Topic 9: Linear Independence

02-680: Essentials of Mathematics and Statistics

October 1, 2024

For a set $S = \{v_1, v_2, ..., v_n\}$ of vectors from vector space V, we say the set is ***linearly independent*** iff

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = \mathbf{0} \quad \implies \quad \alpha_1 = \alpha_2 = ... = \alpha_n = \mathbf{0}.$$

That is to say, for any vector $v_i \in S$ there is no ***linear combination*** of $S \setminus \{v_i\}$ that is equal to v_i. Here $\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$ with $\forall i : \alpha_i \in \mathbb{R}$ is a linear combination of the vectors in S. Usually simplified to:

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

Assume the opposite, a vector α' such that $v_i = \sum_{j \in [n] \setminus \{i\}} \alpha'_j v_j$. We could set $\alpha'_i = -1$ and $\sum_{j \in [n]} \alpha'_j v_j = \mathbf{0}$, but the α''s are non-zero. Which would violate the definition above.

Example. Consider the following set of vectors in $\mathbb{R}^2$:

$$\left\{ \langle 1, 2 \rangle, \langle 2, 1 \rangle, \langle 8, 7 \rangle \right\}.$$

Is this set linearly independent?

To find out we can test if the following has only one solution:

$$\alpha_1 \langle 1, 2 \rangle + \alpha_2 \langle 2, 1 \rangle + \alpha_3 \langle 8, 7 \rangle = \langle 0, 0 \rangle$$

or more familiarly:

$$\begin{array}{rcrcrcl} \alpha_1 & + & 2\alpha_2 & + & 8\alpha_3 & = & 0 \\ 2\alpha_1 & + & \alpha_2 & + & 7\alpha_3 & = & 0 \end{array}$$

Which we know how to reduce:

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

Meaning that for any arbitrary α_3, setting $\alpha_1 = -2\alpha_3$ and $\alpha_2 = -3\alpha_3$ the linear combination is $\mathbf{0}$. Since α_3 can be non-zero the vectors are not linearly independent. (For instance $2\langle 1, 2 \rangle + 3\langle 2, 1 \rangle = \langle 8, 7 \rangle$.)

1 Span

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

$$\langle 8,7\rangle \in span\left(\left\{\langle 1,2\rangle, \langle 2,1\rangle\right\}\right) = \left\{\alpha_1 \langle 1,2\rangle + \alpha_2 \langle 2,1\rangle \mid \alpha_1, \alpha_2 \in \mathbb{R}\right\}$$

Formally for a set of vectors D,

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

The span of any set is a vector space.

If D is a subset of elements from a vector space V, then D is a subspace of V.

So in the example above since $\left\{\langle 1,2\rangle, \langle 2,1\rangle\right\} \subseteq \mathbb{R}^2$, the vector space defined by the set is also a subpace of $\mathbb{R}^2$.

2 Basis

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 $B = \{b_1, b_2, ..., b_n\} \subseteq V$ is a ***basis*** of vector space V iff:

- $V = span(B)$
- $\nexists b_i \in B : V = span(B \setminus \{b_i\})$, that is we cannot remove any element and have it still span all of V.

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

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

We call the vector α above the ***coordinate representation*** of vector v with respect to the basis. For those who have worked with PCA before, you can think of the basis as your PCs and the coordinate representation as the transformation into PC space.

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

2.1 Orthonormal Basis

Recall that two vectors are orthogonal if the dot product is 0. 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) $\forall b_i, b_j \in B : b_i \dot{b}_j = 0$ (meaning all bases are orthogonal), and

(2) $\forall b \in B : \|b\|_2 = 1$ (all vectors are normal).

The nice think 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:

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

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

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

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

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.

Example. In vector space $\mathbb{R}^2$ let

$$v_1 = \langle 6, 2 \rangle \quad \text{and} \quad v_2 = \langle 5, -3 \rangle.$$

If we choose the orthonormal basis B' from above, we find the coordinate representations:

$$v_1' = \langle 4\sqrt{2}, 2\sqrt{2} \rangle \quad \text{and} \quad v_2' = \langle \sqrt{2}, 4\sqrt{2} \rangle.$$

by solving the following systems of equations:

$$\begin{array}{ccccc} \frac{\alpha_1}{\sqrt{2}} & + & \frac{\alpha_2}{\sqrt{2}} & = & 6 \\ \frac{\alpha_1}{\sqrt{2}} & - & \frac{\alpha_2}{\sqrt{2}} & = & 2 \end{array} \quad \text{and} \quad \begin{array}{ccccc} \frac{\beta_1}{\sqrt{2}} & + & \frac{\beta_2}{\sqrt{2}} & = & 5 \\ \frac{\beta_1}{\sqrt{2}} & - & \frac{\beta_2}{\sqrt{2}} & = & 3 \end{array}$$

We can see that the lengths of the vectors are preserved:

$$\begin{array}{rcl} \|v_1\|_2 & \stackrel{?}{=} & \|v_1'\|_2 \\ \sqrt{6^2 + 2^2} & \stackrel{?}{=} & \sqrt{(4\sqrt{2})^2 + (2\sqrt{2})^2} \\ \sqrt{36+4} & \stackrel{?}{=} & \sqrt{32+8} \\ \sqrt{40} & = & \sqrt{40} \end{array} \quad \text{and} \quad \begin{array}{rcl} \|v_1\|_2 & \stackrel{?}{=} & \|v_1'\|_2 \\ \sqrt{5^2 + (-3)^2} & \stackrel{?}{=} & \sqrt{(1\sqrt{2})^2 + (4\sqrt{2})^2} \\ \sqrt{25+9} & \stackrel{?}{=} & \sqrt{2+32} \\ \sqrt{34} & = & \sqrt{34} \end{array}$$

as is the angle (dot product)

$$\begin{aligned}
v_1 \cdot v_2 &\stackrel{?}{=} v_1' \cdot v_2' \\
\langle 6, 2\rangle \cdot \langle 5, -3\rangle &\stackrel{?}{=} \langle 4\sqrt{2}, 2\sqrt{2}\rangle \cdot \langle \sqrt{2}, 4\sqrt{2}\rangle \\
6 \cdot 5 + 2 \cdot -3 &\stackrel{?}{=} 4\sqrt{2} \cdot \sqrt{2} + 2\sqrt{2} \cdot 4\sqrt{2} \\
30 - 6 &\stackrel{?}{=} 8 + 16
\end{aligned}$$

Matrix 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: $vT = \alpha$.

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

$$\begin{bmatrix} 6 \\ 2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 4\sqrt{2} \\ 2\sqrt{2} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 5 \\ -3 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ 4\sqrt{2} \end{bmatrix}$$

This will be more useful when talking about eigenelements and decomposition in the next topic.

3 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 (proof omitted).

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

Because of the property above $rank(A) = rank(A^T)$.

Then for some second matrix $B \in \mathbb{R}^{m \times p}$: $rank(AB) \leq \min(rank(A), rank(B))$.

And for a third matrix $C \in \mathbb{R}^{n \times m}$: $rank(A + B) \leq rank(A) + rank(B)$.

Note that in the reduced form we talked about in the Topic 7, the number of non-zero rows is the rank (and the rows are going to be linearly independent). This also means that any nonsingular (invertable) matrix is full rank (i.e. rank is equal to n).

Useful References

Isaak and Monougian, "Basic Concepts of Linear Algebra". §3.1-3.4
Wilder, "10-606-f23:Lecture 4" GitHub repository, https://github.com/bwilder0/10606-f23/blob/main/files/notes_vectorspace.pdf
Kolter, "Linear Algebra Review and Reference", https://www.cs.cmu.edu/~zkolter/course/15-884/linalg-review.pdf §3.6