Math 1b — TA: Padraic Bartlett

Determinants: An Introduction

Week 5 — *Caltech 2011*

1 Random Questions

1. Consider the following game, called "scream-toes:"

   - To start, place n people so that they're standing in a circle.
   - Have everyone look down at the ground (the "down" phase.)
   - Then, have each player randomly select another player's toes, and look at them (the "toes" phase.)
   - Then, have all players look up at whichever player's toes they were just looking at. If two players are looking at each other, then they both scream (or shout loudly, whichever they prefer.)

   Show that the average number of screams for a given round of scream-toes is about $2/e$, given enough players.

2. Suppose that A is a matrix with integer entries, such that the sum of the entries in each of the rows of A is a multiple of 7. Prove that the determinant is divisible by 7.

2 The Determinant: Basic Definitions, and an Example

Definition. For a $n \times n$ matrix A, let A_{ij} denote the matrix formed from A by deleting the i-th row and j-th column from A.

Then, we can define the **determinant** of A recursively[1] as follows: for 1×1 matrices, we define $\det(A) = a_{11}$, and for larger $n \times n$ matrices A, we define

$$\det(A) = \sum_{i=1}^{n} (-1)^{i-1} a_{1i} \cdot \det(A_{1i}).$$

To give an idea of how we use this recursive definition in practice, consider the following example:

[1]A **recursive** definition is one that defines some object in terms of itself. In this case, we define the determinant of a $n \times n$ matrix in terms of the determinants of smaller $n-1 \times n-1$ matrices. So, to find the determinant of (say) a 3×3 matrix, we use our recursive definition to reduce the problem to finding the determinants of 3 different 2×2 matrices, and then apply the recursive definition on each of those matrices to reduce further to the case of 6 different 1×1 matrices, which we know how to do.

Example. Find the determinant of the following matrix:

$$A = \begin{pmatrix} 3 & 1 & 4 & 1 \\ 5 & 9 & 2 & 6 \\ 5 & 3 & 5 & 8 \\ 9 & 7 & 9 & 3 \end{pmatrix}$$

Solution. By our definition, we know that

$$\begin{aligned} \det(A) &= \sum_{i=1}^{n} (-1)^{i-1} a_{1i} \cdot \det(A_{1i}) \\ &= 3 \cdot \det \begin{pmatrix} 9 & 2 & 6 \\ 3 & 5 & 8 \\ 7 & 9 & 3 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 5 & 2 & 6 \\ 5 & 5 & 8 \\ 9 & 9 & 3 \end{pmatrix} + 4 \cdot \det \begin{pmatrix} 5 & 9 & 6 \\ 5 & 3 & 8 \\ 9 & 7 & 3 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 5 & 9 & 2 \\ 5 & 3 & 5 \\ 9 & 7 & 9 \end{pmatrix} \end{aligned}$$

Use the definition of the determinant again to expand each of these three by three matrices:

$$\begin{aligned} \det \begin{pmatrix} 9 & 2 & 6 \\ 3 & 5 & 8 \\ 7 & 9 & 3 \end{pmatrix} &= 9 \cdot \det \begin{pmatrix} 5 & 8 \\ 9 & 3 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 3 & 8 \\ 7 & 3 \end{pmatrix} + 6 \cdot \det \begin{pmatrix} 3 & 5 \\ 7 & 9 \end{pmatrix} \\ &= 9(5 \cdot 3 - 8 \cdot 9) - 2(3 \cdot 3 - 8 \cdot 7) + 6(3 \cdot 9 - 5 \cdot 7) \\ &= -467. \\ \det \begin{pmatrix} 5 & 2 & 6 \\ 5 & 5 & 8 \\ 9 & 9 & 3 \end{pmatrix} &= 5 \cdot \det \begin{pmatrix} 5 & 8 \\ 9 & 3 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 5 & 8 \\ 9 & 3 \end{pmatrix} + 6 \cdot \det \begin{pmatrix} 5 & 5 \\ 9 & 9 \end{pmatrix} \\ &= 5(5 \cdot 3 - 8 \cdot 9) - 2(5 \cdot 3 - 8 \cdot 9) + 6(5 \cdot 9 - 5 \cdot 9) \\ &= -171. \\ \det \begin{pmatrix} 5 & 9 & 6 \\ 5 & 3 & 8 \\ 9 & 7 & 3 \end{pmatrix} &= 5 \cdot \det \begin{pmatrix} 3 & 8 \\ 7 & 3 \end{pmatrix} - 9 \cdot \det \begin{pmatrix} 5 & 8 \\ 9 & 3 \end{pmatrix} + 6 \cdot \det \begin{pmatrix} 5 & 3 \\ 9 & 7 \end{pmatrix} \\ &= 5(3 \cdot 3 - 8 \cdot 7) - 9(5 \cdot 3 - 8 \cdot 9) + 6(5 \cdot 7 - 3 \cdot 9) \\ &= 326. \\ \det \begin{pmatrix} 5 & 9 & 2 \\ 5 & 3 & 5 \\ 9 & 7 & 9 \end{pmatrix} &= 5 \cdot \det \begin{pmatrix} 3 & 5 \\ 7 & 9 \end{pmatrix} - 9 \cdot \det \begin{pmatrix} 5 & 5 \\ 9 & 9 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 5 & 3 \\ 9 & 7 \end{pmatrix} \\ &= 5(3 \cdot 9 - 5 \cdot 7) - 9(5 \cdot 9 - 5 \cdot 9) + 6(5 \cdot 7 - 3 \cdot 9) \\ &= -24. \end{aligned}$$

Now, take these values and plug them into our original equation:

$$\begin{aligned}
\det(A) &= \sum_{i=1}^{n} (-1)^{i-1} a_{1i} \cdot \det(A_{1i}) \\
&= 3 \cdot \det \begin{pmatrix} 9 & 2 & 6 \\ 3 & 5 & 8 \\ 7 & 9 & 3 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 5 & 2 & 6 \\ 5 & 5 & 8 \\ 9 & 9 & 3 \end{pmatrix} + 4 \cdot \det \begin{pmatrix} 5 & 9 & 6 \\ 5 & 3 & 8 \\ 9 & 7 & 3 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 5 & 9 & 2 \\ 5 & 3 & 5 \\ 9 & 7 & 9 \end{pmatrix} \\
&= 3 \cdot (-467) - 1 \cdot (-171) + 4(326) - 1(-24) \\
&= 98.
\end{aligned}$$

3 The Determinant: Some Exploratory Theorems

The determinant is a rather strange-looking thing. At first glance, it's not remotely clear why we'd ever want to study it; it seems complex and convoluted, and hardly like the kind of thing we would ever intentionally want to work with.

Yet, as it turns out, the determinant is an incredibly useful object! In specific, we have the following theorem:

Theorem 1 *The determinant of a $n \times n$ matrix A is nonzero if and only if A is nonsingular.*

In other words, the determinant – a single number that we can pretty quickly find from a matrix – can instantly tell us if a matrix has an inverse, without bothering with all of the row-reduction nonsense we normally have to do. So this is remarkably useful!

How can we prove such a thing? Well, at the moment, we really can't: we barely understand what the determinant does in general, and thus aren't probably going to have much luck starting on this theorem right now.

So, what we'll do instead is see what we *can* prove about the determinant, and see if we can use these insights to eventually try and prove this result.

To start: if we want to understand what the determinant does to matrices, we should begin by looking at what it does to the simplest matrix we can think of: the identity matrix!

Theorem 2 *If I_n is the $n \times n$ identity matrix, then $\det(I_n) = 1$.*

Proof. So: how do we prove things about the determinant? Well, we defined the determinant recursively, with a definition that told us how to find 1×1 determinants and how to build up this knowledge to find $n \times n$ determinants. A natural idea, then, would be to prove things in a similar way: to demonstrate thing for a base case, and then to build up these results for larger matrices. In other words, we want to use induction! Specifically, for this proof, let's proceed by induction on n.

For the 1×1 matrix $I_1 = (1)$, our claim is trivially true: $\det(I_1) = 1$.

For our inductive step, we assume that our hypothesis holds for n, and seek to prove our claim for $n + 1$.

So: by definition, we know that

$$\begin{aligned}\det(I_{n+1}) &= (-1)^0 \cdot 1 \cdot \det((I_{n+1})_{11}) + (-1)^1 \cdot 0 \cdot \det((I_{n+1})_{12}) + \ldots + (-1)^{n-1} \cdot 0 \cdot \det((I_{n+1})_{1,n+1}) \\ &= \det((I_{n+1})_{11}).\end{aligned}$$

However, removing the first row and first column of the $n+1 \times n+1$ identity matrix just leaves I_n, the $n \times n$ identity matrix: so this is just $\det(I_n)$, which we know to be 1 from our inductive hypothesis.

We now understand the identity matrix. What should we look at now?

Well: one of the most basic things we can do to a matrix are the various **row operations**: i.e. given a matrix, we will often either

- multiply some row by a constant λ,
- swap two rows, or
- add λ times one row to another.

What do these three properties do to the determinant? I.e. if we have a matrix and perform one of these row operations, how does the determinant change?

We explore this in the next three theorems:

Theorem 3 *Suppose that A is a $n \times n$ matrix. If A' is the matrix acquired by multiplying the k-th row of A by some constant λ, then* $\det(A') = \lambda \det(A)$.

Proof. We proceed by induction on n.

For $n = 1$, this is trivial: if $A = (a_{11})$, then A' is necessarily (λa_{11}), and thus

$$\det(A') = \det((\lambda a_{11})) = \lambda a_{11} = \lambda \cdot \det(A).$$

Assume that our result holds for all $n \times n$ matrices. We now seek to prove our claim for all $n+1 \times n+1$ matrices: i.e. for any $n+1 \times n+1$ matrix A, constant λ, and row k, we want to show that the matrix A' formed by multiplying A's k-th row by λ has determinant $\lambda \cdot \det(A)$.

There are two possible cases. Either $k = 1$, in which case A' is of the form

$$\begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \ldots & \lambda a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix},$$

and therefore

$$\det(A') = \sum_{i=1}^{n+1} (-1)^{i-1} \lambda a_{1i} \cdot \det(A'_{1i}) \qquad = \lambda \cdot \sum_{i=1}^{n+1} (-1)^{i-1} a_{1i} \cdot \det(A_{1i}) = \lambda \cdot \det(A),$$

where we justify saying that A_{1i} and A'_{1i} are the same by saying that the only places where A and A' differ is in their first row, and we've deleted that in both A_{1i} and A'_{1i}.

The only case remaining to consider is when $k > 1$; in this situation, we have that A' is of the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda a_{k1} & \lambda a_{k2} & \lambda a_{k3} & \dots & \lambda a_{kn} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix},$$

and thus that

$$\det(A') = \sum_{i=1}^{n+1} (-1)^{i-1} a_{1i} \cdot \det(A'_{1i}).$$

However, we know that (because $k > 2$ each of the A'_{1i} are just the matrix A_{1i} with one row multipled by λ. Because these are $n \times n$ matrices, we can then use our inductive hypothesis to note that $\det(A'_{1i}) = \lambda \det(A_{1i}$, and thus that

$$\begin{aligned} \det(A') &= \sum_{i=1}^{n+1} (-1)^{i-1} a_{1i} \cdot \det(A'_{1i}) \\ &= \sum_{i=1}^{n+1} (-1)^{i-1} a_{1i} \cdot \lambda \det(A_{1i}) \\ &= \lambda \cdot \det(A). \end{aligned}$$

Thus, for any k, we've proven that multiplying the k-th row of A by λ multiples the determinant by λ as well.

Theorem 4 *Suppose that A is a $n \times n$ matrix. If A' is the matrix acquired by swapping the two rows i, j in A, then $\det(A') = -\det(A)$.*

Proof. We again proceed by induction on n.

Our base case is $n = 2$, as this is the first case where we have two distinct rows to swap. However, our claim is still trivial, as when we swap the only two rows in any 2×2 matrix, we have that for any matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ac - bd = -1 \cdot (bd - ac) = -1 \cdot \det \begin{pmatrix} c & d \\ a & b \end{pmatrix}.$$

So: we now assume that our result holds for any $n \times n$ matrix and rows $i < j$, and seek to prove it for $n+1 \times n+1$ matrices. To do this, we first notice the following fact: it suffices to prove that our claim holds in the special case when $j = i+1$. Why is this? Well, notice that if you swap the rows

$$(i \text{ and } i+1), (i+1 \text{ and } i+2), \ldots (j-1 \text{ and } j),$$

you've put the i-th row in the j-th spot and decremented all of the other rows between i and j by one spot. Now, if you swap

$$(j-1 \text{ and } j-2), (j-2 \text{ and } j-3), \ldots (i+1 \text{ and } i),$$

this moves all of the other rows back to their original position *except* for the j-th row, which is now in the i-th spot. So this process switches the i-th and j-th rows, and does so with $2(j-i)-1$-many swaps, which is always an odd number of swaps. So, if our conjecture – that switching rows multiplies the determinant by -1 – holds for just adjacent rows, this process shows that it must hold for all rows, because we can "create" any swap of rows out of an odd number of swaps of adjacent rows, and $(-1)^{\text{odd}}$ is always -1.

Excellent. Let A be a $n+1 \times n+1$ matrix, and let $i, i+1$ be the pair of rows we seek to swap. Then there are two cases:

$i = 1$. In this case, A' is of the form

$$\begin{pmatrix} a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} \end{pmatrix}.$$

Now, examine $\det(A')$:

$$\begin{aligned} \det(A') &= \sum_{l=1}^{n+1} (-1)^{l-1} a_{2l} \cdot \det(A'_{1l}) \\ &= \sum_{l=1}^{n+1} (-1)^{l-1} a_{2l} \cdot \left(\sum_{k<l} (-1)^{k-1} a_{1k} \det(A'_{12,lk}) + \sum_{k>l} (-1)^{k-2} a_{1k} \det(A'_{12,lk}) \right) \end{aligned}$$

Why did we write $\det(A'_{1l})$ as this crazy two-part sum above? Well, if you look at what A'_{1l} is, it's the matrix formed from A' after deleting the first row and the l-th column. So, if you're expanding the determinant of this matrix, what are you doing? You're starting from the first entry in the second row of A' (because you deleted the first row), and then adding entries $\times$ some determinant in alternating sign all the way down A'_{1l}. Until you get to the l-th column, you can just find the sign of these entries by taking $(-1)^{k-1}$, as you just started at $k = 0$ and have alternated: but once you get to the l-th column, you skip it, because it was deleted in A'_{1l}! So, if you're alternating sign, once you get past the l-th column you have to adjust your signs for the column you skipped: i.e. the sign of an entry past the l-th column is $(-1)^{k-2}$, instead of $(-1)^{k-1}$.

For similar reasons, we can expand $\det(A)$ into such a sum as well:

$$\begin{aligned}\det(A) &= \sum_{k=1}^{n+1}(-1)^{k-1}a_{1k}\cdot\det(A_{1k})\\ &= \sum_{k=1}^{n+1}(-1)^{k-1}a_{1k}\cdot\left(\sum_{l<k}(-1)^{l-1}a_{2l}\det(A_{12,lk})+\sum_{l>k}(-1)^{l-2}a_{2j}\det(A_{12,lk})\right).\end{aligned}$$

We claim that the first one of these sums is the negative of the other. To see this, simply pick any pair $l \neq k$. If $l < k$, then the $a_{1k}a_{2l}$ term in $\det(A')$'s determinant is

$$(-1)^{l-1}\cdot a_{2l}\cdot(-1)^{k-2}a_{1k}\det(A'_{12,lk}) =$$

and the $a_{1k}a_{2l}$ term in $\det(A)$'s determinant is

$$(-1)^{k-1}\cdot a_{1k}\cdot(-1)^{l-1}a_{2l}\det(A_{12,lk}).$$

Because the matrices A and A' only differ in their first two rows, we know that $\det(A'_{12,lk}) = \det(A_{12,lk})$; therefore, by combining powers of -1, we can see that the A' terms are precisely -1 times the A terms.

Similarly, if we examine all of the other terms – i.e. the terms where the $a_{1k}a_{2l}$ is such that $k < l$ – we have that the $\det(A')$ terms are of the form

$$(-1)^{l-1}\cdot a_{2l}\cdot(-1)^{k-1}a_{1k}\det(A'_{12,lk}) =$$

and the $\det(A)$ terms are of the form

$$(-1)^{k-1}\cdot a_{1k}\cdot(-1)^{l-2}a_{2l}\det(A_{12,lk}).$$

Again, they disagree by simply their sign. So, because every individual term in the $\det(A')$ sum is precisely -1 times the corresponding term in the $\det(A)$ sum, we can say that $\det(A) = -\det(A')$.

So, the case when $i = 1$ is done! Don't worry: the case when $i > 1$ is far easier.

In fact, because when $i > 1$ our matrix is of the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & \cdots & a_{(i+1)n}\\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{in}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn}\end{pmatrix},$$

we have that

$$\det(A') = \sum_{k=1}^{n+1}(-1)^{k-1}a_{1k}\cdot\det(A'_{1k}),$$

where the A'_{1j}'s are all just the matrices A_{1j} with two rows swapped. By our inductive hypothesis, we then have that $\det(A'_{1k}) = -\det(A_{1k})$, and thus that

$$\begin{aligned}\det(A') &= \sum_{k=1}^{n+1}(-1)^{k-1}a_{1k}\cdot\det(A'_{1k})\\ &= -\sum_{k=1}^{n+1}(-1)^{k-1}a_{1k}\cdot\det(A'_{1k})\\ &= -\det(A).\end{aligned}$$

We now study the last remaining row operation:

Theorem 5 *Suppose that A is a $n \times n$ matrix. If A' is the matrix acquired by taking A and adding λ times the j-th row of A to its k-th row, then* $\det(A') = \det(A)$.

Proof. We actually don't need to proceed by induction here, though we do need to consider two cases.

When $n = 2$, this is trivial, as for $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we either have

$$\det(A') = \det\begin{pmatrix} a+\lambda c & b+\lambda c \\ c & d \end{pmatrix} = (a+\lambda c)b - (b+\lambda c)d = ac - bd = \det(A), \text{ or}$$
$$\det(A') = \det\begin{pmatrix} a & b \\ c+\lambda a & d+\lambda b \end{pmatrix} = a(d+\lambda b) - b(c+\lambda a) = ac - bd = \det(A).$$

When $n > 2$, we can always assume that k, the row we're adding to, is 1. Why is this? Well, if it's not 1, we can simply swap the k-th and first rows, and then swap two other non-k rows, which changes the determinant by $(-1)\cdot(-1) = 1$. (Note that we need three rows to do this; hence $n = 2$ is treated as a special case.)

In this case, we have that A' is of the form

$$\begin{pmatrix} a_{11}+\lambda a_{j1} & a_{12}+\lambda a_{j2} & \dots & a_{1n}+\lambda a_{jn} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Therefore, if we expand $\det(A')$, we get

$$\begin{aligned}\det(A') &= \sum_{i=1}^{n+1}(-1)^{i-1}(a_{1i}+\lambda a_{ji})\cdot\det(A'_{1i})\\ &= \sum_{i=1}^{n+1}(-1)^{i-1}a_{1i}\det(A'_{1i}) + \lambda\sum_{i=1}^{n+1}(-1)^{i-1}a_{ji}\cdot\det(A'_{1i})\end{aligned}$$

If we use the observation that (because A and A' agree on every row but the first) $A_{1i} = A'_{1i}$, we can then see that

$$\det(A') = \sum_{i=1}^{n+1}(-1)^{i-1}a_{1i}\det(A_{1i}) + \lambda\sum_{i=1}^{n+1}(-1)^{i-1}a_{ji}\cdot\det(A_{1i})$$

$$= \det(A) + \det\begin{pmatrix} a_{j1} & a_{j2} & a_{j3} & \dots & a_{jn} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & a_{j3} & \dots & a_{jn} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}.$$

Notice that because the matrix above has identical first and j-th rows, we can swap those two rows without changing the matrix (and thus without changing its determinant.) However, we know that swapping any two rows of a matrix multiplies the determinant by -1: therefore, we know that the matrix above must have determinant 0 (as 0 is the only number x such that $(-1)x = x$.)

Thus, we have that $\det(A') = \det(A)$, as claimed.

To summarize, we've proven the following four things:

- The determinant of the identity matrix is 1.
- Multiplying one of the rows of a matrix by some constant λ also multiplies that matrix's determinant by λ.
- Switching two rows in a matrix switches the sign of that matrix's determinant.
- Adding a multiple of one row to another in a matrix does not change its determinant.

These are relatively simple facts, that we were able to deduce in about a hour's worth of blindly using induction and expanding sums: yet, they will allow us to prove a host of rather difficult theorems without even breaking a sweat! To see why, it's useful to use the language of **elementary matrices**, which we review in the next section:

4 A Quick Aside: Elementary Matrices

So, we have all of these different kinds of row operations we can perform on matrices. A natural question to ask, then, is the following: Can we do these row operations to matrices by simply multiplying by some sort of special matrix? In other words, is there a $n \times n$ matrix that will swap the first two rows of anything you multiply it with? Or a matrix that will add 3 times the second row to the fourth row of any matrix you multiply it with?

As it turns out, yes! These matrices are called **elementary matrices**, which we review below:

Definition. There are three different kinds of **elementary matrices**, corresponding to the three different types of row operations. We list them here:

$$E_{\text{multiply row k by }\lambda} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

If the λ is in the (i,i)-th spot, multiplying A on the left by this matrix multiplies A's i-th row by λ.

$$E_{\text{switch rows i and j}} = \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & \vdots & \dots & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 & \dots & \vdots & \dots & 0 \\ 0 & 0 & \ddots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

The matrix above is the standard identity matrix with its i-th and j-th columns (highlighted) switched. Multiplying A on the left by this matrix switches A's i-th and j-th rows.

$$E_{\text{add }\lambda\cdot\text{row j to row i}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

If the λ is in the (i,j)-th spot, multiplying A on the left by this matrix adds λ times row j of A to row i of A.

Theorem 6 *The matrices above do what they say they do.*

Proof. Left to the reader! Seriously, though, it's easy; just check it.

So, in the language of elementary matrices, the three results we've proven about row operations and the determinant can be restated as follows:

1. For any matrix A,

$$\det(E_{\text{multiply row k by }\lambda} \cdot A) = \det(E_{\text{multiply row k by }\lambda}) \cdot \det(A) = \lambda \det(A).$$

2. For any matrix A,

$$\det(E_{\text{switch rows i and j}} \cdot A) = \det(E_{\text{switch rows i and j}}) \cdot \det(A) = -\det(A).$$

3. For any matrix A,

$$\det(E_{\text{add } \lambda\cdot\text{row j to row i}} \cdot A) = \det(E_{\text{add}\lambda\cdot\text{row j to row i}}) \cdot \det(A) = \det(A).$$

5 The Determinant: Key Results

With these results in hand, we move to our first large theorem:

Theorem 7 *For any $n \times n$ matrix A, A is nonsingular if and only if* $\det(A) \neq 0$.

Proof. To prove this theorem, we need to prove two statements: (1) whenever A is nonsingular, $\det(A) \neq 0$, and (2) whenever A is singular, $\det(A) = 0$.

We start by proving the first statement. Assume that A is nonsingular. Then, we know that A's reduced row-echelon form is the identity matrix, by definition. In other words, we know that we can start with the identity matrix I, and perform a series of row operations $r_1, \ldots r_n$ to transform I into A.

If we let $E_1, \ldots E_n$ be the elementary matrices corresponding to these row operations, this just says that we have elementary matrices $E_1, \ldots E_n$ such that

$$A = E_1 \cdot \ldots \cdot E_n \cdot I.$$

Now, if we take determinants of both sides, we have that

$$\det(A) = \det(E_1 \cdot \ldots \cdot E_n \cdot I);$$

finally, because we've proven that the determinant distributes across all three kinds of elementary matrices, we have at last that

$$\det(A) = \det(E_1) \cdot \ldots \cdot \det(E_n) \cdot \det(I).$$

However, we know that the determinants of elementary matrices are either $\lambda \neq 0$, -1, or 1 for the three various kinds of elementary matrices: therefore, we know that the product of all of the values on the right-hand side above is nonzero, as none of the individual terms are 0. Therefore, we've shown that $\det(A) \neq 0$, as claimed.

Now, all we have to do is prove the opposite direction: that A being singular implies that $\det(A) = 0$. To do this, simply recall that A being singular implies that the reduced row-echelon form of A has an all-zero row: i.e. that there are a series of row operations $r_1, \ldots r_n$ and some matrix M with an all-zeroes row such that performing these row operations on M creates A.

As before, if we let $E_1, \ldots E_n$ be the row operations corresponding to these row operations, we then have that

$$\begin{aligned} A &= E_1 \cdot \ldots \cdot E_n \cdot M \\ \Rightarrow \det(A) &= \det(E_1 \cdot \ldots \cdot E_n \cdot M) \\ &= \det(E_1) \cdot \ldots \cdot \det(E_n) \cdot \det(M). \end{aligned}$$

By swapping rows if necessary, we can insure that M's zero row is the first row in M; therefore, by the definition of the determinant, we know that $\det(M) = 0$, and thus that $\det(A) = 0$ as well. This establishes the other side of our "if and only if" claim; so we have proven our theorem.

Theorem 8 *For any pair of $n \times n$ matrices A, B, $\det(AB) = \det(A) \cdot \det(B)$.*

Proof. The language of elementary matrices makes this proof as simple as the one above it!

To start, notice that if either A or B are singular, AB must also be singular (as $B(AB)^{-1}$ is a right inverse for A whenever $(AB)^{-1}$ exists; similarly, $(AB)^{-1}A$ is always a left inverse for B whenenver $(AB)^{-1}$ exists.)

Therefore, whenever either $\det(A)$ or $\det(B) = 0$, we know that $\det(AB)$ is also 0, by our earlier theorem. So it suffices to prove our claim when both $\det(A)$ and $\det(B)$ are nonzero: i.e. when both A and B are nonsingular.

In this case, just as before, notice that we can always find elementary matrices $E_1^A, \ldots E_n^A, E_1^B, \ldots E_m^B$ such that

$$
\begin{aligned}
A &= E_1^A \cdot \ldots \cdot E_n^A, \text{ and} \\
B &= E_1^B \cdot \ldots \cdot E_m^B.
\end{aligned}
$$

Then, we have that

$$
\begin{aligned}
\det(AB) &= \det(E_1^A \cdot \ldots \cdot E_n^A \cdot E_1^B \cdot \ldots \cdot E_m^B) \\
&= \det(E_1^A) \cdot \ldots \det(E_n^A) \cdot \det(E_1^B) \cdot \ldots \det(E_m^B) \\
&= \det(E_1^A \cdot \ldots \cdot E_n^A) \cdot \det(E_1^B \cdot \ldots \cdot E_m^B) \\
&= \det(A)\det(B).
\end{aligned}
$$

So we've proven our claim!

Theorem 9 *For any $n \times n$ matrix A, $\det(A^T) = \det(A)$.*

Proof. This is the third (and final) theorem that we can prove trivially with our elementary matrix notation.

To start, first notice that

$$
\begin{array}{rcl}
(E_{\text{swap rows i and j}})^T & = & E_{\text{swap rows i and j}}, \\
(E_{\text{multiply row k by } \lambda})^T & = & E_{\text{multiply row k by } \lambda}, \text{ and} \\
(E_{\text{add } \lambda \text{ times row j to row k}})^T & = & E_{\text{add } \lambda \text{ times row k to row j}}.
\end{array}
$$

(Check these by hand if you're not persuaded; it's pretty quick!)

Because of the above observations, we know that taking the transpose of any elementary matrix doesn't change its type or constant λ (for those that involve a constant λ: therefore, we know that transposition doesn't change the determinant with respect to elementary matrices. In other words,

$$\begin{aligned}
\det((E_{\text{swap rows i and j}})^T) &= \det(E_{\text{swap rows i and j}}) \\
\det((E_{\text{multiply row k by }\lambda})^T) &= \det(E_{\text{multiply row k by }\lambda}), \text{ and} \\
\det((E_{\text{add }\lambda\text{ times row j to row k}})^T) &= \det(E_{\text{add }\lambda\text{ times row j to row k}}).
\end{aligned}$$

Now, we proceed just like we did in our earlier problems. First, notice that because $A \cdot A^{-1} = I = (A \cdot A^{-1})^T = (A^{-1})^T A^T$ whenever either A or A^T have an inverse, we know that A being singular is equivalent to A^T being singular, and thus $\det(A) = 0$ iff $\det(A^T) = 0$.

Now, suppose that $\det(A) \neq 0$. Then, as before, we can write

$$\begin{aligned}
A &= E_1 \cdot \ldots \cdot E_n \\
\Rightarrow \det(A) &= \det(E_1 \cdot \ldots \cdot E_n) \\
&= \det(E_1) \cdot \ldots \cdot \det(E_n) \\
&= \det(E_1^T) \cdot \ldots \cdot \det(E_n^T) \\
&= \det(E_n^T) \cdot \ldots \cdot \det(E_1^T) \\
&= \det(E_n^T \cdot \ldots \cdot E_1^T) \\
&= \det(A^T).
\end{aligned}$$

So we've proven our claim.

This last theorem then allows us to pretty much replace the word "row" with the word "column" in all of our previous theorems, just by taking the transpose and working there! The following theorem makes this explicit:

Theorem 10 *We can expand the determinant of any matrix along its first column, instead of its first row. In other words,*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i-1} a_{i1} \det(A_{i1}).$$

Furthermore, the three theorems we proved about row operations on matrices and determinants have analogues with respect to column operations:

- *Multiplying one of the columns of a matrix by some constant λ also multiplies that matrix's determinant by λ.*
- *Switching two columns in a matrix switches the sign of that matrix's determinant.*
- *Adding a multiple of one column to another column in a matrix does not change its determinant.*

Proof. To prove the first claim, simply note that because $\det(A) = \det(A^T)$, we have

$$\det(A) = \det(A^T) = \sum_{i=1}^{n} (-1)^{i-1} a_{i1} \det(A_{i1}).$$

Similarly, if we multiply the column of a matrix by some constant λ, we just need to take its transpose, multiply by the appropriate row-elementary matrix, and then take the transpose again. As taking the transpose never changes the matrix, the only step that changes the determinant is when we multipled by the row-elementary matrix, which multiplied the determinant by λ. Identical reasoning proves these results for the other two column operations.