LECTURE 1

The inverse of a matrix

Last Term, in our brief introduction to matrices, we took a brief look at the idea of an inverse matrix. Specifically, we asked the question: if $\mathbf{r}$ is transformed to $\mathbf{r}'$ by the (square) matrix A, i.e.

$$\mathbf{r}' = \mathsf{A}\mathbf{r},$$

then is is possible to transform $\mathbf{r}'$ back to $\mathbf{r}$ in some way? For example, if A rotates the vector, can we *invert* the rotation and get back to where we started?

We saw that it is only possible to invert the transformation if $\det \mathsf{A} \neq 0$. If $\det \mathsf{A} = 0$, the forward transformation loses some information about the vector. Then there isn't enough information to recreate the starting vector from an inverse transformation.

More precisely, if $\det \mathsf{A} = 0$, the transformation maps the vector into a lower dimensional space.

If $\det \mathsf{A} = 0$, we say that A is a **singular matrix**. Otherwise, we say that it is non-singular (or 'invertible').

If an $n \times n$ matrix is non-singular, an inverse matrix A^{-1} can always be found such that

$$\mathsf{A}\mathsf{A}^{-1} = \mathsf{A}^{-1}\mathsf{A} = \mathsf{I}_\mathsf{n}$$

where I_n is the $n \times n$ identity matrix.

Today, we will explore the matrix inverse in more detail. Our first question will be: how can we calculate it? But before we can do that, I need to point out something important about dealing with matrix equations.

Pre- and post-multiplying by matrices

Suppose I know that two 3×3 matrices are equal:

$$\mathsf{A} = \mathsf{B}$$

Then multiplying them both by a third 3×3 matrix will preserve the equality, *but only if I multiply in the same way on both sides of the equation*. Specifically, I can multiply both sides on the left by C, to get

$$\mathsf{C}\mathsf{A} = \mathsf{C}\mathsf{B}. \tag{1.1}$$

This is called 'pre-multiplying' by C. I could also 'post-multiply' to get

$$\mathsf{A}\mathsf{C} = \mathsf{B}\mathsf{C}. \tag{1.2}$$

What I **cannot** do is multiply one side on the left and the other side on the right, e.g.

$$\mathsf{C}\mathsf{A} = \mathsf{B}\mathsf{C}. \tag{1.3}$$

If we could do this, then from (2.2) and (2.3) we would have $\mathsf{AC} = \mathsf{CA}$, which we know is not generally true in general.

The order of matrix multiplication matters. You have to do the same thing *in exactly the same way* to both sides of an equation.

Calculating the matrix inverse

Some are designed to be very efficient for computers to do: a lot of computer 'number crunching', for example in computational chemistry, involves doing calculations with matrices.

There are several methods available for inverting matrices. We will outline the simplest method; it's not the most efficient method for large matrices, but for 2×2 and 3×3 matrices it works fine.

We'll show the derivation for a 3×3 matrix, but will do it in a way that is hopefully obvious to generalise. The derivation itself is not crucial (so don't worry if you find it hard-going), but it does give us some practice of some of the previous things we have covered.

Start with the matrix

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} ,$$

By expanding along the top row, the determinant of A takes the form

$$\det \mathsf{A} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

where C_{ij} denotes the cofactor, i.e. signed minor, of a_{ij} (see Lecture 1 last Term). For example,

$$C_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \qquad \text{and} \qquad C_{21} = -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} .$$

Now let A' be the matrix obtained from A by replacing row 1 with a copy of row 3:

$$\mathsf{A}' = \begin{pmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} .$$

Its determinant, also expanded along the top, is

$$\det \mathsf{A}' = a_{31}C_{11} + a_{32}C_{12} + a_{33}C_{13}$$

with the same cofactors as A, because we haven't changed rows 2 and 3. But because A' has two identical rows, we also know that $\det \mathsf{A}' = 0$. Hence

$$a_{31}C_{11} + a_{32}C_{12} + a_{33}C_{13} = 0.$$

By generalising the arguments above,

1. Expanding $\det \mathsf{A}$ along row i gives

$$a_{i1}C_{i1} + a_{i2}C_{i2} + a_{i3}C_{i3} = \det \mathsf{A}.$$

2. Replacing row j of det A with row i and then expanding gives

$$a_{i1}C_{j1} + a_{i2}C_{j2} + a_{i3}C_{j3} = 0 \qquad \text{if } i \neq j.$$

Now, let adj A be the transposed matrix of cofactors:

$$\text{adj}\,\mathsf{A} = \begin{pmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{pmatrix}$$

adj A is called the **adjugate** of A. Multiplying adj A by A gives:

It is often called the 'adjoint', for historical reasons. The term 'adjoint' is also used for the conjugate transpose of a matrix, so beware!

$$(\text{adj}\,\mathsf{A})\mathsf{A} = \begin{pmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31} & \cdots & \cdots \\ a_{11}C_{12} + a_{21}C_{22} + a_{31}C_{32} & \cdots & \cdots \\ a_{11}C_{13} + a_{21}C_{23} + a_{31}C_{33} & \cdots & \cdots \end{pmatrix}$$

I've only written out the first column, for clarity, but you can work out for yourself what the other columns will contain. Using the results 1. and 2. derived above, this simplifies to

$$(\text{adj}\,\mathsf{A})\mathsf{A} = \begin{pmatrix} \det\mathsf{A} & 0 & 0 \\ 0 & \det\mathsf{A} & 0 \\ 0 & 0 & \det\mathsf{A} \end{pmatrix}$$

and hence

$$(\text{adj}\,\mathsf{A})\mathsf{A} = (\det\mathsf{A})\mathrm{I}_3.$$

If $\det\mathsf{A} \neq 0$ (i.e. A is non-singular), we can divide through to get

$$\left(\frac{1}{\det\mathsf{A}}\text{adj}\,\mathsf{A}\right)\mathsf{A} = \mathrm{I}_3.$$

From here, we can multiply both sides on the right by A^{-1} to get

$$\left(\frac{1}{\det\mathsf{A}}\text{adj}\,\mathsf{A}\right)\mathsf{A}\mathsf{A}^{-1} = \mathsf{A}^{-1}.$$

Using that $\mathsf{A}\mathsf{A}^{-1} = \mathrm{I}_3$, we get

$$\boxed{\mathsf{A}^{-1} = \frac{1}{\det\mathsf{A}}\text{adj}\,\mathsf{A}. \qquad (1.4)}$$

This is the formula for calculating the inverse of a non-singular matrix.

Example 1. Show that

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

is invertible, and use (2.4) to find the inverse.

The determinant is 3, so the matrix is invertible.

To use the formula in (2.4), we start by calculating the adjugate. It is easiest to do this in a few steps. First, we calculate the matrix of cofactors. For our first example, I am going to write them all out in gory detail, to make sure it is clear what I'm doing.

$$\text{matrix of cofactors } = \begin{pmatrix} (+1)\begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} & (-1)\begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} \\ (-1)\begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} & (+1)\begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} \end{pmatrix}$$

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

Then, the adjugate matrix is the transpose of this, i.e.

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

Finally, the inverse, using (2.4), is just the adjugate divided by the determinant. This gives

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

as the final answer.

It is worth checking that we've done this right. If this is the correct inverse, it should satisfy $AA^{-1} = I$. Let's just verify this:

$$AA^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix} \frac{1}{3}\begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = I_3.$$

Let's do a more interesting example.

Example 2. Calculate the inverse of the rotation matrix

$$\mathsf{R}_z = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

As you may recall, this matrix rotates a vector around the z-axis by θ.

We won't write out the answer to this in so much detail. The top-left element of the cofactor matrix is going to be

$$(+1)\begin{vmatrix} \cos\theta & 0 \\ 0 & 1 \end{vmatrix} = \cos\theta.$$

Make sure you can see why. Filling in the rest of the cofactor matrix, we get

$$\text{matrix of cofactors} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(How bizarre! The matrix of cofactors is the same as the original matrix!) The adjugate is the transpose of this, i.e.

$$\operatorname{adj}\mathsf{R}_z = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Then, we need to divide by the determinant. We showed last Term that the determinant of this matrix is 1, so

$$\mathsf{R}_z^{-1} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

This shouldn't come as too much of a surprise (especially because it was an example in an previous lecture). We can use the oddness/evenness of sin and cos to write our result as

$$\mathsf{R}_z^{-1} = \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) & 0 \\ \sin(-\theta) & \cos(-\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and we see that it is therefore a rotation around the z-angle by $-\theta$. This is clearly the inverse of a rotation around the same axis by $+\theta$.

The examples show that it's possible to calculate the inverse of any non-singular matrix by a formula. The formula is somewhat tedious to apply, and it turns out that there are more efficient ways for large matrices, but this will be enough for us.

Application of matrix inversion

One application of matrix inversion for chemists is in the area of group theory. I will leave that for you to discover next Year. I wanted to show an example which ties back in to what we covered last Term.

Suppose we wanted to solve the simultaneous linear equations.

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= c_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= c_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= c_3.
\end{aligned}$$

This system of equations can be written as the following matrix equation

$$\mathsf{A}\mathbf{x} = \mathbf{c}$$

where A is the matrix of the coeffıcents c_{ij}, and $\mathbf{x}$ and $\mathbf{c}$ are the following column vectors

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \qquad \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

Make sure you are happy why the matrix form is equivalent to the set of equations.

Suppose that A is invertible, i.e. that there is an inverse matrix A^{-1}. We can then multiply both sides of (2) by this inverse.

Notice it wouldn't make sense to multiply on the right. Both sides are (overall) just column vectors, and we have not defined how to multiply a column vector by a 3×3 matrix on the right.

It's clear that we want to multiply both sides of (2) on the *left* by A^{-1}, i.e.

$$\mathsf{A}^{-1}\mathsf{A}\mathbf{x} = \mathsf{A}^{-1}\mathbf{c}.$$

so that the left-hand side simplifies to

$$\mathbf{I}\mathbf{x} = \mathsf{A}^{-1}\mathbf{c}.$$

Then, transforming x on the LHS doesn't change it, so

$$\mathbf{x} = \mathsf{A}^{-1}\mathbf{c}. \tag{1.5}$$

If we remember, $\mathbf{x}$ was the vector of unknowns. The equation above gives us a very simple formula to calculate it!

Example 3. Solve the simultaneous equations

$$\begin{aligned} x_1 - x_2 + x_3 &= 1 \\ x_1 + x_2 + x_3 &= 1 \\ x_1 + 2x_2 + 4x_3 &= 0 \end{aligned}$$

I'll leave it to you to verify that, if

$$\mathsf{A} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}$$

then

$$\mathsf{A}^{-1} = \frac{1}{6}\begin{pmatrix} 2 & 6 & -2 \\ -3 & 3 & 0 \\ 1 & -3 & 2 \end{pmatrix}$$

Therefore, using (2.5), we have

$$\begin{aligned} \mathbf{x} &= \frac{1}{6}\begin{pmatrix} 2 & 6 & -2 \\ -3 & 3 & 0 \\ 1 & -3 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{6}\begin{pmatrix} 8 \\ 0 \\ -2 \end{pmatrix} \end{aligned}$$

Hence $x_1 = \frac{4}{3}$, $x_2 = 0$, and $x_3 = -\frac{1}{3}$.

When you are comfortable with the matrix inversion formula, you'll find that this method usually much quicker than using Cramer's rules for 3×3 and larger systems.