02-680 Module 9

Essentials of Mathematics and Statistics

Aidan Jan

October 21, 2025

Linear Independence

First, we need to first define linear dependence.

- A set of vectors is **linearly dependent** if at least one of the vector lies in the span of the others, meaning it adds no new dimension
- In other words, by summing or scalar multiplying the other vectors, it is possible to produce the given vector.

Linear Combinations of Vectors

for a list of vectors $v_1, \ldots, v_m$, a linear combination is written in the form:

- $a_1 v_1 + \cdots + a_m v_m$
- where $a_1, \ldots, a_m$ are scalars.

Building a Vector Space

Vectors are elements of a **vector space**, and we can:

- Add them
- Scale them with scalars
- Use them to build other vectors through combinations

Vector spaces follow the **closure property**:

- Addition and scalar multiplication are closed
- The result always stays within the *same* vector space.

Why Linear Independence Matters

To build the whole space efficiently from a vector set, we want:

- No redundancy in our set.
- No vector in the set can be written as a combination of the others.
- Each vector adds a new direction.

This leads to the concept of **Linear Independence** - a foundational idea before we define a basis.

Definition of Linear Independence

Let $S = \{v_1, v_2, \ldots, v_n\} \subseteq V$ be a set of vectors in a vector space V. We say that S is linearly independent if:

$$a_1v_1 + a_2v_2 + \cdots a_nv_n = 0 \Rightarrow a_1 = a_2 = \cdots = a_n = 0$$

A linear combination of vectors in S equals the zero vector only **when all the coefficients are zero**.

- This is because linear independence is defined this way to ensure that each vector contributes a new, non-redundant direction to the span. It guarantees a minimal set of generators, uniqueness in representation, and full dimensionality of the space spanned by the vectors.
- In other words, the equation says that for any vector $v_i \in S$ there is **no linear combination** of $S \setminus \{v_i\}$ that is equal to v_i. Here, $a_1v_1 + a_2v_2 + \cdots + a_nv_n$ with $\forall i \ : \ a_i \in \mathbb{R}$ is a linear combination of the vectors in S.
- This is usually simplified to:

$$\sum_{i=1}^{n} \alpha_i v_i$$

Note that linear dependence and linear independence are notions that apply to a **collection** of vectors. It does not make sense to say things like "this vector is linearly dependent on these other vectors," or "this matrix is linearly independent."

Example

Is the set

$$\left\{ \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\2\end{pmatrix}, \begin{pmatrix}3\\1\\4\end{pmatrix} \right\}$$

linearly independent? Equivalently, we are asking ifthe homogeneous vector equation

$$x\begin{pmatrix}1\\1\\1\end{pmatrix} + y\begin{pmatrix}1\\-1\\2\end{pmatrix} + z\begin{pmatrix}3\\1\\4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$$

We can solve this by forming a matrix and row reducing.

$$\begin{pmatrix}1 & 1 & 3\\1 & -1 & 1\\1 & 2 & 4\end{pmatrix} \longrightarrow \begin{pmatrix}1 & 0 & 2\\0 & 1 & 1\\0 & 0 & 0\end{pmatrix}$$

This says $x = -2z$ and $y = -z$. So there exist nontrivial solutions: for instance, taking $z = 1$ gives this equation of **linear dependence**:

$$-2\begin{pmatrix}1\\1\\1\end{pmatrix} - \begin{pmatrix}1\\-1\\2\end{pmatrix} + \begin{pmatrix}3\\1\\4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$$

$$-2v_1 - v_2 + v_3 = 0 \Rightarrow v_3 = 2v_1 + v_2$$

This shows that the three vectors are linearly dependent.

- We can test whether any set of vectors are independent using this method. If the only solution to $a_1v_1 + a_2v_2 + a_3v_3 = 0$ is the trivial solution, $a_1 = a_2 = a_3 = 0$, then the vectors are linearly independent.
- If a non-zero solution exists (meaning that one vector can be written in terms of the other two), then it is not.

Facts about Linear Independence

1. Two vectors are linearly dependent if and only if they are colinear, i.e., one is a scalar multiple of the other.
2. Any set containing the zero vector is linearly dependent.
3. If a subset of $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent, then $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent as well.

Span

A **span** is the set of all possible linear combinations of a list of vectors $v_1, \ldots, v_m$:

$$\text{span}(v_1, \ldots, v_m) = \{a_1 v_1 + \ldots a_m v_m \;:\; a_1, \ldots, a_m \in \mathbb{R}\}$$

The span is the smallest subspace that contains all the elements of the list.

- Drawing a picture of $\text{span}\{v_1, v_2, \ldots, v_k\}$ is the same as drawing a picture of all linear combinations of $v_1, v_2, \ldots, v_k$.

Finding Span

For a set of vectors, we say that **span** is another set of vectors that consists of all linear combinations. So in the case above:

$$\langle 8, 7\rangle \in \text{span}(\{\langle 1, 2\rangle, \langle 2, 1\rangle\}) = \{a_1\langle 1, 2\rangle + a_2\langle 2, 1\rangle | a_1, a_2 \in \mathbb{R}\}$$

We want to know **whether** $\langle 8, 7\rangle$ can be written as such a linear combination.

To do this, we assume:

$$\langle 8, 7\rangle = a_1\langle 1, 2\rangle + a_2\langle 2, 1\rangle$$

First, we distribute the variables:

$$\langle a_1 + 2a_2, 2a_1 + a_2\rangle = \langle 8, 8\rangle$$

Now, we can convert it into a linear system:

$$\begin{aligned} a_1 + 2a_2 &= 8 \\ 2a_1 + a_2 &= 7 \end{aligned}$$

We can solve this using substitution, elimination, or converting it to a matrix and reducing. In this case, $a_1 = 2$ and $a_2 = 3$.

- Since the solution exists, yes, $\langle 8, 7\rangle$ is in the span of $\{\langle 1, 2\rangle, \langle 2, 1\rangle\}$
- Formally, for a set of vectors D with cardinality $|D| = n$,

$$\text{span}(D) \;:= \left\{ \sum_{i=1}^{n} \alpha_i D_i \,\middle|\, \forall i \in [n] \;:\; \alpha_i \in \mathbb{R} \right\}$$

- If $D \subseteq V$ for a vector space V, then D is always a valid subspace of V.
- So in the example above since $\{\langle 1, 2\rangle, \langle 2, 1\rangle\} \subseteq \mathbb{R}^2$, the vector space defined by the set is also a subspace of $\mathbb{R}^2$.

Linear Independence

By definition, a set of vectors are linearly independent if for the equation

$$a_1 v_1 + \cdots + a_m v_m = 0$$

for m vectors v_m and constants a_m, the only solution of the constants that satisfies the equation is $a_1 = \cdots = a_m = 0$.

Linear Dependence Lemma

Suppose $v_1, \dots, v_m$ is a linearly dependent list of vectors, then there is some $j \in \{1, \dots, m\}$ such that:

- $v_j \in \text{span}(v_1, \dots, v_{j-1})$
- if the j-th term is removed from $v_1, \dots, v_m$, the span of the remaining list equals $\text{span}(v_1, \dots, v_m)$.

The length of every **linearly independent** list of vectors is less than or equal to the length of every spanning list of vectors.

Bases

A **basis** is a list of vectors that is both **independent** and **spans the space**.

- In a vector space V, we are particularly interested in sets of vectors A that possess the property that any vector $v \in V$ can be obtained by a linear combination of vectors in A.
- These vectors are special vectors, and in the following, we will characterize them.

Simply stated, the **basis** of a vector space is the **smallest set** of linearly independent vectors that span the space.

- More formally, we say a set $B = \{b_1, b_2, \dots, b_n\} \subseteq V$ is a **basis** of vector space V if and only if:
  - $V = \text{span}(B)$
  - $\nexists \text{bi} \in B \,:\, V = \text{span}(B \setminus \{\text{bi}\})$, that is we cannot remove any element and have it still span all of V.

This means for every element in V, there is a unique linear combination of the elements in B that is equal:

$$\forall v \in V \,:\, \forall \alpha \,:\, \sum_{i=1}^{n} \alpha_i b_i = v$$

We call the vector a above the coordinate representation of vector v with respect to the basis.

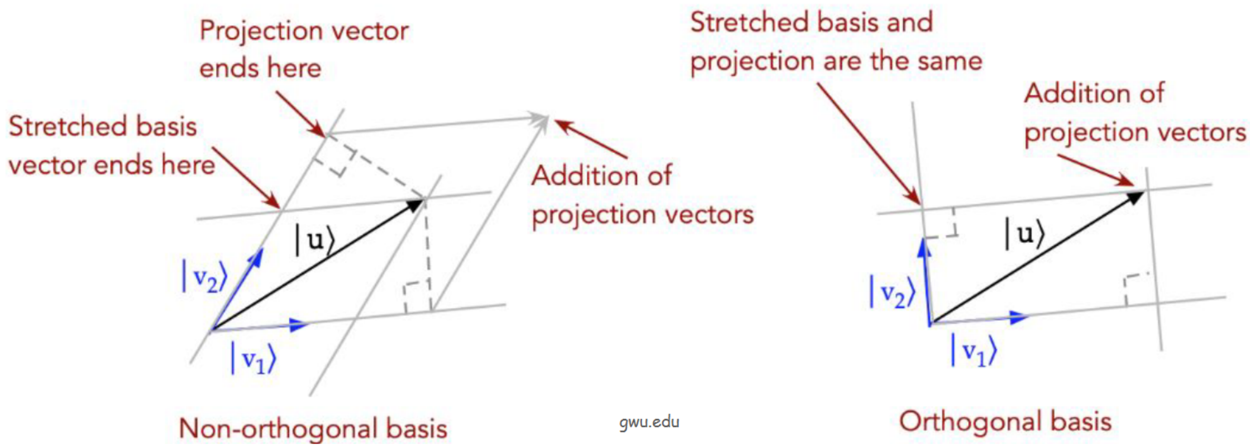
- We say the **size of basis set** is the **dimension of the vector space.**

Orthogonal Vectors

Recall that two vectors are orthogonal if the dot product is 0. Two vectors $x, y \in \mathbb{R}^n$ are **orthogonal** or **perpendicular** if $x \cdot y = 0$.

- Notation: $x \perp y$ means $x \cdot y = 0$.
- Since $0 \cdot x = 0$ for any vector x, the zero vector is orthogonal to every vector in $\mathbb{R}^n$

Orthonormal Basis



We can also say that a vector v is **normal** if $\|v\|_2 = 1$ (graphically it means it lies on the unit (hyper)sphere).

A basis B is considered an **orthonormal basis** if

1. $\forall b_i, b_j \in B \,:\, b_i b_j = 0$ (meaning all bases are orthogonal)
2. $\forall b \in B \,:\, \|b\|_2 = 1$ (all vectors are normal).

The nice thing about an orthonormal basis is that when you convert to coordinate representations, you preserve length and angles between vectors. The most common orthonormal basis for $\mathbb{R}^n$ is the standard basis:

$$\left\{ \langle x_1, x_2, \cdots, x_n \rangle \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^{n} x_i = 1, x_i \in \{0,1\} \forall i \right\}$$

This is, for $\mathbb{R}^3$, the standard basis is $\{\langle 1,0,0,\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$

Multiple Orthonormal Bases

Notice though that the $\mathbb{R}^2$ standard basis is $\{\langle 1,0,\rangle, \langle 0,1\rangle\}$, we can use a secondary basis:

$$B' = \left\{ \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle, \left\langle \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right\rangle \right\}$$

is also an orthonormal basis for $\mathbb{R}^2$.

This means that for any two vectors the coordinate representation of the vectors in both basis will maintain the norms and dot-products.

- Intuitively, in this case, we are applying a rotation of the standard basis about the origin to obtain any other orthonormal basis. This is because orthonormal vectors must have length 1 and be perpendicular to each other.
- We are basically rotating our coordinate plane since the basis vectors define it.
- As a result, the most that would happen would be all the vectors on the plane are rotated about the origin. Rotations are linear transformations, which preserve lengths and angles.

Matrices as Transformation

Note that if we construct a matrix T for which the columns are our basis vectors, we can use matrix multiplication to find the coordinate representation: $vT = a$. So in the example above:

$$T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

and it follows that

$$\begin{bmatrix} 6 \\ 2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 4\sqrt{2} \\ 2\sqrt{2} \end{bmatrix}$$

(We are converting the $\langle 6, 2 \rangle$ in the original coordinate system to the same vector but in terms of the new basis vectors.)

This will be more useful when talking about eigenelements and decomposition later.

Rank

Looking back to matrices quickly, remember we can think of a matrix $A \in \mathbb{R}^{n \times m}$ as a set of vectors (either n length m vectors of rows, or m length n vectors of columns). The **rank** is the number of linearly independent vectors. While you may see both row rank and column rank independently, it turns out these are actually always the same.

For a matrix $A \in \mathbb{R}^{m \times n}$

- The rank of a matrix A is denoted as $\text{rk}(A)$, and is defined as the dimension of its column space. A critical property is that the dimension of the column space is always equal to the dimension of the row space.
- Therefore, the rank is also the dimension of the row space.
- $\dim(\text{Column Space}) = \dim(\text{Row Space}) = \text{rk}(A)$
- This formally explains why "row rank equals column rank" - the number of linearly independent columns is equal to the number of linearly independent rows.

We say a matrix is **full rank** if $\text{rank}(A) = \min(m, n)$. Notice it is always the case that $\text{rank}(A) \leq \min(m, n)$.

- Because of the property above $\text{rk}(A) = \text{rk}(AT)$, then for some second matrix $B \in \mathbb{R}^{m \times p}$: $\text{rank}(AB)$ $\leq \min(\text{rank}(A), \text{rank}(B))$.
- And for a third matrix $B \in \mathbb{R}^{n \times m}$: $\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B)$. Note that in the reduced form we talked about previously, the number of non-zero rows is the rank (and the rows are going to be linearly independent). This also means that any nonsingular (invertible) matrix is full rank (i.e., rank is equal to n).
- The most common method for calculating rank of a matrix is Gaussian Elimination. Use row operations to transform the matrix into its row-echelon form. The rank of the matrix is equal to the number of non-zero rows in its row-echelon form.