Chapter 8. Determinants

This material is in Chapter 2 of Anton Linear Algebra.

Contents

8.1 Introductory remarks

The determinant of a square matrix A is a number $\det(A)$ associated with the matrix A, and one of its main properties is that A^{-1} exists exactly when $\det(A) \neq 0$.

Unfortunately the calculation of $\det(A)$, and the explanation of what it is, turns out to be tricky. Certainly it is harder than the trace of A. Very vaguely $\det(A)$ is the number you end up dividing by when you compute A^{-1} (and that ties in with the fact that you can't divide by it if it is zero, so that the inverse matrix of A won't make sense if $\det(A) = 0$).

We can make that rather less vague for 2×2 matrices

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

In this case you can calculate A^{-1} as a formula. You can do it either by row-reducing

$$[A \,|\, I_2] = \left[\begin{array}{ccccc} a_{11} & a_{12} & : & 1 & 0 \\ a_{21} & a_{22} & : & 0 & 1 \end{array} \right]$$

and you should end up with

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Having got this formula somehow, you could also check that it works out. (To do that, multiply the supposed A^{-1} by A to see you do indeed get I_2.)

8.1.1 Definition. For a 2×2 matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, the *determinant* of A is defined to be the number

$$\det(A) = a_{11}a_{22} - a_{21}a_{21}$$

In words, this is the product of the two diagonal entries minus the product of the two off-diagonal entries.

It is possible to work out a formula for the inverse of a 3×3 matrix, though it would is quite a bit more messy. There are a number of ways to say what $\det(A)$ is for matrices that are larger than 2×2. I think there is no one way that is really neat. All the approaches either use a lot of ancillary theory, or else have some significant drawback. The way we will choose now is easy enough to explain, but tricky enough to use as a way of showing that determinants do what they are meant to do. In other words proofs are not so nice when we start the way we are going to do, but we won't really notice that problem because we will skip the proofs!

8.2 Cofactor expansion approach to determinants

A quick way to define a determinant is via what is called cofactor expansion along the first row. For 3×3 matrices this means

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

In words the idea is to multiply each entry of the first row times the determinant of the matrix you get by covering over the first row and the column of the entry. Then add these up with alternating signs $+, -, + \dots$.

When we start with a 3×3 matrix A, we end up with $\det(A)$ in terms of 2×2 determinants. And we already know how to evaluate them.

For the 4×4 case, this idea works out like this

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

You should note the alternating signs and that what we now end up with is 4 determinants of size 3×3 to calculate. If we expand each of these via cofactors along the first row, we end up with $12 = 4 \times 3$ determinants of size 2×2 to calculate.

If we use the same approach for 5×5 determinants, we end up with even more work to do. So this method may be simple enough in principle but it is laborious. We will soon explain a more efficient approach for large matrices.

8.3 A formula for determinants

While the above explanation is fine, it is what is called a reduction formula for a determinant. It says how to work out a determinant (of an $n \times n$ matrix A with $n \geq 3$) in terms of smaller determinants. When you keep using the reduction formula enough you get down to 2×2 determinants and we have a nice tidy formula for them. You might like to have a formula for bigger determinants, not a reduction formula.

Such a thing is available and is described in many standard books (but not in the latest edition of the book by Anton). One snag with it is that it requires a bit of theory to explain how it works. I'll outline it below.

In fact the cofactor expansion idea (the reduction formula) works out for 2×2 determinants as well. If you wanted to use it on a 2×2 determinant, it would tell you the answer in terms of 1×1 determinants! It would say

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11} \det[a_{22}] - a_{12} \det[a_{21}]$$

and we might have to worry a bit about what the 1×1 determinants $\det[a_{22}]$ and $\det[a_{21}]$ might mean. Well, a 1×1 matrix has just one entry, one number in it. If you want to be fussy you can insist that a 1×1 matrix is a matrix and an number is a scalar, not the same as a matrix. But actually there is rarely a reason to be so fussy and we can usually safely ignore the difference between a 1×1 matrix and a scalar. The determinant of a 1×1 matrix is just that scalar. Then the reduction formula works out the right determinant for 2×2 matrices!

Technically, we should have said what determinants of 1×1 matrices are and starting with 2×2 meant that we were not being totally complete. So we've filled in that small detail now, though it is not real important.

Back to a formula for general determinants. If you think about it for a while, it is not hard to see that what you get when you expand out $\det(A)$ completely is a sum of products of entries of A times ± 1. In fact what happens is that, if A is an $n \times n$ matrix, then all products of n entries of A show up which satisfy the restriction that the product contains just one entry from each row of A and one from each column. This is kind of apparent from the cofactor expansion approach. At the beginning we get an entry from the first row times a determinant of a matric where the first row is no longer there (and the column of the entry you have is also no longer present in the smaller determinant).

By arguing in this way, you can establish that what you would get if you multiplied out all the reduction formulae for

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

would be a big sum of terms of the form

$$\pm a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

where $j_1, j_2, \ldots, j_n$ are all of the n column numbers in some order.

So $j_1, j_2, \ldots, j_n$ must be all n of the column numbers $1, 2, \ldots, n$, but not necessarily in that order. In fact all possible orders appear. The possible way to reorder $1, 2, \ldots, n$ are called the permutations of these n numbers. It is possible to see fairly easily that the total number of these permutations is a number called 'n factorial'. We write in as $n!$ and it is the product of the numbers $1, 2, \ldots, n$.

$$n! = n(n-1)(n-2)\cdots(3)(2)(1)$$

So this approach gives a formula for the determinant, a formula with $n!$ terms. One snag is that $n!$ gets big pretty fast.

$$3! = 6, \quad 5! = 120, \quad 10! = 3628800$$

So for a 10×10 matrix this formula would have more than 3.6 million terms, a lot. Even for 5×5, you'd have more than 100 terms, each involving a product of 5 terms.

Then there is the problem of which terms get a plus sign and which get a minus. There is a theory about this, and it comes down to something called the 'sign' of a permutation. It would be a digression for us to try and explain what that is in a satisfactory way. So here is a quick explanation. Starting with a permutation

$$j_1, j_2, \ldots, j_n$$

of $1, 2, \ldots, n$, (so that $j_1, j_2, \ldots, j_n$ are all the first n whole numbers written in some order), we are going to write down a matrix called the matrix for this permutation. In each row (and column) the permutation matrix has just one single entry equal to 1, all the others are 0. To be specific, in row number i, there is a 1 in column j_i, and zeros elsewhere. (Another way to say it is that there are entries $= 1$ at the positions (i, j_i) for $i = 1, 2, \ldots, n$, but every other entry is 0.) The sign of the permutation is the same as the determinant of its permutation matrix.

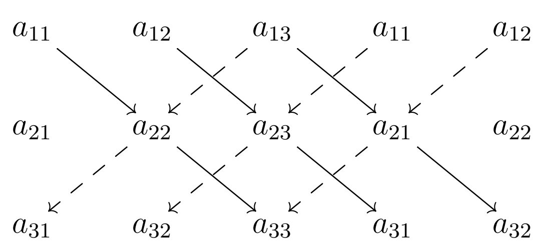
Well, that is a true statement, but it is a bit unsatisfactory. Our long formula for a determinant still has some determinants in it, the ones that give the ± 1 signs.

There is a way to say how the whole formula works out for 3×3 matrices, and it is a fairly satisfactory way of working out 3×3 determinants. The drawback is that it does not extend to bigger determinants in any very similar way.

Starting with a 3×3 matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

write it down with the first two columns repeated



Add the products diagonally to the right and subtract those diagonally to the left as indicated by the arrows

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$
$$-a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Notice that there are the correct number of terms here ($3! = 6$). And each product of 3 has one entry from each row, one entry from each column.

As mentioned above, this method of repeating the first and second columns does **not** work for sizes apart from 3×3, and there is nothing really like this for 4×4 or bigger matrices. The cofactor expansion method does work for any size of (square) matrix.

8.4 Properties of determinants

Here are the key properties of determinants. We'll explain why they are true in the case of 2×2 determinants, and give short shift to the explanations of why these properties still work for $n \times n$ determinants.

(i) $\det(I_2) = 1$

This is pretty easy to see.

(ii) $\det(AB) = \det(A)\det(B)$

Proof. We should show that this is true for any 2×2 matrices A and B, without knowing what the entries are in the matrices. What we do is write out the matrices

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \quad AB = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

Then multiply out

$$\det(AB) = (a_{11}b_{11} + a_{12}b_{21})(a_{21}b_{12} + a_{22}b_{22}) - (a_{11}b_{12} + a_{12}b_{22})(a_{21}b_{11} + a_{22}b_{21})$$

and

$$\det(A)\det(B) = (a_{11}a_{22} - a_{12}a_{21})(b_{11}b_{22} - b_{12}b_{21})$$

to show that the answers are the same.

It is not really hard to do, though maybe not profitable to write the remaining steps out for you. You might like to convince yourself that it does work out as claimed. Doing it for yourself is more likely to convince you that it works. $\square$

(iii) $\det(A^{-1}) = \dfrac{1}{\det(A)}$

Proof. Using the previous result

$$\det(A^{-1})\det(A) = \det(A^{-1}A) = \det(I_2) = 1$$

and so $\det(A^{-1}) = 1/\det(A)$. □

(iv) $\det(A^t) = \det(A)$

This is not at all hard (for the 2×2 case).

(v) The determinants of elementary matrices are as follows

$1'$. E the elementary matrix for the row operation "multiply row 1 by $k \neq 0$"
$E = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$, $\det(E) = k$.

$1''$. E the elementary matrix for the row operation "multiply row 2 by $k \neq 0$"
$E = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$, $\det(E) = k$.

$2'$. E the elementary matrix for the row operation "add k times row 2 to row 1"
$E = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$, $\det(E) = 1$.

$2''$. E the elementary matrix for the row operation "add k times row 1 to row 2"
$E = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$, $\det(E) = 1$.

3. E the elementary matrix for the row operation "swop rows 1 and 2"
$E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\det(E) = -1$.

For general $n \times n$ determinants, all these statements remain true, when suitably phrased.

8.4.1 Lemma. *The determinants of elementary matrices are as follows*

1. *E the elementary matrix for the row operation "multiply a row by $k \neq 0$" has* $\det(E) = k$.
2. *E the elementary matrix for the row operation "add k times one row to a different row" has* $\det(E) = 1$.
3. *E the elementary matrix for the row operation "swop two specified rows" has* $\det(E) = -1$.

This leads us to a way to calculate $n \times n$ determinants. Starting with the matrix A, do row operations on A to row reduce A. At the first row operation we are replacing the matrix A by EA for some elementary matrix A. So the determinant of the matrix we have changes to

$$\det(EA) = \det(E)\det(A)$$

We've seen that $\det(E)$ is easy to figure out. So it is quite easy to keep track of the changes in the determinant as we do each row operation. We'll organise this better next.

If we keep doing row operations, keeping track of how the determinant changes as we go along, we will get to an upper triangular matrix and we can say this:

8.4.2 Lemma. *If A is upper triangular or if A is lower triangular then $\det(A)$ is the product of the diagonal entries.*

Proof. For a lower triangular matrix like

$$\begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{32} & 0 \\ a_{41} & a_{42} & a_{42} & a_{44} \end{pmatrix}$$

it is quite easy to see that expanding along the first row gives a_{11} times some determinant (in fact the determinant of a lower triangular 3×3 matrix in this case) plus zero. An expanding further we just get the product of the diagonals.

For the upper triangular case we could relay on the fact that the transpose will be lower triangular (and has the same determinant and the same diagonal entries). So for example

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix} = \det \begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{12} & a_{22} & 0 & 0 \\ a_{13} & a_{23} & a_{33} & 0 \\ a_{14} & a_{24} & a_{34} & a_{44} \end{pmatrix} = a_{11}a_{22}a_{33}a_{44}$$

This is not a fully correct proof as we have not dealt with arbitrary sized matrices, but it should be enough to explain the idea for the general case. □

8.4.3 Theorem. *(Simplification rules and properties of determinants) Let A be an $n \times n$ matrix throughout.*

(i) Adding a multiple of one row of A to another row results in a matrix with unchanged determinant.

(ii) Factoring a scalar $k \neq 0$ out from a single row of A divides the determinant by k. That is

$$\det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = k \det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{i1}/k & a_{i2}/k & \cdots & a_{in}/k \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

(iii) Swapping two rows of A changes the determinant by a factor -1.

(iv) *An efficient way to work out* $\det(A)$ *is to use Gaussian elimination to row reduce* A *to row-echelon form, keeping track of how the determinant changes after each row operation (see properties (i) – (iii)). Notice that the row-echelon form will be upper triangular with either all 1's on the diagonal or some diagonal entries are 0 (so that the determinant of the row-echelon form is 1 in the case where the diagonal entries are all 1, which happens when* A *is invertible, 0 othewise).*

Proof. The first 3 rules follow from what we said already about determinants of products and determinants of elementary matrices.

For (ii) it may be a bit confusing to see it, but notice that

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = E \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{i1}/k & a_{i2}/k & \cdots & a_{in}/k \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

where E is the diagonal elementary matrix with k at position (i,i) and 1's in the other diagonal places (the elementary matrix for the row operation that multiplies row i across by k).

As for the efficiency claim, working out the determinant directly by expanding will typically take around $n!$ (multiplication) operations, whereas Gaussian elimination will take more like $n^3/3$. We mentioned already that $10!$ is 3.6 million and 10^3 is 'only' a thousand. For n that is any way large $n!$ is prohibitive, while $n^3/3$ is much smaller. $\square$

8.4.4 Examples. 1. Find $\det \begin{bmatrix} 0 & 1 & 5 \\ 1 & 2 & 4 \\ 4 & 5 & 7 \end{bmatrix}$ via row reduction.

$$\begin{aligned} \det \begin{bmatrix} 0 & 1 & 5 \\ 1 & 2 & 4 \\ 4 & 5 & 7 \end{bmatrix} &= -\det \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 5 \\ 4 & 5 & 7 \end{bmatrix} \\ &= -\det \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 5 \\ 0 & -3 & -9 \end{bmatrix} \\ &= -\det \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 5 \\ 0 & 0 & 6 \end{bmatrix} \\ &= -6\det \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix} = -6 \end{aligned}$$

2. Show $\det \begin{bmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{bmatrix} \neq 0$ if x, y and z are all different.

   (Notice that the determinant would be zero if any two of x, y and z were equal. In this case the matrix would have two identical columns, and so determinant zero.)

   To solve this we will first transpose the matrix and then use row operations on the transpose.

$$\begin{aligned}
\det \begin{bmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{bmatrix} &= \det \begin{bmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{bmatrix} \\
&= \det \begin{bmatrix} 1 & x & x^2 \\ 0 & y-x & y^2-x^2 \\ 1 & z & z^2 \end{bmatrix} \\
&= \det \begin{bmatrix} 1 & x & x^2 \\ 0 & y-x & y^2-x^2 \\ 0 & z-x & z^2-x^2 \end{bmatrix} \\
&= \det \begin{bmatrix} 1 & x & x^2 \\ 0 & y-x & (y-x)(y+x) \\ 0 & z-x & (z-x)(z+x) \end{bmatrix} \\
&= (y-x) \det \begin{bmatrix} 1 & x & x^2 \\ 0 & 1 & y+x \\ 0 & z-x & (z-x)(z+x) \end{bmatrix} \\
&= (y-x)(z-x) \det \begin{bmatrix} 1 & x & x^2 \\ 0 & 1 & y+x \\ 0 & 1 & z+x \end{bmatrix} \\
&= (y-x)(z-x) \det \begin{bmatrix} 1 & x & x^2 \\ 0 & 1 & y+x \\ 0 & 0 & z+x-(y+x) \end{bmatrix} \\
&= (y-x)(z-x) \det \begin{bmatrix} 1 & x & x^2 \\ 0 & 1 & y+x \\ 0 & 0 & z-y \end{bmatrix} \\
&= (y-x)(z-x)(z-y)
\end{aligned}$$

   If x, y and z are all different, then all 3 factors in the determinant are different from 0. So their product is not zero.

   By the way, this determinant we just worked out has a special name. It is called a *Vandermonde determinant.*

8.4.5 Theorem. *If A is an $n \times n$ matrix, then the following are two equivalent statements about A:*

(a) A *is invertible*

(b) $\det(A) \neq 0$.

This can then be added as an extra part to a theorem you have seen before. Recall that there are several ways to recognise invertible matrices:

> *Let* A *be an* $n \times n$ *(square) matrix.*
>
> *The following are equivalent statements about* A*, meaning that is any one of them is true, then the other have to be true as well. (And if one is not true, the others must all be not true.)*
>
> *(a)* A *is invertible (has an inverse)*
>
> *(c) the equation* $A\mathbf{x} = \mathbf{0}$ *(where* $\mathbf{x}$ *is an unknown* $n \times 1$ *column matrix,* $\mathbf{0}$ *is the* $n \times 1$ *zero column) has only the solution* $\mathbf{x} = \mathbf{0}$
>
> *(d) the reduced row echelon for* A *is* I_n
>
> *(e)* A *can be written as a product of elementary matrices*
>
> *(f) there is an* $n \times n$ *matrix* B *with* $BA = I_n$
>
> *(g) there is an* $n \times n$ *matrix* B *with* $AB = I_n$

Proof. If you recall the equivalence of (a) and (d) [something discussed last semester and so we take it as known] we can deduce the equivalence of (a) and (b).

In fact we have that (a) implies (b) because of the fact that if A has an inverse then $\det(A^{-1}) = 1/\det(A)$ — so that $\det(A) \neq 0$. So if we know (a) is true about A, then (b) must also be true.

On the other hand suppose we know that (b) holds, that is that $\det(A) \neq 0$. If we row reduce A, the determinant may change by various factors for some of the row operations, but the factors are always nonzero. So when we end up with a row echelon form (triangular) the diagonal entries must be all 1 (otherwise we would have zero for $\det(A)$). If we continue with Gauus-Jordan eleimination to reduced row echelon form we must get the identity matrix. So A satisfies (d), hence satisfies (a). $\square$

8.5 A determinant formula for the inverse matrix

If A is an $n \times n$ matris, then we know more or less that $\det(A)$ is the number that you have to dvide by to calculate A^{-1}. We can now give a formula that makes that very clear.

8.5.1 Definition (Matrix of cofactors)**.** If A is an $n \times n$ matrix with (i, j) entry $a_{i,j}$ we define the *cofactor for the* (i, j) *entry* to be

$$C_{ij} = (-1)^{i+j} \det(M_{i,j})$$

where $M_{i,j}$ is the matrix you get by covering over row i and column j of A. (So the matrices $M_{i,j}$ are $(n-1) \times (n-1)$ matrices.)

The *matrix of cofactors of* A is the $n \times n$ matrix with (i, j) entry $C_{i,j}$.

8.5.2 Remark. In this terminology we can say that the way we defined the determinant was

$$\det(A) = a_{1,1}C_{1,1} + a_{1,2}C_{1,2} + \cdots + a_{1,n}C_{1,n}$$

That is multiply the entries in the first row times their cofactors, and add up the results.

The signs $(-1)^{i+j}$ that go on the cofactors start with plus in the top left corner and alternate their signs along rows or down columns, like this

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & & & \ddots \end{bmatrix}$$

8.5.3 Definition (Adjugate matrix)**.** If A is an $n \times n$ matrix, then the *adjugate matrix of* A is the transpose of the matrix of cofactors. We will denote this matrix by $\operatorname{adj}(A)$.

8.5.4 Theorem (Determinant formula for the inverse)**.** *If A is an $n \times n$ matrix, then*

$$A \operatorname{adj}(A) = \det(A) I_n$$

From this we see that

(i) If $\det(A) \neq 0$*, then*

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$$

(ii) We can expand $\det(A)$ *by cofactor expansion along any row because multiplying row i of A times column i of* $\operatorname{adj}(A)$ *gives*

$$a_{i,1}C_{i,1} + a_{i,2}C_{i,2} + \cdots + a_{i,n}C_{i,n} = \det(A)$$

(iii) We can expand $\det(A)$ *by cofactor expansion along any column.*

Proof. We won't make any attempt to prove that $A \operatorname{adj}(A) = \det(A) I_n$.

The rest are really more or less immediate consequences. If $\det(A) \neq 0$ then we can say

$$A \left(\frac{1}{\det(A)} \operatorname{adj}(A) \right) = I_n$$

That says that $(1/\det(A)) \operatorname{adj}(A)$ is the inverse. So the product works the other way around also

$$\left(\frac{1}{\det(A)} \operatorname{adj}(A) \right) A = I_n$$

and therefore $\operatorname{adj}(A) A = \det(A) I_n$.

The formula $A \operatorname{adj}(A) = \det(A) I_n$ does imply that row i of A times column i of $\operatorname{adj}(A)$ gives $\det(A)$ (the (i,i) entry of $\det(A) I_n$). Since $\operatorname{adj}(A)$ is the transpose of the cofactor matrix, column i of $\operatorname{adj}(A)$ is row i of the cofactor matrix.

If you use $\operatorname{adj}(A) A = \det(A) I_n$ and work out what that says for the (i,i) entry, you end up with the cofactor expansion along column i. □

8.5.5 Example. In the case of a 2×2 matrix

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

then the matrix of cofactors (just a matrix of 1×1 determinants with $\pm$ signs attached is

$$\text{cofactor matrix}(A) = \begin{bmatrix} +d & -c \\ -b & +a \end{bmatrix}$$

and taking the transpose we get the adjugate

$$\operatorname{adj}(A) \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

So we get again the familiar formula for the inverse of a 2×2 matrix

$$A = \frac{1}{\det(A)} \operatorname{adj} A = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

8.5.6 Example. Find $\operatorname{adj}(A)$ when

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$$

Solution: The cofactor matrix for A is

$$\begin{bmatrix} +\det \begin{bmatrix} 4 & 5 \\ 0 & 6 \end{bmatrix} & -\det \begin{bmatrix} 0 & 5 \\ 0 & 6 \end{bmatrix} & +\det \begin{bmatrix} 0 & 4 \\ 0 & 0 \end{bmatrix} \\ -\det \begin{bmatrix} 2 & 3 \\ 0 & 6 \end{bmatrix} & +\det \begin{bmatrix} 1 & 3 \\ 0 & 6 \end{bmatrix} & -\det \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \\ +\det \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} & -\det \begin{bmatrix} 1 & 3 \\ 0 & 5 \end{bmatrix} & +\det \begin{bmatrix} 1 & 2 \\ 0 & 4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 24 & 0 & 0 \\ -12 & 6 & 0 \\ -2 & -5 & 4 \end{bmatrix}$$

Take the transpose to get the adjugate:

$$\operatorname{adj}(A) = \begin{bmatrix} 24 & -12 & -2 \\ 0 & 6 & -5 \\ 0 & 0 & 4 \end{bmatrix}$$

(We could verify that $A \operatorname{adj}(A) = \det(A) I_3$. In fact $\det(A) = 24$ in this case because A is triangular and so the determinant is the product of the diagonal elements (of A). We also have $A^{-1} = (1/\det(A)) \operatorname{adj}(A) = (1/24) \operatorname{adj}(A)$ and A^{-1} is upper trainagular — as you should know would be the case since A is upper triangular.)

8.6 Cramer's rule

This is a topic that is treated in many textbooks, a fairly direct consequence of the formula $A^{-1} = (1/\det(A))\,\mathrm{adj}(A)$ for the inverse.

We recall that if we have a system of n equations in n unknowns, then it can be written as a single matrix equation of the form

$$A\mathbf{x} = \mathbf{b}$$

where A is the matrix of coefficients in the left hand sides of the equations, $\mathbf{x}$ are the n unknowns written as a column matrix ($n \times 1$ column) and $\mathbf{b}$ is the right-hand sides again written as a column.

More precisely stated, the idea is that a system of m linear equations in n unknowns

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}\right.$$

can be written as a single matrix equation

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

which is of the form $A\mathbf{x} = \mathbf{b}$ we mean. However, I am concentrating now on the situation where $m = n$ and A is a square matrix.

Now recall that in the case of the same number of equatioins as unknowns, **if we know** that the (square) matrix A has an inverse, and we know what A^{-1} is, then we can easily solve $A\mathbf{x} = \mathbf{b}$ by multiplying both sides on the lect by A^{-1} to get $A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}$ or

$$\mathbf{x} = A^{-1}\mathbf{b}$$

We can combine that with our formula for the iverse to get

$$\mathbf{x} = \frac{\mathrm{adj}(A)\mathbf{b}}{\det(A)}$$

and further we can express the result in terms of $n \times n$ determinants. To calculate $\mathrm{adj}(A)\mathbf{x}$ we multiply the rows of $\mathrm{adj}(A)$ times the column $\mathbf{b}$. But the rows of A are the columns of the cofactor matrix. So, for instance the first unknown x_1 is

$$x_1 = \frac{C_{1,1}b_1 + C_{2,1}b_2 + \cdots + C_{n,1}b_n}{\det(A)}$$

The top line of this fraction is actually a determinant also (expanded along column 1 via cofactor expansion). We have

$$x_1 = \frac{\det \begin{bmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \\ b_n & a_{n2} & \cdots & a_{nn} \end{bmatrix}}{\det \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}}$$

The denominator is the determinant of the matrix A of the coefficients, while the determinant in the numerator has the first column replaced by the column $\mathbf{b}$ of right hand sides.

What Cramer's rule says is that

$$x_j = \frac{\det(\text{matrix got from } A \text{ by replacing column } j \text{ by } \mathbf{b})}{\det(A)}$$

8.6.1 Example. Write a determinant formula for x_3 if (x_1, x_2, x_3, x_4) solves the system

$$\begin{cases} 5x_1 + 8x_2 - 2x_3 + x_4 = -9 \\ \qquad\quad x_2 + 4x_3 - 7x_4 = 3 \\ \qquad\qquad\quad 11x_3 + 2x_4 = -1 \\ x_1 + x_2 - 8x_3 + 3x_4 = 0 \end{cases}$$

(assuming that the matrix of coefficients is invertible).

Solution: By Cramer's rule we have

$$x_3 = \frac{\det \begin{bmatrix} 5 & 8 & \mathbf{-9} & 1 \\ 0 & 1 & \mathbf{3} & -7 \\ 0 & 0 & \mathbf{-1} & 2 \\ 1 & 1 & \mathbf{0} & 3 \end{bmatrix}}{\det \begin{bmatrix} 5 & 8 & -2 & 1 \\ 0 & 1 & 4 & -7 \\ 0 & 0 & 11 & 2 \\ 1 & 1 & -8 & 3 \end{bmatrix}}$$

(Note that column 3 on the top is the one that is different from the one below. The determinant below is the determinat of the matrix of coefficients. On the top column 3 is replaced by the column of right hand sides.)

April10: Fix typo in Theorem 8.5.4.