18.06.12: ‘Kernels and images’

Lecturer: Barwick

Wednesday 2 March 2016

Suppose we can apply some row operations:

$$\left(\begin{array}{c|c} A & B \end{array} \right) \rightsquigarrow \left(\begin{array}{c|c} C & D \end{array} \right).$$

Here, A and C are $m \times n$ matrices, and B and D are $m \times p$ matrices. What this really means is that there's an invertible $m \times m$ matrix M such that $MA = C$ and $MB = D$. (And it turns out that any M can be built this way!)

So you can use row operations to whittle your favorite matrix down, and then solve.

First, let's apply row operations to

$$(A|0) = \left(\begin{array}{ccccc|c} -3 & 6 & -1 & 1 & -7 & 0 \\ 1 & -2 & 2 & 3 & -1 & 0 \\ 2 & -4 & 5 & 8 & -4 & 0 \end{array}\right)?$$

When we get A into *reduced row echelon form* (rref, as the cool kids say) we get

$$(C|0) = \left(\begin{array}{ccccc|c} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Why is this good? First, *we haven't changed the kernel*. A vector $\vec{x} \in \mathbf{R}^5$ is in $\ker(A)$ if and only if $\vec{x}$ is in $\ker(C) = \ker(MA)$. (Why?)

Second, we can use the rref above gives us this system of linear equations

$$\begin{aligned} x_1 &= 2x_2 + x_4 - 3x_5 \\ x_3 &= -2x_4 + 2x_5 \end{aligned}$$

So x_1 and x_3 can each be written in terms of x_2, x_4, x_5, and there's no dependence among x_2, x_4, x_5. So pick variables s, t, u, and let $x_2 = s$, $x_4 = t$, and $x_5 = u$.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = s\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + u\begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

There’s your basis!

Let's find the kernel via row reduction

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

When we get $(A|0)$ into rref, we obtain

$$\left(\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

So if $x_2 = s$ and $x_5 = t$, then

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = s \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

There’s your basis!

Column operations work exactly dual to row operations. (Just think of transposing, doing row operations, and transposing back!) So suppose we can apply some column operations:

$$\left(\frac{A}{B}\right) \rightsquigarrow \left(\frac{C}{D}\right).$$

Here, A and C are $m \times n$ matrices, and B and D are $p \times n$ matrices. What this really means is that there's an invertible $n \times n$ matrix N such that $AN = C$ and $BN = D$.

Why is that a good idea? Well, we’re looking for vectors such that $A\vec{x} = \vec{0}$. So if we take

$$\left(\frac{A}{I}\right)$$

where I is the $n \times n$ identity matrix, then we can start using column operations to get it to some

$$\left(\frac{C}{D}\right)$$

So $AD = C$. So if C has a column of zeroes, then the corresponding column of D will be a vector in the kernel. Furthermore, if you get C into column echelon form, then the nonzero column vectors of D lying under the zero columns of C form a basis of $\ker(A)$. (Properly speaking, to prove this, you need the

Rank-Nullity theorem, which we’ll come to soon.)

Let's do this one:

$$A = \begin{pmatrix} 1 & 0 & 3 & 0 & 2 & -8 \\ 0 & 1 & 5 & 0 & -1 & 4 \\ 0 & 0 & 0 & 1 & 7 & -9 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Let us get the top of $\left(\frac{A}{I}\right)$ into column echelon form.

We obtain

$$\left(\frac{C}{D}\right) = \left(\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 3 & -2 & 8 \\ 0 & 1 & 0 & -5 & 1 & -4 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -7 & 9 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

The last three columns of D are our basis.

Dual to finding a basis of the kernel, we can find a basis of the *image* of a matrix, $\mathrm{im}(A)$. The image of A is the span of the column vectors of A.

If A is an $m \times n$ matrix, then A eats a vector of $\mathbf{R}^n$, and it poops a vector of $\mathbf{R}^m$. The kernel of A is thus a subspace of $\mathbf{R}^n$, and the image of A is a subspace of $\mathbf{R}^m$.

The way we *compute* a basis of the image is not wildly different from the way in which we compute a basis of the kernel, but the operations are dual, and that can get confusing. To clear up our confusion, we'll need some theorems!!

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

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

We are going to spend some quality time with this result.