18.06.08: ‘Inverting matrices’

Lecturer: Barwick

Monday 22 February 2016

Suppose

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

an $n \times n$ (square!!) matrix. We contemplate A via the map $T_A \colon \mathbf{R}^n \longrightarrow \mathbf{R}^n$.

Recall that T_A is defined by the rule

$$T_A(\vec{v}) = A\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^{n} v_i \vec{A}^i = \begin{pmatrix} \sum_{i=1}^{n} a_{1i}v_i \\ \sum_{i=1}^{n} a_{2i}v_i \\ \vdots \\ \sum_{i=1}^{n} a_{ni}v_i \end{pmatrix} = \begin{pmatrix} \vec{A}_1 \cdot \vec{v}_1 \\ \vec{A}_2 \cdot \vec{v}_2 \\ \vdots \\ \vec{A}_n \cdot \vec{v}_n \end{pmatrix}.$$

We're interested in *inverting* T_A; that is, we're interested in finding, for each vector $\vec{w} \in \mathbf{R}^n$, a vector $\vec{x} \in \mathbf{R}^n$ such that $\vec{w} = A\vec{x}$. When we can do this, we write

$$\vec{x} = T_A^{-1}(\vec{w}).$$

In other words, we want to have a way of solving uniquely *any* system of linear equations that looks like

$$
\begin{aligned}
w_1 &= \sum_{i=1}^{n} a_{1i} x_i; \\
w_2 &= \sum_{i=1}^{n} a_{2i} x_i; \\
&\vdots \\
w_n &= \sum_{i=1}^{n} a_{ni} x_i,
\end{aligned}
$$

regardless of what numbers $w_1, \ldots, w_n$ *are!*

Note that we can't always do this …

Here’s a 5×5 matrix:

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 7 & 9 & 11 \\ 4 & 7 & 10 & 13 & 16 \\ 10 & 16 & 22 & 28 & 34 \end{pmatrix}$$

I know straightaway that I won’t be able to solve just any old equation $\vec{w} = A\vec{x}$. (How?)

Here's the bit that is genuinely surprising: suppose A is *invertible* – that is, suppose we can find, for each vector $\vec{w} \in \mathbf{R}^n$, a vector $\vec{x} \in \mathbf{R}^n$ such that $\vec{w} = A\vec{x}$. Then there's a *matrix* A^{-1} such that

$$\vec{x} = A^{-1}\vec{w}.$$

Let us reflect upon this miracle!

What we're saying is that if A is invertible, then *any* system of linear equations

$$\begin{aligned} w_1 &= \sum_{i=1}^{n} a_{1i} x_i; \\ w_2 &= \sum_{i=1}^{n} a_{2i} x_i; \\ &\vdots \\ w_n &= \sum_{i=1}^{n} a_{ni} x_i \end{aligned}$$

admits a *unique solution*, and moreover that there's some matrix

$$A^{-1} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

such that for any i,

$$x_i = \sum_{j=1}^{n} b_{ij} w_j.$$

What are the columns of this matrix?

When $n = 1$, this is familiar. A 1×1 matrix A is just a *number*. When is A invertible? What's the inverse matrix A^{-1}?

When $n = 2$, the action gets a bit more interesting:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

In order for A to be invertible, I need to know that any time I have a couple of numbers u and v, I can solve the system of linear equations

$$\begin{aligned} u &= ax + by; \\ v &= cx + dy \end{aligned}$$

for x and y. Let's work our way through the computation.

So our 2×2 matrix A is invertible if and only if $ad - bc \neq 0$, in which case we have

$$A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}.$$

The number $ad - bc$ is called the *determinant* of A.

The 2×2 case is, sadly, a bit misleading. General formulæ are not always so pleasant, or so useful. The problem is that the complexity of these formulæ increases like $n^3 n!$. So the 3×3 case isn't *so* bad: the matrix

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

is invertible if and only if

$$a(ei - fh) - b(di - fg) + c(dh - eg) \neq 0,$$

and in that case,

$$A^{-1} = \frac{1}{a(ei - fh) - b(di - fg) + c(dh - eg)} \begin{pmatrix} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{pmatrix}.$$

However, the 4×4 case already pushes the utility of general formulæ to their breaking point: consider the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}.$$

Now A is invertible if and only if

$$\begin{aligned}\det(A) = {} & a_{14}a_{23}a_{32}a_{41} - a_{13}a_{24}a_{32}a_{41} - a_{14}a_{22}a_{33}a_{41} + a_{12}a_{24}a_{33}a_{41} \\ + {} & a_{13}a_{22}a_{34}a_{41} - a_{12}a_{23}a_{34}a_{41} - a_{14}a_{23}a_{31}a_{42} + a_{13}a_{24}a_{31}a_{42} \\ + {} & a_{14}a_{21}a_{33}a_{42} - a_{11}a_{24}a_{33}a_{42} - a_{13}a_{21}a_{34}a_{42} + a_{11}a_{23}a_{34}a_{42} \\ + {} & a_{14}a_{22}a_{31}a_{43} - a_{12}a_{24}a_{31}a_{43} - a_{14}a_{21}a_{32}a_{43} + a_{11}a_{24}a_{32}a_{43} \\ + {} & a_{12}a_{21}a_{34}a_{43} - a_{11}a_{22}a_{34}a_{43} - a_{13}a_{22}a_{31}a_{44} + a_{12}a_{23}a_{31}a_{44} \\ + {} & a_{13}a_{21}a_{32}a_{44} - a_{11}a_{23}a_{32}a_{44} - a_{12}a_{21}a_{33}a_{44} + a_{11}a_{22}a_{33}a_{44} \quad \neq 0.\end{aligned}$$

And I just refuse to write down the formula for the inverse. Ain't nobody got time for that.

Still, the idea of inversion has conceptual value. Recall that the following are logically equivalent for an $n \times n$ matrix A:

1. the column vectors $\vec{A}^1, \vec{A}^2, \ldots, \vec{A}^m$ are linearly independent;

2. the row vectors $\vec{A}_1, \vec{A}_2, \ldots, \vec{A}_m$ are linearly independent;

3. the system of linear equations

$$\begin{aligned} 0 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n; \\ 0 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n; \\ &\vdots \\ 0 &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n; \end{aligned}$$

has *exactly one* solution (namely 0).

We have more:

4. for any real numbers $w_1, \ldots, w_n$, the system of linear equations

$$\begin{aligned} w_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n; \\ w_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n; \\ &\vdots \\ w_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n; \end{aligned}$$

   has *exactly one* solution;

5. the matrix A in invertible;

6. there exists an $n \times n$ matrix A^{-1} such that $A^{-1}A = AA^{-1} = I$.

There's even one more:

7. the determinant of A is nonzero.

(We'll talk more about determinants later in the course.)

The point is, we have a notion that's conceptually convenient, but not so very computable. How do we manage?

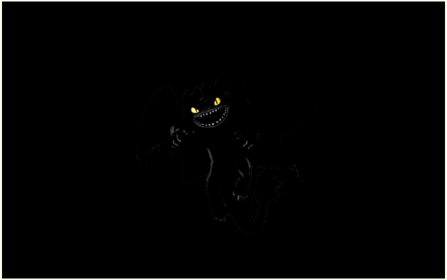

Figure 1: Mother Mathematics. Artist’s rendition.

When Mother Mathematics gives you a problem that’s too hard to solve in general, you solve it in an easy special case, and you try to reduce to that case.

So …which matrices *can* we invert easily?

Diagonal matrices are pretty easy: the matrix $A = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ is invertible if and only if none of the λ_is are zero, in which case

$$A^{-1} = \mathrm{diag}(\lambda_1^{-1}, \dots, \lambda_n^{-1}).$$

More generally, if I've got myself an upper triangular matrix, we should be able to work out whether it's invertible:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

When does that happen?

So, what about something that isn't quite upper triangluar? Is

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

invertible?

Is

$$A = \begin{pmatrix} 2 & 1 & a & b & c & d \\ 1 & 3 & e & f & g & h \\ 0 & 0 & 8 & 5 & i & j \\ 0 & 0 & 8 & 1 & k & l \\ 0 & 0 & 0 & 0 & 5 & -7 \\ 0 & 0 & 0 & 0 & 9 & 1 \end{pmatrix}$$

invertible? (Does it matter what $a, b, c, d, e, f, g, h, i, j, k$, or l are?)

Is

$$A = \begin{pmatrix} 2 & 1 & 1 & a & b & c & d & e \\ 1 & 2 & 1 & f & g & h & i & j \\ 1 & 1 & 2 & k & l & m & n & o \\ 0 & 0 & 0 & 5 & 6 & p & q & r \\ 0 & 0 & 0 & -4 & 9 & s & t & u \\ 0 & 0 & 0 & 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 & 1 \end{pmatrix}$$

invertible?