The rank nullity theorem

This handout includes two of the longer proofs of results obtained in Chapter 2. The first of these is the rank-nullity formula, which was introduced last week. The second result allows us to relate the two notions of rank (for matrices and for linear maps) that we have seen.

Theorem 2.8 (The rank nullity formula) *Let U be a finite-dimensional vector space over $\mathbb{F}$ and $f : U \to V$ be a linear map. Then*

$$\dim U = \operatorname{rank}(f) + \operatorname{nullity}(f).$$

Proof: Let $\dim(W) = n$. As $\operatorname{Ker}(f) \leq U$ we have that $\dim \operatorname{Ker}(f) = r \leq n$ for some r by Corollary 1.30. Let $\mathbf{u}_1, \dots, \mathbf{u}_r$ be a basis of $\operatorname{Ker}(f)$. By Corollary 1.33 we can extend this to a basis $\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_n$ of U. We want to show that $\dim \operatorname{Im}(f) = n - r$. For this it will be enough to show that

$$f(\mathbf{u}_{r+1}), \dots, f(\mathbf{u}_n) \quad \text{is a basis of } \operatorname{Im}(f). \tag{1}$$

To prove (1) we first note that all of these elements are in the image of f. We now need to show that they span and are linearly independent.

Span: For each $\mathbf{v} \in \operatorname{Im}(f)$ there exists $\mathbf{u} \in U$ such that $f(\mathbf{u}) = \mathbf{v}$. Now $\mathbf{u} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i$ for some $\lambda_i \in \mathbb{F}$ as the $\mathbf{u}_i$ are a basis of U. Therefore

$$\begin{aligned} \mathbf{v} = f(\mathbf{u}) &= \textstyle\sum_{i=1}^{n} \lambda_i f(\mathbf{u}_i) \\ &= \textstyle\sum_{i=1}^{r} \lambda_i f(\mathbf{u}_i) + \sum_{i=r+1}^{n} \lambda_i f(\mathbf{u}_i) \\ &= \mathbf{0} + \textstyle\sum_{i=r+1}^{n} \lambda_i f(\mathbf{u}_i) \end{aligned}$$

and hence $\mathbf{v}$ is a linear combination of the vectors in (1).

Linear independence: Suppose that our set of vectors in (1) is not independent. Then there exist $\lambda_i \in \mathbb{F}$ not all zero with $\sum_{i=r+1}^{n} \lambda_i f(\mathbf{u}_i) = \mathbf{0}$. As f is a linear map we have that $f(\sum_{i=r+1}^{n} \lambda_i \mathbf{u}_i) = \mathbf{0}$ and hence that $\sum_{i=r+1}^{n} \lambda_i \mathbf{u}_i$ is in the kernel of f.

Therefore there exist $\mu_i \in \mathbb{F}$ such that

$$\sum_{i=r+1}^{n} \lambda_i \mathbf{u}_i = \sum_{i=1}^{r} \mu_i \mathbf{u}_i$$

And hence

$$\sum_{i=r+1}^{n} \lambda_i \mathbf{u}_i - \sum_{i=1}^{r} \mu_i \mathbf{u}_i = \mathbf{0}.$$

But the $\mathbf{u}_i$ form a basis of U and hence μ_i and λ_i are zero for all i. This contradicts our choice of λ_i, and hence our set of vectors must have been independent.

This completes the proof that (1) is a basis of $\operatorname{Im}(f)$, and so we are done. $\square$

Theorem 2.15 *Let A be an $m \times n$ matrix over $\mathbb{F}$. Then*

$$\dim \operatorname{Ker}(A) = n - \operatorname{rr}(A).$$

Proof: Let B be a matrix in echelon form corresponding to A. Then we have that $\operatorname{Ker}(A) = \operatorname{Ker}(B)$ and also that $\operatorname{rr}(A) = \operatorname{rr}(B)$. Therefore it is enough to show that $\dim \operatorname{Ker}(B) = n - \operatorname{rr}(B)$.

Let B have b non-zero rows and let i_t be the column with the first non-zero entry of row t in it. For example, if

$$B = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Then we have $i_1 = 3$, $i_2 = 4$, and $i_3 = 6$. We now want to consider a vector $\mathbf{x} = (x_1, \dots, x_n)$ such that $B\mathbf{x} = \mathbf{0}$. Pick at random the $n - b$ entries x_j where $j \notin \{i_1, \dots, i_b\}$. Then I claim we can find unique values for $x_{i_1}, \dots, x_{i_b}$ such that $\mathbf{x}$ satisfies $B\mathbf{x} = \mathbf{0}$. We do this by working backwards from x_{i_b} up to x_{i_1}, just as we do when using echelon form to solve a set of equations.

For example, let B be as above. Pick $x_1 = x_2 = x_5 = x_7 = 1$ (for example). Now the last row of B gives $x_6 = 0$. The next row up gives $x_4 + 3x_7 = 0$; i.e. $x_4 = -3x_7 = -3$. The final row gives $x_3 + x_4 + x_6 + 2x_7 = 0$; i.e. $x_3 = 3 - 0 - 2 = 1$. So in this case $\mathbf{x} = (1, 1, 1, -3, 1, 0, 1)^T$ satisfies $B\mathbf{x} = \mathbf{0}$.

In particular we can find elements $\mathbf{x}_1, \dots, \mathbf{x}_{n-b}$ such that $B\mathbf{x} = \mathbf{0}$ and our $n - b$ freely chosen elements in each vector are all 0 except for one which is 1, and this element 1 is in a different position each time. My final claim is that these vectors form a basis of $\operatorname{Ker}(B)$, which will complete the proof.

Clearly my chosen vectors are linearly independent. To see that they span let $\mathbf{y}$ be a solution of $B\mathbf{y} = \mathbf{0}$. Let λ_j be the jth entry of $\mathbf{y}$ which is not in position $i_1, i_2, \dots i_b$. Then $\mathbf{x} = \sum_{j=1}^{n-b} \lambda_j \mathbf{x}_j$ has the same choices as $\mathbf{y}$ for the $n - b$ freely chosen elements, and $B\mathbf{x} = \mathbf{0}$, so by uniqueness we have $\mathbf{x} = \mathbf{y}$.

We have shown that the elements $\mathbf{x}_1, \dots, \mathbf{x}_{n-b}$ form a basis for $\operatorname{Ker}(B)$, and so we are done. $\square$