MAT188 WEEK 7 Basis and coordinates

Vocabulary: Basis, dimension, coordinate, change of basis, matrix of a linear map with respect to a given basis, similar matrices

Reading from the textbook: : Sec 3.3, Sec 3.4

Introduce

Last week we defined the concept of a subspace of $\mathbb{R}^n$. Typically subspaces are infinitely large spaces. It is handy to describe them via a set of vectors whose linear combinations create every vector in our subspace, a.k.a via a spanning set. Given a spanning set for a subspace, there might be some redundancy, in the sense that not all the vectors are needed to generate our subspace. Technically such sets with redundancy are linearly dependent. We can use row reduction as a tool to identify which vectors we can remove from a given spanning set without changing the substance it spans. This method allows us to boil a spanning set down to a linearly independent spanning set, a.k.a a basis. This week we explore the fact that for a given subspace there are infinitely many bases (*bases* is the plural of *basis*).

Despite the fact that there are many different bases for a given subspace, all of them have the same number of vectors. This number is an intrinsic property of the subspace itself and does not depend on our choice of basis. In a sense, this number tells us how "large" our subspace is. We call this number the *dimension* of our subspace. The algebraic definition of *dimension* was introduced in the 18th century and was a major advance in the development of mathematics: It allows us to concretely conceive of spaces with more than three dimensions, without having to visualize them. Today dimension is a standard concept in mathematics, physics, statistics, data science, and machine learning.

Let's start by building some intuition on the number of vectors required to span a subspace by looking at a familiar example.

Example Consider the plane $P = \{ \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \mid x+y+z=0\}$. Pause reading and justify why P is a subspace of $\mathbb{R}^3$. Let $\vec{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -2 \\ -2 \\ 4 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \vec{v}_4 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ be vectors in P. Pause and verify that these are indeed vectors in P. Let's algebraically verify that every vector in P can be written as a linear combination of $\vec{v}_1, \cdots, \vec{v}_4$. A typical vector $\vec{w}$ in P is of the form $\vec{w} = \begin{bmatrix} -y-z \\ y \\ z \end{bmatrix}$, where y, z range over $\mathbb{R}$. To check if $\vec{w} \in \text{span}(\vec{v}_1, \cdots, \vec{v}_4)$ we should verify that the matrix-vector equation $[\vec{v}_1\ \vec{v}_2\ \vec{v}_3\ \vec{v}_4]\vec{c} = \begin{bmatrix} -y-z \\ y \\ z \end{bmatrix}$ is consistent for any value of y, z. Creating an augmented matrix and going back to the good old row reduction technique we get

$$\left[\begin{array}{cccc|c} -1 & -2 & 1 & 0 & -y-z \\ -1 & -2 & -1 & 1 & y \\ 2 & 4 & 0 & -1 & z \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 2 & 0 & -1/2 & (z+y)/2 \\ 0 & 0 & 1 & -1/2 & (z+y)/2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

This system is always consistent. Hence $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ is a spanning set for P. However, not all these vectors are needed to span P. *Almost* the same computation shows that $\{\vec{v}_1, \cdots, \vec{v}_4\}$ are linearly dependent.

Exercise Let $A = [\vec{v}_1 \cdots \vec{v}_4]$. Explain what the computation above say about the general solution to $A\vec{x} = \vec{0}$. Use your answer to show that $\{\vec{v}_1, \cdots \vec{v}_4\}$ are linearly dependent.

Finally, verify that $\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1/2 \\ 0 \\ 1/2 \\ 1 \end{bmatrix}$ are particular solutions to $A\vec{x} = \vec{0}$.

Indeed, we can span P with $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$, or even better, with only $\{\vec{v}_1, \vec{v}_3\}$. The reason is that

$$\left[\begin{array}{ccc|c} -1 & -2 & 1 & -y-z \\ -1 & -2 & -1 & y \\ 2 & 4 & 0 & z \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 2 & 0 & (z+y)/2 \\ 0 & 0 & 1 & (z+y)/2 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \text{and} \quad \left[\begin{array}{cc|c} -1 & 1 & -y-z \\ -1 & -1 & y \\ 2 & 0 & z \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & (z+y)/2 \\ 0 & 1 & (z+y)/2 \\ 0 & 0 & 0 \end{array}\right]$$

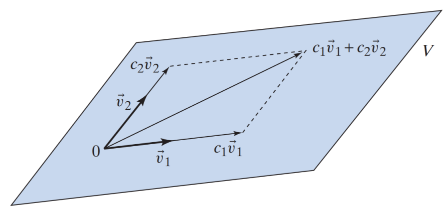
are still consistent. However, here is were our shrinking process stops. We can not go any lower. Removing more vectors will result in an inconsistent system. For instance, removing $\vec{v}_2$ results in the system

$$\left[\begin{array}{c|c} -1 & -y-z \\ -1 & y \\ 2 & z \end{array}\right] \sim \left[\begin{array}{c|c} 1 & (z+y)/2 \\ 0 & (z+y)/2 \\ 0 & 0 \end{array}\right],$$

which is inconsistent when $(z+y)/2 \neq 0$. We showed

$$P = \text{span}(\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4) = \text{span}(\vec{v}_1, \vec{v}_2, \vec{v}_3) = \text{span}(\vec{v}_1, \vec{v}_3) \neq \text{span}(\vec{v}_1).$$

Let's switch to thinking about the same concept geometrically. We started with 4 vectors on P, but to reach to all the points on P we only need 2 linearly independent vectors, and 2 is the least number of vectors that spans P. Any linearly independent set with few vectors fails to span P.



That is a basis, which is a linearly independent spanning set, is a sweet spot between just a spanning set or just a linearly independent set of vectors in a subspace. Spanning sets can have too many vectors, and linearly independent sets can have too few vectors. In general, given a spanning set for a subspace V we can always remove "enough" redundant vectors to obtain a linearly independent spanning set, i.e a basis for V. On the flip side, given any linearly independent set of vectors in V, we can always add "enough" vectors to obtain a linearly independent spanning set for V.

Example Back to our example $P = \{ \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \mid x+y+z=0\}$. Take $\{\vec{v}_2\}$ in P from the previous example. This is a linearly independent set in P that does not span P. Let's choose a vector in P that is not in $(\vec{v}_2)$, say for instance $\vec{v}_4$. Note that there are infinitely many options for this choice. The set $\{\vec{v}_2, \vec{v}_4\}$ is still linearly independent and it also spans P (why?). We found another basis for P!

The following theorem says any set of linearly independent vectors in V is "smaller in size or equal in size to" any spanning set for V.

Theorem (Every spanning set is larger or equal to every linearly independent set in V)**.** *Consider vectors $\vec{v}_1, \cdots, \vec{v}_p$ and $\vec{w}_1, \cdots, \vec{w}_q$ in a subspace V of $\mathbb{R}^n$. If the vectors$\vec{v}_1, \cdots, \vec{v}_p$ are linearly independent, and the vectors $\vec{w}_1, \cdots, \vec{w}_q$ span V, then $q \geq p$.*

Given a nonzero subspace V of $\mathbb{R}^n$, there are infinitely many bases for V. The theorem above guarantees that ALL bases of a subspace V have the same number of vectors. We call this number the dimension of V.

Theorem (Number of vectors in a basis)**.** *All bases of a subspace V of $\mathbb{R}^n$ consist of the same number of vectors.*

Definition (Dimension)

Consider a subspace V of $\mathbb{R}^n$. The number of vectors in **any** basis of V is called the dimension of V and is denoted by $\dim(V)$.

Example Back to $P = \{ \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \mid x+y+z=0\}$ from before. We found two basis for P: $\{\vec{v}_1, \vec{v}_3\}$ and $\{\vec{v}_2, \vec{v}_4\}$. Note that they both have the same number of vectors: two. P is a 2-dimensional subspace of $\mathbb{R}^3$. This matches our intuition. Recall that P is plane in $\mathbb{R}^3$.

Exercise Consider a subspace V of $\mathbb{R}^n$ with $\dim(V) = m$. Use the theorem [Every spanning set is larger or equal to every linearly independent set in V] to justify why the following statements are all true.

a We can find at most m linearly independent vectors in V.
b We need at least m vectors to span V.
c If m vectors in V are linearly independent, then they form a basis of V.
d If m vectors in V span V, then they form a basis of V.

Let's revisit the two important subspaces that are attached to a linear transformation: the kernel and image of a linear transformation. In the next example, we will compute a basis for the kernel and the image of linear map and make an observation about their dimensions.

Example Consider $T_A : \mathbb{R}^4 \to \mathbb{R}^3$, $T_A(\vec{x}) = A\vec{x}$, where $A = \begin{bmatrix} -1 & -2 & 1 & 0 \\ -1 & -2 & -1 & 1 \\ 2 & 4 & 0 & -1 \end{bmatrix} = [\vec{v}_1 \cdots \vec{v}_4]$ We have already seen the computation that goes into finding a basis for the image and kernel of T_A in the first example of this reading. Remember that $\mathrm{im}(T_A) = \mathrm{span}(\vec{v}_1, \cdots \vec{v}_n)$ for which we already computed a basis $\{\vec{v}_1, \vec{v}_3\}$ in our first example. Note that the number of vectors in this basis is exactly the number of pivots in $RREF(A)$, that is the rank of A. Now, let's consider the $\ker(T_A)$. Remember that $\ker(T_A)$ is the general solution to $A\vec{x} = \vec{0}$. Solving the system, we see $\ker T_A = (\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 0 \\ 1/2 \\ 1 \end{bmatrix})$. Note that these two vectors are linearly independent. **Indeed, it is always true that the vectors we get when we write the vector parametric form of a solution to a homogeneous system are linearly independent** (why? It is not obvious! It takes some thinking about what they look like in general!). Hence, $\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 0 \\ 1/2 \\ 1 \end{bmatrix}\}$ is a basis for $\ker(T_A)$. Note that the number of vectors in a basis for $\ker T_A$ is precisely the number of the non-pivot columns of $\mathrm{RREF}(A)$.

To summarize, $\dim \ker T$ is the number of non-pivot columns in $\mathrm{RREF}(A)$, $\dim \mathrm{im}\, T$ is the number of pivot columns of in $\mathrm{RREF}(A)$, and the total number of columns in A is the dimension of the domain of T. This observation is stated in the rank-nullity theorem below.

Now we are ready to give a more sophisticated definition of the rank of a matrix.[1]

[1]Earlier in this course, we defined the rank of a matrix as the number of leading ones in the reduced row echelon form of a matrix. If T is the linear transformation given by multiplication by some matrix A, then these definitions agree!

Definition (Rank & Nullity)

The **rank** of a matrix A is the dimension of $\mathrm{im}(T_A)$, the **nullity** of a matrix A is the dimension of $\ker(T_A)$, where $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is the linear transformation defined by $T_A(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^n$.

Theorem (Rank+Nullity)**.** *Let $T_A : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation defined by $T_A(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^n$. Then*

$$\dim(\ker T_A) + \dim(\operatorname{im} T_A) = n, \quad \textit{equivalently} \quad Rank(A) + Nullity(A) = n.$$