Math 425 Applied & Comput. Lin. Algebra Fall 2024 Lecture

Span, Linear Independence, Bases, Dimension

We begin this lecture with a motivating idea. Let's consider two vectors in $\mathbb{R}^3$ that are not scalar multiples of each other. This means that when we draw these vectors as arrows in space with tails at the origin they are not aligned on the same line. For concreteness, we take

$$\mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \qquad \mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}.$$

Taking scalar multiples of the first vector $c\mathbf{u}$ as c ranges over all real numbers will generate every vector whose terminal point (the "tip" of the corresponding arrow) lies on the line that goes through the origin and is paralllel to $\mathbf{u}$. Note that these include vectors obtained by using a negative scalar c and they lie on the same line but in the opposite direction to $\mathbf{u}$. We will denote this line generated by $\mathbf{u}$ with $L_{\mathbf{u}}$. We do the same thing for the vector $\mathbf{v}$ that generates another line $L_{\mathbf{v}}$ that goes through the origin. Now we "mix" these constructions. We imagine scaling $\mathbf{u}$ with a scalar c and scaling $\mathbf{v}$ with a scalar d (independently from c), and adding up the resulting vectors:

$$c\,\mathbf{u} + d\,\mathbf{v} \quad = \quad c\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + d\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}.$$

The important point to understand here is that the scalars c and d are not chosen once and for all and then fixed. One needs to imagine that both c and d are taking values ranging from $-\infty$ to $+\infty$ (independently). For *each* choice of c and d we compute the above sum to obtain a vector. Then we draw this vector. And we repeat; gazillions of times! What do we get? More precisely, what kind of geometric figure emerges? We get a plane in $\mathbb{R}^3$ *spanned* by the two vectors $\mathbf{u}$ and $\mathbf{v}$. It contains both of the lines $L_{\mathbf{u}}$ and $L_{\mathbf{v}}$. But it also contains all of the lines going through $(0,0,0)^T$ which "sweep" from $L_{\mathbf{u}}$ to $L_{\mathbf{v}}$. Any vector on this plane (i.e. any vector drawn from the origin lying completely in this plane) can be written as a *linear combination* of $\mathbf{u}$ and $\mathbf{v}$ with suitable scalars c and d. It is great that we can capture infinitely many vectors with just two vectors $\mathbf{u}$ and $\mathbf{v}$.
What we have done above can be generalized to more than two vectors in *any* vector space.

Definition 1. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ be vectors in a vector space V. A sum of these vectors of the form

$$\sum_{i=1}^{k} c_i\mathbf{v}_i \; = \; c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

where the coefficients $c_1, c_2, \ldots, c_k$ are scalars, is called a *linear combination* of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$. The set of all such linear combinations as $c_1, c_2, \ldots, c_k$ take all possible values is called the *span* of $\mathbf{v}_1, \ldots, \mathbf{v}_k$, and it is denoted by $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$.

Example 1. If we take $\mathbf{u}$ and $\mathbf{v}$ as above, by choosing $c = 1$ and $d = -2$ we get the linear combination

$$1\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + (-2)\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ -6 \end{pmatrix}.$$

On the other hand, the vector $\mathbf{w} = \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix}$ is *not* a linear combination of $\mathbf{u}$ and $\mathbf{v}$. How do we know? If $\mathbf{w}$ were a linear combination of $\mathbf{u}$ and $\mathbf{v}$, then we would be able to find scalars c and d so that

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

Note that we could find such c and d by solving a linear system

$$\begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix}.$$

You can easily check (by Gaussian elimination) that this system is inconsistent. In other words, $\mathbf{w}$ is not a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Proposition 1. Let $\mathbf{v}_1, \ldots, \mathbf{v_k}, \mathbf{v}$ be vectors in $\mathbb{R}^m$, and let $A = (\mathbf{v}_1\,\mathbf{v}_2 \cdots \mathbf{v}_k)$ be the $m \times k$ matrix whose ith column is $\mathbf{v}_i$. Then $\mathbf{v}$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ (i.e. $\mathbf{v}$ is in the span$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$) if and only if the system $A\mathbf{x} = \mathbf{v}$ has a solution.

Example 2. Let's go back to $\mathbb{R}^3$ and consider sets that are spanned by various vectors.

- What is the span of a single vector $\mathbf{v}$ that is not the zero vector? Well, any linear combination of $\mathbf{v}$ is just a scalar multiple $c\mathbf{v}$. Therefore span$\{\mathbf{v}\} = \{c\mathbf{v} : c \in \mathbb{R}\}$. This gives us a line in $\mathbb{R}^3$ consisting of all vectors parallel to $\mathbf{v}$. Note that if span$\{\mathbf{0}\} = \{\mathbf{0}\}$.
- span$\{\mathbf{v}_1,\ \mathbf{v}_2\}$ is the set of vectors of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. As in the beginning of the lecture, if $\mathbf{v}_1$ is not a scalar multiple of $\mathbf{v}_2$ (or vice versa), then they will span a plane in $\mathbb{R}^3$. Otherwise, they will span a line as in the above case.
- span$\{\mathbf{v}_1,\ \mathbf{v}_2,\ \mathbf{v}_3\}$ is typically equal to $\mathbb{R}^3$, unless all three vectors lie on a plane, or if all three vectors do not lie on a line. These three vectors will span $\mathbb{R}^3$ if and only if $A = (\mathbf{v}_1\,\mathbf{v}_2\,\mathbf{v}_3)$ is nonsingular.

Linear Dependence and Independence. When we talk about the span of a bunch of vectors span$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in a vector space V, it is almost always the case that none of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are redundant. In other words, if we dropped any one of the these vectors the span of the remaining vectors will be a strictly smaller set. However, in some instances, we might realize that, one or more of the vectors *are* redundant, i.e. dropping these redundant vectors will not change the span.

Example 3. Let $\mathbf{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and $\mathbf{w} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. We note that span$\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} = \mathbb{R}^2$. In fact, one way of seeing this is by observing that

$$\begin{pmatrix} a \\ b \end{pmatrix} = (a-b)\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

In other words, $\mathbf{w}$ is redundant: span$\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} =$ span$\{\mathbf{u}, \mathbf{v}\}$. There is another way of seeing this. Since $\mathbf{w} = 2\mathbf{u} + (-1)\mathbf{v}$, the vector $\mathbf{w}$ is already in the span$\{\mathbf{u}, \mathbf{v}\}$, and by including $\mathbf{w}$ we will not add anything new to the span.

In the above example the fact $\mathbf{w} = 2\mathbf{u} + (-1)\mathbf{v}$ could also be expressed by $2\mathbf{u} + (-1)\mathbf{v} + (-1)\mathbf{w} = \mathbf{0}$. Hence, saying that $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$ (hence redundant) is the same as saying that the zero vector is a linear combination of the three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

Definition 2. The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in a vector space V are called *linearly dependent* if there exists scalars $c_1, \ldots, c_k$, *not all zero*, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \;=\; \mathbf{0}.$$

Vectors that are not linearly dependent are called *linearly independent.* In other words, we say $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent, if $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \;=\; \mathbf{0}$ implies that $c_1 = c_2 = \cdots = c_k = 0$.

Note that if one wants to check linear independence, one has to show that the *only* the only way one can produce the zero vector as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is the trivial one: $c_1 = c_2 = \cdots = c_k = 0$.
In the section on linear combinations we answered the question whether a given vector is a linear combination of some other vectors by solving a system of equations. We employ the same strategy to decide whether a set of vectors are linearly dependent/independent.

Theorem 1. Let $\mathbf{v}_1, \ldots, \mathbf{v_k}$ be vectors in $\mathbb{R}^m$, and let $A = (\mathbf{v}_1\, \mathbf{v}_2 \cdots \mathbf{v}_k)$ be the $m \times k$ matrix whose ith column is $\mathbf{v}_i$. Then $\mathbf{v}_1, \ldots, \mathbf{v_k}$ are linearly dependent if and only if there is a non-zero solution $\mathbf{c} \neq \mathbf{0}$ to the homogeneous linear system $A\mathbf{c} = \mathbf{0}$. Equivalently, these vectors are linearly independent if and only if the only solution to the system $A\mathbf{c} = \mathbf{0}$ is $\mathbf{c} = \mathbf{0}$.

Example 4. For instance, let's take $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -4 \\ 6 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}$, and $\mathbf{v}_3 = \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}$. In order to figure out whether these vectors are linearly dependent or not we solve $A\mathbf{c} = \mathbf{0}$ where $A = (\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3)$.

$$\begin{pmatrix} 1 & 3 & 4 \\ -4 & 0 & 2 \\ 6 & 4 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 3 & 4 \\ 0 & 12 & 18 \\ 0 & 0 & 0 \end{pmatrix}.$$

This shows that $\operatorname{rank}(A) = 2$ and the system has infinitely many solutions. Conclusion: the vectors are linearly dependent.

Corollary 1. Any collection of $k > n$ vectors in $\mathbb{R}^n$ is linearly dependent.

Proof. The matrix A one would form is $n \times k$ and $\operatorname{rank}(A) \leq n < k$. So the reduced echelon form of A has free variables. This means $A\mathbf{c} = \mathbf{0}$ has nonzero solutions. $\square$

Bases and Dimension. In many situations we will be interested in finding the most economical way of describing a vector space V as a span of a set of vectors. Clearly, we need enough vectors so that the span is indeed the entire V. But we do not want redundant vectors either. In other words, we should use a set of vectors that is linearly independent.

Definition 3. A finite collection of elements $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of a vector space V is called a *basis* of V if this set spans V and it is linearly independent.

Example 5. If $V = \mathbb{R}^n$ we have the *standard basis* consisting of

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \ldots, \quad \mathbf{e}_{n-1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

where $\mathbf{e}_i$ is the vector with 1 in the ith row and 0s elsewhere. This set spans $\mathbb{R}^n$ because if $\mathbf{x} = (x_1\, x_2\, \ldots\, x_n)^T$ is *any* vector in $\mathbb{R}^n$ then $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$. These n vectors are also linearly independent since $\mathbf{0} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$ implies that $x_1 = x_2 = \cdots = x_n = 0$.

Of course, $\mathbb{R}^n$ has many bases other than the standard basis. The next theorem tells us *precisely* which set of vectors are bases of $\mathbb{R}^n$.

Theorem 2. Every basis of $\mathbb{R}^n$ consists of exactly n vectors. A set of n vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of $\mathbb{R}^n$ if and only if $A = (\mathbf{v}_1\, \ldots\, \mathbf{v}_n)$ is nonsingular.

Proof. A basis of $\mathbb{R}^n$ cannot contain more than n vectors since such a set of vectors is always linearly dependent. If one takes $k < n$ vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ we claim that the span of these vectors cannot be equal to $\mathbb{R}^n$. To show the claim, let's put these vectors into a $n \times k$ matrix $A = (\mathbf{v}_1\, \ldots\, \mathbf{v}_k)$. We apply Gaussian elimination to A, in other words, by multiplying A with an (invertible) matrix E from the left (corresponding to elementary row operations of type 1 and 2) we get $EA = U$ where U is in row echelon form. Since A has more rows than columns U must have a zero row in its nth row. Now, clearly $U\mathbf{x} = \mathbf{e}_n$ is inconsistent, and therefore $A\mathbf{x} = E^{-1}U\mathbf{x} = E^{-1}\mathbf{e}_n = \mathbf{b}$ is inconsistent. This shows the vector $\mathbf{b}$ is not in the span of $\mathbf{v}_1, \ldots, \mathbf{v}_k$, hence if $k < n$ then $\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \neq \mathbb{R}^n$. This proves the first statement: Every basis of $\mathbb{R}^n$ must have n (linearly independent) vectors. For n vectors to be a basis, the system $A\mathbf{x} = \mathbf{b}$ must have a solution for *every* $\mathbf{b} \in \mathbb{R}^n$. But this is possible if and only if A is nonsingular. $\square$

Note that the theorem suggests a very efficient way of checking whether $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ is a basis of $\mathbb{R}^n$: check using Gaussian elimination (or by some other method) whether $A = (\mathbf{v}_1\, \ldots\, \mathbf{v}_n)$ is nonsingular. In other words, check whether $\text{rank}(A) = n$.
As the above theorem shows that every basis of $\mathbb{R}^n$ contains exactly n elements. This is a phenomenon that is not unique to $\mathbb{R}^n$ but it is true for every vector space V.

Theorem 3. If a vector space V has a basis consisting of n vectors, then *every* basis of V has n elements. This number is called the *dimension* of V and is denoted by $\dim V = n$.

Proof. The theorem can be proved using the following fact: if $\mathbf{v_1}, \ldots, \mathbf{v}_n$ span V then every set of $k > n$ vectors $\mathbf{w}_1, \ldots, \mathbf{w}_k$ is linearly dependent. Assuming the fact we prove the theorem. If there are two bases of V with different number of elements, the smaller basis (by definition) spans V, and the fact implies that the vectors in the bigger basis need to be linearly dependent. But, of course, that is absurd. $\square$

Example 6. Using this result we can compute the dimension of some of the vector spaces we have seen. Note that it is enough to have one basis for each example to determine the basis. For instance, $\dim\mathcal{P}^{(n)} = n + 1$ since $\{1, x, x^2, \ldots, x^n\}$ is a basis. The dimension of $\mathcal{M}_{m\times n}$, the vector space of $m \times n$ matrices, is mn, since $\{E_{ij} : i = 1, \ldots, m,\;\; j = 1, \ldots, n\}$ is a basis, where E_{ij} is the $m \times n$ matrix whose (i, j) entry is 1 and all the other entries 0.

We started the lecture with the search of an economical way of describing a vector (sub)space. We finish with a result in the same direction.

Theorem 4. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis for V. Then any $\mathbf{v}$ in V is represented as a linear combination of this basis in a *unique* way. In other words, there exist a *unique* set of scalars $c_1, \ldots, c_n$ such that $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$.

Proof. Suppose there exist another set of scalars $d_1, \ldots, d_n$ such that $\mathbf{v} = d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n$. Then $\mathbf{0} = (c_1 - d_1)\mathbf{v}_1 + \cdots + (c_n - d_n)\mathbf{v}_n$, and since at least one of $c_j \neq d_j$, at least one of $c_j - d_j \neq 0$. But this is a contradiction to the fact that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is linearly independent, as a basis should be. $\square$

The unique coefficients $c_1, c_2, \ldots, c_n$ above are called *coordinates* of $\mathbf{v}$ with respect to the basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. This is the place to point out a subtlety about bases. So far we have been working with bases as *sets* of vectors. But we should think about them as *ordered sets*. After all, if we permute the vectors in a basis, we will get different coordinates (we have to permute the coefficients as well).

Example 7. Let $\mathbf{v} = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$. With respect to the standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of $\mathbb{R}^3$ the coordinates of $\mathbf{v}$ are $(1, -3, 2)^T$. This is what we called coordinates all along. But if we take

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

as our basis of $\mathbb{R}^3$, the coordinates of $\mathbf{v}$ with respect to this basis are $(4, -5, 2)^T$.

As the above example shows one can get the coordinates of $\mathbf{v} \in \mathbb{R}^n$ with respect to a basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ by solving $A\mathbf{c} = \mathbf{v}$ where $A = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$.