitely many solutions.** There's no $0 = 1$, but there's at least one free variable.

Geometrically: The solution set looks like one of:

(A) The empty set. (i.e.: The set $\{\}$ with nothing inside it.)
(B) A single vector.
(C) A line, or a plane, or a 3-dimensional space, or... etc.

Null Space

Def: Let A be an $m \times n$ matrix, so $A\colon \mathbb{R}^n \to \mathbb{R}^m$.
The **null space** of A is:

$$N(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}.$$

So: $N(A)$ is the set of solutions to the linear system $A\mathbf{x} = \mathbf{0}$.

Fact: Either $A\mathbf{x} = \mathbf{b}$ has no solutions, or at least one solution (logic!).
If $A\mathbf{x} = \mathbf{b}$ has at least one solution, then the solution set of $A\mathbf{x} = \mathbf{b}$ is a translation of $N(A)$. Therefore, in this case:
◦ $A\mathbf{x} = \mathbf{0}$ has exactly one solution $\iff$ $A\mathbf{x} = \mathbf{b}$ has exactly one solution.
◦ $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions $\iff$ $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions.
⋆ Careful: This fact assumes $A\mathbf{x} = \mathbf{b}$ has at least one solution. If $A\mathbf{x} = \mathbf{b}$ has no solutions, then we cannot draw these conclusions!

Column Space

There are two **equivalent** definitions of the column space.

Def 1: Let A be an $m \times n$ matrix. Let A have columns $[\mathbf{v}_1 \cdots \mathbf{v}_n]$.
The **column space** of A is

$$C(A) = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

So: the column space is the span of the columns of A.

Def 2: Let A be an $m \times n$ matrix, so $A\colon \mathbb{R}^n \to \mathbb{R}^m$.
The **column space** of A is

$$C(A) = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}.$$

So: the column space is just the range of A. (i.e., the set of all actual outputs.)

Therefore: The linear system $A\mathbf{x} = \mathbf{b}$ has a solution $\iff$ $\mathbf{b} \in C(A)$.

Important:
◦ $N(A)$ is a subspace of the domain of A.
◦ $C(A)$ is a subspace of the codomain of A.

Two Crucial Facts

Fact 1 (Prop 8.3): Let A be an $m \times n$ matrix. The following are equivalent:
(i) $N(A) = \{\mathbf{0}\}$.
(ii) The columns of A are linearly independent.
(iii) rref(A) has a pivot in each column.
Further: If any of these hold, then $n \leq m$.

Fact 2 (Prop 9.2): Let A be an $m \times n$ matrix. The following are equivalent:
(i) $C(A) = \mathbb{R}^m$.
(ii) The columns of A span $\mathbb{R}^m$.
(iii) rref(A) has a pivot in each row.
Further: If any of these hold, then $n \geq m$.

Subspaces

Def: A **(linear) subspace** of $\mathbb{R}^n$ is a subset $V \subset \mathbb{R}^n$ such that:
(1) $\mathbf{0} \in V$.
(2) If $\mathbf{v}, \mathbf{w} \in V$, then $\mathbf{v} + \mathbf{w} \in V$.
(3) If $\mathbf{v} \in V$, then $c\mathbf{v} \in V$ for all scalars $c \in \mathbb{R}$.

N.B.: For a subset $V \subset \mathbb{R}^n$ to be a (linear) subspace, all three properties must hold. If any one fails, then the subset V is not a (linear) subspace!

Fact: For any $m \times n$ matrix A:
(a) $N(A)$ is a subspace of $\mathbb{R}^n$.
(b) $C(A)$ is a subspace of $\mathbb{R}^m$.

So, the set of solutions to $A\mathbf{x} = \mathbf{0}$ is a linear subspace. But what about the set of solutions to $A\mathbf{x} = \mathbf{b}$? Assuming there are solutions to $A\mathbf{x} = \mathbf{b}$, then the set of solutions is an *affine subspace*.

Def: An **affine subspace** of $\mathbb{R}^n$ is a translation of a (linear) subspace.

Important: In this class, when we say "subspace," we mean *linear subspace*. This is more specific than the broader concept of "affine subspace."

Solutions of Linear Systems (again)

For a linear system $A\mathbf{x} = \mathbf{b}$, there are three possibilities:

No solutions	There is a $0 = 1$ equation	$\mathbf{b} \notin C(A)$
Exactly one solution	No $0 = 1$ equation, and No free variables	$\mathbf{b} \in C(A)$ and $N(A) = \{\mathbf{0}\}$
Infinitely many solutions	No $0 = 1$ equation, and At least one free variable	$\mathbf{b} \in C(A)$ and $N(A) \neq \{\mathbf{0}\}$

Basis & Dimension

Def: A **(linear) subspace** of $\mathbb{R}^n$ is a subset $V \subset \mathbb{R}^n$ such that:
(i) $\mathbf{0} \in V$.
(ii) If $\mathbf{v}, \mathbf{w} \in V$, then $\mathbf{v} + \mathbf{w} \in V$.
(iii) If $\mathbf{v} \in V$, then $c\mathbf{v} \in V$ for all scalars $c \in \mathbb{R}$.

Def: A **basis** for a subspace $V \subset \mathbb{R}^n$ is a set of vectors $\{\mathbf{v_1}, \ldots, \mathbf{v}_k\}$ such that:
(1) $V = \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$.
(2) $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent.

∘ Condition (1) ensures that every vector $\mathbf{v}$ in the subspace V can be written as a linear combination of the basis elements: $\mathbf{v} = x_1\mathbf{v}_1 + \cdots + x_k\mathbf{v}_k$.
∘ Condition (2) ensures that these coefficients are unique – that is, for a given vector $\mathbf{v}$, there is only one possible choice of $x_1, \ldots, x_k$.

Def: The **dimension** of a subspace $V \subset \mathbb{R}^n$ is the number of elements in any basis for V.

But what if one basis for V has (say) 5 elements, but another basis for V had 7 elements? Then how could we make sense of the dimension of V? Fortunately, that can never happen, because:

Fact: For a given subspace, every basis has the same number of elements.

Rank-Nullity Theorem: Let A be an $m \times n$ matrix, so $A\colon \mathbb{R}^n \to \mathbb{R}^m$. Then

$$\dim(C(A)) + \dim(N(A)) = n.$$

This is fantastic! (We call $\dim(C(A))$ the **rank**, and $\dim(N(A))$ the **nullity**.)

Finding Bases for Null Spaces & Column Spaces

Given a matrix A. How can we find a basis for the null space $N(A)$? How can we find a basis for the column space $C(A)$?

Basis of $N(A)$: Solve $A\mathbf{x} = \mathbf{0}$ for the pivot variables in terms of the free variables. This will give a basis for $N(A)$.

Basis of $C(A)$: The pivot columns of $\text{rref}(A)$ form a basis for $C(\text{rref}(A))$. The *corresponding columns* of A form a basis for $C(A)$.

Careful: For every matrix A, we have $N(\text{rref}(A)) = N(A)$.
However, the column spaces $C(\text{rref}(A))$ and $C(A)$ are usually different.

Rank-Nullity Theorem (revisited)

Fact: Let A be an $m \times n$ matrix. Then:

$$(\#\text{ of free variables}) = \dim(N(A)).$$
$$(\#\text{ of pivot variables}) = \dim(C(A)).$$

So: The rank-nullity theorem

$$\dim(C(A)) + \dim(N(A)) = n$$

is exactly the same as the statement

$$(\#\text{ of pivot variables}) + (\#\text{ of free variables}) = n.$$

Dimension (revisited): Linear Independence & Span

Fact: Let $V \subset \mathbb{R}^n$ be a k-dim subspace. Let S be a set of vectors in V.
(a) If S is linearly independent, then S has <u>at most</u> k elements.
(b) If S spans V, then S has <u>at least</u> k elements.
(c) If S is a basis of V, then S has <u>exactly</u> k elements.

Moral: We can think of a basis for a subspace $V \subset \mathbb{R}^n$ as a set of vectors $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ in V which is...
(1) A lin-indep set in V having the maximum possible number of vectors.
(2) A spanning set of V having the minimum possible number of vectors.
(3) Linearly independent and spans V. (This is the definition of "basis.")

Linear Transformations

The two basic vector operations are addition and scaling. From this perspective, the nicest functions are those which "preserve" these operations:

Def: A **linear transformation** is a function $T\colon \mathbb{R}^n \to \mathbb{R}^m$ which satisfies:
(1) $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$
(2) $T(c\mathbf{x}) = cT(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Fact: If $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then $T(\mathbf{0}) = \mathbf{0}$.

We've already met examples of linear transformations. Namely: if A is any $m \times n$ matrix, then the function $T\colon \mathbb{R}^n \to \mathbb{R}^m$ which is matrix-vector multiplication

$$T(\mathbf{x}) = A\mathbf{x}$$

is a linear transformation.

(Wait: I thought matrices *were* functions? Technically, no. Matrices are literally just arrays of numbers. However, matrices *define* functions by matrix-vector multiplication, and such functions are always linear transformations.)

Question: Are these all the linear transformations there are? That is, does every linear transformation come from matrix-vector multiplication? Yes:

Prop 13.2: Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then the function T is just matrix-vector multiplication: $T(\mathbf{x}) = A\mathbf{x}$ for some matrix A.

In fact, the $m \times n$ matrix A is

$$A = \begin{bmatrix} | & & | \\ T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \\ | & & | \end{bmatrix}.$$

Terminology: For linear transformations $T\colon \mathbb{R}^n \to \mathbb{R}^m$, we use the word "kernel" to mean "nullspace." We also say "image of T" to mean "range of T." So, for a linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$:

$$\ker(T) = \{\mathbf{x} \in \mathbb{R}^n \mid T(\mathbf{x}) = \mathbf{0}\} = T^{-1}(\{\mathbf{0}\})$$
$$\operatorname{im}(T) = \{T(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n\} = T(\mathbb{R}^n).$$