Span

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

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 $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is the set of all linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_k$.

i.e.: $\mathrm{span}(\mathbf{v}_1, \dots, \mathbf{v}_k) = \{c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \mid c_1, \dots, c_k \in \mathbb{R}\}$.

Linear Independence

The definition in the textbook is:

Def: A set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is **linearly independent** if none of the vectors is a linear combination of the others.

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

Caveat: This definition only applies to a set of two or more vectors.

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

Def: A set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 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 $c_1 = \cdots = c_k = 0$.

$\therefore$ A set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 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 **<u>not all</u>** the scalars $c_1, \dots, c_k$ are zero.

Point: Linear independence of $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ 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 often 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.
(ii) The origin.
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.
(ii) A line.
(iii) The origin.
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."
(ii) A plane.
(iii) A line.
(iv) The origin.
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 $\mathbb{R}^2$?

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** of a vector $\mathbf{v} \in \mathbb{R}^n$ is:

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

Note that $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$. (Q: Why is this a reasonable definition of length?)

Cauchy-Schwarz Inequality: For any non-zero vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$:

$$|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\|\|\mathbf{w}\|.$$

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

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

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|.$$

Equality holds if and only if $\mathbf{w} = c\mathbf{v}$ for some positive scalar c.

Dot Products - Geometry

Prop: Let $\mathbf{w}, \mathbf{w} \in \mathbb{R}^n$ 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

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

$$\mathbf{v} \times \mathbf{w} = \begin{bmatrix} v_2w_3 - v_3w_2 \\ v_3w_1 - v_1w_3 \\ v_1w_2 - v_2w_1 \end{bmatrix}.$$

Note: Dot products make sense in $\mathbb{R}^n$ for any dimension n. But cross products only really work 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 $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$. The area of the parallelogram formed by $\mathbf{v}$ and $\mathbf{w}$ is $\|\mathbf{v} \times \mathbf{w}\|$.

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.

Linear Systems as Matrix-Vector Products

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

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

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.

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 $A\mathbf{x} = \mathbf{0}$.

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.

$\star$ 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 necessarily 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 <u>domain</u> of A.
- $C(A)$ is a subspace of the <u>codomain</u> 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 $N(A) \neq \{\mathbf{0}\}$

Subspace & 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: $\mathbf{v} = x_1\mathbf{w}_1 + \cdots + x_k\mathbf{w}_k$.
∘ 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**.)