1111: Linear Algebra I

Dr. Vladimir Dotsenko (Vlad)

Lecture 15

Solution sets and spans

Yesterday we considered the following example of conversion between two different types of subspaces. Consider the matrix $A = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 3 & -5 & 3 & -1 \end{pmatrix}$, and the corresponding system of equations $Ax = 0$. The reduced row echelon form of this matrix is $\begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & 0 & -1 \end{pmatrix}$, so the free unknowns are x_3 and x_4. Setting $x_3 = s$, $x_4 = t$, we obtain the solution $\begin{pmatrix} -s+2t \\ t \\ s \\ t \end{pmatrix}$, which we can represent as $s\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}$. We conclude that the solution set to the system of equations is the linear span of the vectors $v_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}$.

Let us implement this approach in general. Suppose A is an $m \times n$-matrix. As we know, to describe the solution set for $Ax = 0$ we bring A to its reduced row echelon form, and use free unknowns as parameters. Let $x_{i_1}, \ldots, x_{i_k}$ be free unknowns. For each $j = 1, \ldots, k$, let us define the vector v_j to be the solution obtained by putting the j-th free unknown to be equal to 1, and all others to be equal to zero. Note that the solution that corresponds to arbitrary values $x_{i_1} = t_1, \ldots, x_{i_k} = t_k$ is the linear combination $t_1v_1 + \cdots + t_kv_k$. Therefore the solution set of $Ax = 0$ is the linear span of $v_1, \ldots, v_k$.

The conversion in the other direction, that is computing, for a given set of vectors, a system of equations for which the solution set is the span of the given vectors, also can be done without much trouble. Since we shall not be needing that, we omit the corresponding construction. If you feel comfortable with the course, you can attempt this as an exercise.

Note that in fact the vectors $v_1, \ldots, v_k$ constructed above are linearly independent. Indeed, the linear combination $t_1v_1 + \cdots + t_kv_k$ has t_i in the place of i-th free unknown, so if this combination is equal to zero, then all coefficients must be equal to zero. Therefore, it is sensible to say that these vectors form a basis in the solution set: every vector can be obtained as their linear combination, and they are linearly independent. However, we only considered bases of $\mathbb{R}^n$ so far, and the solution set of a system of linear equations is not generally equal to $\mathbb{R}^m$ for some m. However, the idea of considering bases in this context is perfectly legitimate, and incredibly useful. We shall now create a theoretical set-up for implementing this idea.

Abstract vector spaces

Definition 1. An *(abstract) vector space* (over real numbers) is a set V equipped with the following data:

- a rule assigning to each elements $v_1, v_2 \in V$ an element of V denoted $v_1 + v_2$, and

- a rule assigning to each element $v \in V$ and each real number c an element of V denoted $c \cdot v$ (or sometimes cv),

for which the following properties are satisfied:

1. for all $v_1, v_2, v_3 \in V$ we have $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$,
2. for all $v_1, v_2 \in V$ we have $v_1 + v_2 = v_2 + v_1$,
3. there is a designated *zero element* of V denoted by 0 for which $0 + v = v + 0 = v$ for all v,
4. for each $v \in V$, there exists $w \in V$, denoted $-v$ and called *the opposite of* v, such that $v+(-v) = (-v)+v = 0$,
5. for all $v_1, v_2 \in V$ and all $c \in \mathbb{R}$, we have $c \cdot (v_1 + v_2) = c \cdot v_1 + c \cdot v_2$,
6. for all $c_1, c_2 \in \mathbb{R}$ and all $v \in V$, we have $(c_1 + c_2) \cdot v = c_1 \cdot v + c_2 \cdot v$,
7. for all $c_1, c_2 \in \mathbb{R}$ and all $v \in V$, we have $c_1 \cdot (c_2 \cdot v) = (c_1 c_2) \cdot v$,
8. for all $v \in V$, we have $1 \cdot v = v$.

Examples of vector spaces

To demonstrate that some set V has a structure of a vector space, we therefore must

- exhibit the rules $v_1, v_2 \mapsto v_1 + v_2$ and $c, v \mapsto c \cdot v$ for $v, v_1, v_2 \in V$, $c \in \mathbb{R}$,
- exhibit the zero element $0 \in V$,
- exhibit, for each $v \in V$, its opposite $-v$,

so that the properties 1-8 hold. (As you will find out in homework, the opposite elements actually can be obtained from the rest for free, but we shall not use it for the moment).

Example 1. Of course, the coordinate vector space $\mathbb{R}^n$, as it says on the tin, is an example of a vector space, with the usual operations on vectors, the usual zero vector, and the opposite vector $-v = (-1) \cdot v$. Properties 1-8 hold, we discussed them on some occasions in the past.

Example 2. Slightly more generally, for given m, n, the set of all $m \times n$-matrices is a vector space with respect to addition and multiplication by scalars. This example is not that different from $\mathbb{R}^n$, we just choose to write numbers not in a column of height mn, but in a rectangular array.

Example 3. Every subspace of $\mathbb{R}^n$ is a vector space. It is *almost* obvious from the definition. Indeed, the definition says that a subspace $U \subset \mathbb{R}^n$ is closed under addition and re-scaling, so properties 1, 2, 5, 6, 7, and 8 hold because they hold for vector operations in $\mathbb{R}^n$. The only things which are not automatic is to check that U contains 0 (for property 3), and that the opposite of every vector of U is in U (for property 4). However, both of these are easy: since U is non-empty, it contains some vector u, and hence it must contain $0 \cdot u = 0$. Also, as we mentioned above, in $\mathbb{R}^n$ we have $-u = (-1) \cdot u$, so the negative of every element of U is in U. Hence properties 3 and 4 hold in U because they hold in $\mathbb{R}^n$.

Example 4. The set $C([0, 1])$ of all continuous functions on the segment $[0, 1]$ is a vector space, with obvious operations $(f+g)(x) = f(x)+g(x)$ and $(c \cdot f)(x) = cf(x)$. The only nontrivial thing (which you either know or will soon know from your analysis module) is that these operations turn continuous functions into continuous functions. All the properties 1-8 are obvious.

Example 5. The set of all polynomials

$$a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0$$

in one variable x with real coefficients is a vector space with respect to addition and re-scaling. If we consider polynomials of degree at most n for some given n, this is also a vector space. Polynomials of degree exactly n do not form a vector space, since the sum of two such polynomials may have smaller degree, e.g. $x^n + (1 - x^n) = 1$, where the sum of two polynomials of degree n is of degree 0.