MA1112: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 2

Rank and nullity of a linear map

Last time we defined the notions of the kernel and the image of a linear map $\varphi\colon V \to W$, as well as the nullity and the rank, the dimensions of those subspaces. We shall now discuss how those dimensions are related.

Theorem 1 (Rank-nullity theorem). *Suppose that V and W are finite-dimensional vector spaces. For a linear map $\varphi\colon V \to W$, we have*

$$\mathrm{rk}(\varphi) + \mathrm{null}(\varphi) = \dim(V).$$

Proof. Let us choose a basis $e_1, \ldots, e_n$ of V and $f_1, \ldots, f_m$ of W, and represent φ by the matrix $A = A_{\varphi,\mathbf{e},\mathbf{f}}$.

Since $\mathrm{Ker}(\varphi)$ consists of vectors v such that $\varphi(v) = 0$, we see that $\mathrm{null}(\varphi)$ is the dimension of the solution space of the system $Ax = 0$, where $x = v_{\mathbf{e}}$. This dimension is equal to the number of free variables, that is the number of non-pivotal columns of the reduced row echelon form.

Also, $\mathrm{Im}(\varphi)$ consists of all vectors of the form $\varphi(v)$ with $v \in V$. Since each v can be written as $v = x_1e_1 + \cdots + x_ne_n$, we can write $\varphi(v) = x_1\varphi(e_1) + \cdots + x_n\varphi(e_n)$, so the subspace $\mathrm{Im}(\varphi)$ is spanned by $\varphi(e_1)$, ..., $\varphi(e_n)$. Columns of coordinates of these vectors are precisely the columns of the matrix A, by its definition, so that — in terms of coordinates — $\mathrm{Im}(\varphi)$ becomes identified with the column space of A.

Let us now demonstrate that the column space of A and the column space of its reduced row echelon form R have the same dimension. Indeed, R is obtained from A by row operations, so $R = MA$ where M is an invertible matrix. Recall that the rule for transforming the matrix of a linear map under change of basis is $A_{\varphi,\mathbf{e}',\mathbf{f}'} = M_{\mathbf{f}',\mathbf{f}}A_{\varphi,\mathbf{e},\mathbf{f}}M_{\mathbf{e},\mathbf{e}'}$. This shows that multiplying A by an invertible matrix on the right merely corresponds to a change of basis in W, and therefore R represents the same linear map φ but with respect to a different basis. As a consequence, the column space of R can also be viewed as representing $\mathrm{Im}(\varphi)$ for a certain coordinate system. The image of a linear map is coordinate-independent, so the dimension of the column space of A is equal to the dimension of the column space of R and is equal to $\mathrm{rk}(\varphi)$.

For a reduced row echelon form matrix R, its columns with pivots are some of the standard unit vectors, and all other columns are their linear combinations, so the dimension of the column space is equal to the number of pivotal variables. Thus, $\mathrm{rk}(\varphi)$ is the number of pivotal variables, and $\mathrm{null}(\varphi)$ is the number of free variables, so $\mathrm{rk}(\varphi) + \mathrm{null}(\varphi)$ is the total number of variables, which is manifestly $\dim(V)$. $\square$

The proof we just discussed combines a lot of things we talked about in semester one. Most importantly, it combines the more "practical" part (reduced row echelon forms, free variables, pivotal variables) with the more "theoretical" part (dimension, change of basis). Following the logic of this argument is a good test of fluency in the first semester material.

Let us now offer another proof of the same result; it is more theoretical and abstract, and is a toy model for some more intricate proofs we shall see in the next few weeks.

Another proof of rank-nullity theorem. This other proof of the rank-nullity theorem will be useful in some subsequent classes.

Let us consider $U = \mathrm{Ker}(\varphi)$; it is a subspace of V. Let us choose a basis $e_1, \ldots, e_k$ of U. We can extend this basis to a basis of V by adjoining vectors $f_1, \ldots, f_l$ (if the linear span of e_1,..., e_k coincides with V, we are done, else take for f_1 a vector outside that linear span; then if the linear span of e_1,..., e_k, f_1 coincides with V, we are done, else take for f_2 a vector outside that linear span, etc.). In this set-up, the vectors f_1, ..., f_l are usually said to be a basis of V *relative to the subspace* U. We shall discuss this notion in detail later.

Let us demonstrate that the vectors $\varphi(f_1)$, ..., $\varphi(f_l)$ form a basis of $\mathrm{Im}(\varphi)$. First, take any vector $v \in V$. We can express it via our basis: $v = x_1 e_1 + \cdots + x_k e_k + y_1 f_1 + \cdots + y_l f_l$. Therefore,

$$\varphi(v) = \varphi(x_1 e_1 + \cdots + x_k e_k + y_1 f_1 + y_l f_l) = x_1 \varphi(e_1) + \cdots + x_k \varphi(e_k) + y_1 \varphi(f_1) + \cdots + y_l \varphi(f_l).$$

Since $e_i \in U$, we have $\varphi(e_i) = 0$, and so $\varphi(v) = y_1 \varphi(f_1) + \cdots + y_l \varphi(f_l)$, which means that the vectors $\varphi(f_1)$, ..., $\varphi(f_l)$ form a spanning set. Let us now show that they are linearly independent. Assume the contrary: let

$$c_1 \varphi(f_1) + \cdots + c_l \varphi(f_l) = 0$$

for some scalars c_1, ..., c_l. This immediately implies $\varphi(c_1 f_1 + \cdots + c_l f_l) = 0$, so that $c_1 f_1 + \cdots + c_l f_l \in U$. But $e_1, \ldots, e_k$ form a basis of U, and thus we can write

$$c_1 f_1 + \cdots + c_l f_l = d_1 e_1 + \cdots + d_k e_k,$$

contradicting the basis property of e_i and f_j taken together. Therefore, the vectors $\varphi(f_1)$, ..., $\varphi(f_l)$ form a basis of $\mathrm{Im}(\varphi)$, and we conclude that $\mathrm{rk}(\varphi) + \mathrm{null}(\varphi) = l + k = \dim(V)$.

We conclude with the following remark.

Remark 1. Each of the two proofs we produced actually allows us to conclude a bit more: for every linear map $\varphi\colon V \to W$, there exist bases of V and W relative to which the matrix representing φ is $\begin{pmatrix} I_l & 0_{(n-l)\times l} \\ 0_{l\times(m-l)} & 0_{(n-l)\times(m-l)} \end{pmatrix}$, where $0_{a\times b}$ is the $a \times b$-matrix whose all entries are equal to 0. Here $l = \mathrm{rk}(\varphi)$. In the context of the first proof, we note that we are allowed to change bases in both the domain and the codomain space φ (and hence do both elementary row and elementary column operations) without changing the rank and the nullity; elementary column operations allow us to cancel all non-pivotal columns, and to bring all pivotal columns to the left. In the context of the second proof, one should re-order the basis of V found in that proof so that the vectors of the relative basis go first, and find the basis of W extending the images of the relative basis vectors.