18.06.16: The four fundamental subspaces

Lecturer: Barwick

Monday 14 March 2016

Here it is again:

Theorem (Rank-Nullity Theorem). *If A is an $m \times n$ matrix, then*

$$\dim(\ker(A)) + \dim(\operatorname{im}(A)) = n.$$

We saw a proof of this by reducing to rref or rcef, and then checking it there. There's just one thing that might bug us here. If I think of the linear map

$$T_A : \mathbf{R}^n \longrightarrow \mathbf{R}^m,$$

then we see that $\ker(A)$ is a subspace of the source $\mathbf{R}^n$, but $\operatorname{im}(A)$ is a subspace of the target $\mathbf{R}^m$. So why should these spaces be related?

To answer this question, there's another matrix we can contemplate, the transpose $A^\intercal$. This is an $n \times m$ matrix, and so it corresponds to a linear map in the other direction:

$$T_{A^\intercal} : \mathbf{R}^m \longrightarrow \mathbf{R}^n.$$

This is the map that takes a column vector $\vec{v}$ and builds the column vector $A^\intercal \vec{v}$, but we can perform a trick here. Instead of thinking about transposing A, we can think about transposing the *vectors*.

So $(\mathbf{R}^m)^\vee$ will be the set of all m-dimensional *row vectors*; equivalently, the set of all $1 \times m$ matrices; equivalently again, the set of all transposes

$$\underset{\rightarrow}{v} \coloneqq (\vec{v})^\intercal$$

of vectors $\vec{v} \in \mathbf{R}^m$. We call $(\mathbf{R}^m)^\vee$ the *dual* $\mathbf{R}^m$.

So, e.g., if $\vec{v} = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}$, then $\underset{\rightarrow}{v} = \begin{pmatrix} 1 & 0 & -2 & 1 & 0 \end{pmatrix}$.

The neat thing about row vectors is that they *do stuff* to column vectors. If $\vec{v} \in \mathbf{R}^n$ and $\underset{\rightarrow}{w} \in (\mathbf{R}^n)^\vee$, then $\underset{\rightarrow}{w}\vec{v}$ is a *number*. (Question: How is this related to the dot product?)

If $V \subseteq \mathbf{R}^n$ is a vector subspace, then we define

$$V^\perp \coloneqq \{\underset{\rightarrow}{w} \in (\mathbf{R}^n)^\vee \mid \text{for any } \vec{v} \in V,\ \underset{\rightarrow}{w}\vec{v} = 0\} \subseteq (\mathbf{R}^n)^\vee.$$

Dually, if $W \subseteq (\mathbf{R}^n)^\vee$ is a vector subspace, then we define

$$W^\perp \coloneqq \{\vec{v} \in \mathbf{R}^n \mid \text{for any } \underset{\rightarrow}{w} \in W,\ \underset{\rightarrow}{w}\vec{v} = 0\} \subseteq \mathbf{R}^n.$$

Fact: $\dim(V) = n - \dim(V^\perp)$, and $\dim(W) = n - \dim(W^\perp)$. (Why?)

Now, since

$$(A^\intercal \vec{v})^\intercal = (\vec{v})^\intercal A = \underset{\to}{v} A,$$

we can leave A just as it is, and we can consider the linear map

$$T_A^\vee \colon (\mathbf{R}^m)^\vee \longrightarrow (\mathbf{R}^n)^\vee$$

given by the formula

$$T_A^\vee(\underset{\to}{v}) \coloneqq \underset{\to}{v} A.$$

So when we contemplate the kernel and image of $A^\intercal$, we can think about it via the map $T_A^\vee$.

For example, $\ker(A^\intercal)$ is the set of all vectors $\underset{\to}{v} \in (\mathbf{R}^n)^\vee$ such that $\underset{\to}{v}A = \underset{\to}{0}$. This space is also called the *cokernel* or the *left kernel* of A. I write $\operatorname{coker}(A)$.

We also have the image of $A^\intercal$, which is the set of all row vectors $\underset{\to}{w} \in (\mathbf{R}^n)^\vee$ such that there exists a row vector $\underset{\to}{v} \in (\mathbf{R}^m)^\vee$ for which $\underset{\to}{w} = \underset{\to}{v}A$. This space is also called the *coimage* of A, or, since its the span of the columns of $A^\intercal$, which is the span of the rows of A, it is also called the *row space* of A. I write $\operatorname{coim}(A)$.

In all, we have four vector spaces that are what Strang call the *fundamental subspaces* attached to A:

$$\ker(A), \quad \operatorname{im}(A), \quad \operatorname{coker}(A) \coloneqq \ker(A^{\intercal}), \quad \operatorname{coim}(A) \coloneqq \operatorname{im}(A^{\intercal}).$$

Here's the abstract statement of the Rank-Nullity Theorem:

(1) $\ker(A) = \operatorname{coim}(A)^{\perp}$, so that

$$\dim(\ker(A)) = n - \dim(\operatorname{coim}(A)).$$

(2) $\operatorname{im}(A) = \operatorname{coker}(A)^{\perp}$, so that

$$\dim(\operatorname{im}(A)) = m - \dim(\operatorname{coker}(A)).$$

(3) A provides a bijection $\operatorname{coim}(A) \cong \operatorname{im}(A)$, so that

$$\dim(\operatorname{coim}(A)) = \dim(\operatorname{im}(A)).$$