Math/CS 103 Professor: Padraic Bartlett

Lecture 14: How to Calculate the Determinant

11/27/13 *UCSB 2013*

On our last problem set, we introduced the "positive determinant" $\det^+$, calculated a number of examples of this object, and proved a number of properties about this determinant! We restate those properties along with their proofs here, with some useful additional notes. Note that the proofs here are simply the ones I came up with; there are many many proofs of these properties, each of which have their own virtues! I simply made these because I like them. We'll likely be discussing these proofs through Wednesday, so don't worry if they seem odd now!

After these proofs, we introduce some new properties on the determinant, that make it much easier to calculate!

1 Background on Elementary Matrices

Remember elementary matrices? They're still really important. In particular, we proved the following theorem about them like two weeks ago:

Theorem 1. *Take any $n \times n$ matrix A. Suppose that we are looking at the composition $E \cdot A$, where E is one of our elementary matrices. Then, we have the following three possible situations:*

- *if $E = E_{\text{multiply entry } k \text{ by } \lambda}$, then $E \cdot A$ would be the matrix A with its k-th row multipled by λ.*
- *if $E = E_{\text{switch entry } k \text{ and entry } l}$, then $E \cdot A$ would be the matrix A with its k-th and l-th rows swapped, and*
- *if $E = E_{\text{add } \lambda \text{ copies of entry } k \text{ to entry } l}$, then $E \cdot A$ would be the matrix A with λ copies of its k-th row added to its l-th row.*

There is a converse theorem to this about products of the form $A \cdot E$, whose proof is identical:

Theorem 2. *Take any $n \times n$ matrix A. Suppose that we are looking at the composition $A \cdot E$, where E is one of our elementary matrices. Then, we have the following three possible situations:*

- *if $E = E_{\text{multiply entry } k \text{ by } \lambda}$, then $A \cdot E$ would be the matrix A with its k-th column multipled by λ.*
- *if $E = E_{\text{switch entry } k \text{ and entry } l}$, then $A \cdot E$ would be the matrix A with its k-th and l-th columns swapped, and*
- *if $E = E_{\text{add } \lambda \text{ copies of entry } k \text{ to entry } l}$, then $A \cdot E$ would be the matrix A with λ copies of its k-th column added to its l-th column.*

This observation is really useful to us! Another observation, which we proved for 3×3 matrices on the HW, was the following:

Theorem. Let A be a $n \times n$ matrix. Then there is some string of elementary matrices $E_1, \dots E_k$ such that

$$A = E_1 \cdot \ldots \cdot E_k$$

The proof is kinda tedious, and mostly identical to the 3×3 case. We put a proof of it at the end of the notes; but for now, simply believe it to be true.

Using these, we can start our proofs from the HW:

2 Elementary Matrices and the Determinant, Round 1

We start by making the following observations:

Theorem. Take any matrix A. Look at the matrix $A \cdot E$, where E is an elementary matrix of the form

$$E_{\text{multiply entry } k \text{ by } \lambda}.$$

Then

$$\det{}^+(A \cdot E) = |\lambda| \cdot \det{}^+(A).$$

Proof. In class/notes, we said that to find the volume of an arbitrary parallelotope, we do the following:

1. First, pick one of the vectors in our parallotope, and calculate its length: this is the "base."
2. Then, pick a vector we haven't yet studied, and find its "height" over the vector we picked earlier by constructing the "orth" vector from our past class, and studying its length.
3. Pick a third vector, and again find its "height" over the previous two vectors by constructing the "orth" vector over those earlier two vectors.
4. Repeat this process until we run out of vectors. The product of the base with all of these heights gives us the volume!

Now, make the following observations:

- The positive determinant of A is just the volume of the parallelotope spanned by the column vectors of A.
- The matrix $A \cdot E$ is just the matrix A with its k-th column multiplied by λ.
- Therefore, the positive determinant of $A \cdot E$ is the volumne of the parallelotope spanned by the column vectors of A, where one of them is multiplied by λ.

- If we pick the λ-multiple as our first vector when calculating the volume, it is clear that the length of the base is scaled by $|\lambda|$, and the length of any height vector is unchanged (as those are calculated by looking at things orthogonal to the base, and therefore do not care about the length of the base!)
- Therefore, the volume of $A \cdot E$ is just the volume of A scaled by $|\lambda|$.

Done! $\square$

Theorem. Take any matrix A. Look at the matrix $A \cdot E$, where E is an elementary matrix of the form

$$E_{\text{switch entry } k \text{ and entry } l}.$$

Then

$$\det{}^+(A \cdot E) = \det{}^+(A).$$

Proof. This is like the above proof, but even easier. First, notice that the matrix $A \cdot E$ is just the matrix A, but with two columns swapped. Therefore, the volume of the parallelepiped spanned by the columns of $A \cdot E$ is the same as the volume of the parallelepiped spanned by the columns of A, because they're both the volumes of the **same** paralellepiped! $\square$

Theorem. Take any matrix A. Look at the matrix $A \cdot E$, where E is an elementary matrix of the form

$$E_{\text{add } \lambda \text{ copies of entry } k \text{ to entry } l}.$$

Then

$$\det{}^+(A \cdot E) = \det{}^+(A).$$

Proof. This is also like the above proof, but slightly harder. Specifically, calculate the volume of the parallelepiped spanned by $A \cdot E$, simultaneously with the volume of the parallelepiped spanned by A's columns, in the following way:

1. First, pick column k to be the base in both cases. Note that both matrices have the same k-th column (as the only column at which they differ is the j0th column), and therefore that the length here is the same.

2. Now, pick column j to be the next vector we take in both cases. We are interested in the "height" of this column vector over the k-th column vector.

   Notice that because the j-th column vector of $A \cdot E$ is just the j-th column of A, plus λ copies of the k-th column, we have that

$$\begin{aligned}&\text{orth}\Big((\text{the } j\text{-th column of } A \cdot E) \text{ onto } (\text{the } k\text{-th column of } A)\Big)\\ =&\text{orth}\Big((\text{the } j\text{-th column of } A) \text{ onto } (\text{the } k\text{-th column of } A)\Big).\end{aligned}$$

This is because adding copies of the k-th column to a vector doesn't change the "amount" of that vector that is orthogonal to that k-th column! (Basically, imagine adding λ copies of a vector $\vec{w}$ to another vector $\vec{v}$. This directly increases the quantity proj($\vec{v}$ onto $\vec{w}$) by $\lambda\vec{v}$; therefore, when we form the vector orth($\vec{v}$ onto $\vec{w}$) $= \vec{v} -$ proj($\vec{v}$ onto $\vec{w}$), we subtract those copies off again!

Therefore, the height of the k-th column over our j-th column is the same in both cases.

3. Now, notice that the spans of the k, j-th columns in the matrices $A, A \cdot E$ are the same in both cases, as they both consist of all multiples of the k-th and j-th columns! Therefore, the "height" of any other vector over these two is unchanged.

Consequently, because the lengths of the base and of the heights are unchanged at each step, these two paralleletopes have the same volume. $\square$

These three proofs are all very similar to question 2 on the HW, which asked you what happened to the volume of various parallelotopes after elementary matrices were applied to them! However, question two is really asking you for the volume of $E \cdot A$, for an arbitrary elementary matrix E and a paralleotope spanned by the columns of A; this is subtly different!

To answer this, then, you need the answer to question 1:

3 The Transpose and the Determinant

The transpose is the following object:

Definition. Take an $n \times n$ matrix A of the form

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

Define the **transpose** of this matrix, A^T, as the matrix where we "flip" A over its top left-bottom right diagonal, i.e. where we switch the rows and columns of A, i.e. where we put the entry a_{ji} in the (i,j)-th entry of A^T, i.e.

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

First, a useful warm-up result:

Theorem. Take any string of matrices $A_1, \dots A_n$, such that $A_1 \cdot \dots \cdot A_n$ is well-defined. Then

$$(A_1 \cdot A_2 \cdot \dots \cdot A_n)^T = A_n^T \cdot A_{n-1}^T \cdot \dots \cdot A_2^T \cdot A_1^T.$$

Proof. First, notice that for any two $n \times n$ matrices A, B, we have $(A \cdot B)^T = (B^T) \cdot (A^T)$. This is not too hard to show: simply notice that due to the definition of matrix multiplication, the entry in (i, j) of $A \cdot B$ is

$$(\text{the } i\text{-th row of } A) \cdot (\text{the } j\text{-th column of } B),$$

which means that the (j, i)-th entry of $(A \cdot B)^T$ is precisely that dot product.

On the other hand, notice that the j-th row of B^T is just $(b_{1,j}, \ldots b_{n,j})$, in other words the j-th column of B. Similarly, the i-th column of A^T is $(a_{i,1}, \ldots a_{i,n})$, the i-th row of A.

Therefore, we have that the (j, i)-th entry of $(B^T) \cdot (A^T)$ is just

$$\begin{aligned}
&(\text{the } j\text{-th row of } B^T) \cdot (\text{the } i\text{-th column of } A^T) \\
=&(\text{the } j\text{-th column of } B) \cdot (\text{the } i\text{-th row of } A) \\
=&(\text{the } i\text{-th row of } A) \cdot (\text{the } j\text{-th column of } B).
\end{aligned}$$

Therefore the two matrices $(A \cdot B)^T$, $(B^T) \cdot (A^T)$ have the same entries, and are therefore the same matrices!

In particular, this tells us that if we look at the transpose of a product of n matrices together, we have that it's just the product of their transposes in reverse order! This is because

$$\begin{aligned}
(A_1 \cdot A_2 \cdot \ldots \cdot A_n)^T &= (A_1 \cdot (A_2 \cdot \ldots \cdot A_n))^T \\
&= (A_2 \cdot \ldots \cdot A_n)^T \cdot A_1^T \\
&= (A_2 \cdot (A_3 \cdot \ldots \cdot A_n))^T \cdot A_1^T \\
&= (A_3 \cdot \ldots A_n)^T \cdot A_2^T \cdot A_1^T \\
&= \quad \vdots \\
&= A_n^T \cdot A_{n-1}^T \cdot \ldots \cdot A_2^T \cdot A_1^T.
\end{aligned}$$

□

This, along with our earlier work on elementary matrices, makes the following result trivial:

Theorem. For any $n \times n$ matrix A, we have $\det^+(A^T) = \det(A)$.

Proof. We proceed as follows:

1. First, write A as the product of elementary matrices $E_1 \cdot \ldots \cdot E_n$. Out of this collection, find all of the elementary matrices of the form $E_{\text{multiply entry } k \text{ by } \lambda}$. Take their corresponding λ's, and enumerate them as $\lambda_1, \ldots, \lambda_k$.

2. Then, on one hand we can express

$$\det{}^+(A) = |\lambda_1 \cdot \ldots \cdot \lambda_k|,$$

. where we regard this product as being 1 if there are no such lambda's. This is because we've shown in our work with elementary matrices that

$$\det{}^+(E_1 \cdot \ldots \cdot E_n) = \det{}^+(E_1 \cdot \ldots \cdot E_{n-1}) \cdot \alpha,$$

where α is equal to 1 if E_n was a "swap" or" add some multiple of an entry to another entry" matrix, and λ if it was a a "multiply an entry by λ" matrix.

Therefore, we can progressively pull off elementary matrices from the right of $E_1 \cdot \ldots \cdot E_n$, with the positive determinant changing by a factor of λ whenever we pull off a "multiply" matrix and unchanging otherwise. At the end, we'll get $|\lambda_1 \cdot \ldots \cdot \lambda_k|$, as claimed!

3. Conversely, notice that we can use our earlier results on the transpose to write

$$A^T = E_n^T \cdot \ldots \cdot E_1^T.$$

Furthermore, notice the following facts:

- The transpose of the "swap entries i and j" matrix is just itself, because this matrix is symmetric across the top left-bottom right diagonal!
- Similarly, the transpose of the "multiply entry k by λ" matrix is just itself, because this matrix is also symmetric across the top left-bottom right diagonal!
- Finally, the transpose of the "add λ copies of entry i to entry j" matrix is the identity matrix except it has a λ in entry (i, j): in other words, it's the "add λ copies of entry j to entry i" matrix!

Therefore, in particular, we know that the transpose of an elementary matrix is an elementary matrix of the same type and same λ! Therefore, repeating our proof from (2) gives us that

$$\det{}^+(E_n^T \cdot \ldots \cdot E_1^T) = |\lambda_1 \cdot \ldots \cdot \lambda_k|.$$

This proves our claim! $\square$

This gives us delightfully trivial proofs of problem (2) on HW#18:

4 Elementary Matrices and the Determinant, Round 2

Theorem. Take any matrix A. Look at the matrix $E \cdot A$, where E is an elementary matrix of the form

$$E_{\text{add } \lambda \text{ copies of entry } k \text{ to entry } l}, \text{ or } E_{\text{switch entry } k \text{ and entry } l}.$$

Then

$$\det{}^+(E \cdot A) = \det{}^+(A).$$

Proof. First, notice that

$$\det{}^+(E \cdot A) = \det{}^+\left((E \cdot A)^T\right).$$

Then, notice that

$$(E \cdot A)^T = A^T \cdot E^T.$$

Now, notice that E^T, as shown before, is still an elementary matrix of the same type as E: i.e. it's still a swap or add-type of elementary matrix.

Therefore, we can use our earlier results on elementary matrices and the determinant to note that

$$\det{}^+(A^T \cdot E^T) = \det{}^+(A^T) = \det{}^+(A),$$

and therefore that

$$\det{}^+(E \cdot A) = \det{}^+(A^T \cdot E^T) = \det{}^+(A^T) = \det{}^+(A).$$

□

Theorem. Take any matrix A. Look at the matrix $E \cdot A$, where E is an elementary matrix of the form

$$E_{\text{multiply entry } k \text{ by } \lambda}.$$

Then

$$\det{}^+(A \cdot E) = |\lambda| \cdot \det{}^+(A).$$

Proof. This is the same argument as that above; we use our earlier results to note that

$$\det{}^+(E \cdot A) = \det{}^+\left((E \cdot A)^T\right) = \det{}^+(A^T \cdot E^T) = \det{}^+(A^T) \cdot |\lambda| = \det{}^+(A) \cdot |\lambda|.$$

□

5 The Determinant and Multiplication

We finally note finally that problem 3 is made pretty trivial by the techniques we used to do problem 1:

Theorem. Take any two $n \times n$ matrices A, B. Then

$$\det{}^+(A \cdot B) = \det{}^+(A) \cdot \det{}^+(B).$$

Proof. First, write A,B as products of elementary matrices:

$$A = E_1 \cdot \ldots \cdot E_n, B = E_{n+1} \cdot \ldots \cdot E_{n+m}.$$

Let $\lambda_1, \ldots \lambda_k$ denote the coefficients corresponding to all of the "multiply an entry by λ" elementary matrices above in A, and $\lambda_{k+1}, \ldots \lambda_{k+l}$ denote those coefficients in B. Then, via the exact proof used earlier in the transpose problem, we have

$$\begin{aligned} \det{}^+(A) &= |\lambda_1 \cdot \ldots \cdot \lambda_k|, \\ \det{}^+(B) &= |\lambda_{k+1} \cdot \ldots \cdot \lambda_{k+l}|, \text{ and} \\ \det{}^+(A \cdot B) &= |\lambda_1 \cdot \ldots \cdot \lambda k + l|. \end{aligned}$$

Therefore, the positive determinant of the product of two matrices is the product of the positive determinants of these two matrices. □

6 The General Determinant

Woo! Ok. That was a lot. But we're nearly done.

Definition. The **determinant** (as opposed to the "positive determinant") of a matrix A is defined as follows:

1. Take A, and write it as the product $E_1 \cdot \ldots \cdot E_n$ of elementary matrices.
2. To find the determinant $\det(A)$ of A, look at these elementary matrices. Let $\lambda_1, \ldots \lambda_k$ denote the constants that show up in the "multiply an entry by λ_i" elementary matrices, and l denote the number of "swap" elementary matrices. Then

$$\det(A) = (-1)^l \cdot \lambda_1 \cdot \ldots \cdot \lambda_k$$

This object, in other words, is just the positive determinant from before, i.e. the volume, except multiplied by a factor of ± 1 depending on the signs of the constants λ_i and the number of swaps performed. This gives us the following observation for free:

Observation. For any matrix A, $|\det(A)| = \det^+(A)$.

This isn't the only thing we get for free! If you go through all of the proofs in this talk and remove the absolute-value signs and add in $(-1)^{\text{number of swaps}}$ in relevant places, you'll find that every proof we just did has an analogue for the determinant! In other words,

- $\det(AB) = \det(A)\det(B)$,
- $\det(A^T) = \det(A)$,
- $\det(EA) = \det(A)$ if E is an "add copies of one entry to another" elementary matrix, $-\det(A)$ if E is a swap matrix, and $\lambda \det(A)$ if E is a "multiply entry by λ'' matrix.

We close this talk by mentioning one property of the determinant that doesn't hold for the positive determinant: **n-linearity**!

Definition. Let T be a map from $n \times n$ matrices of real numbers to $\mathbb{R}$. We say that T is **n-linear** if the following always holds:

- Take any matrix A, with columns $\vec{a_{c_1}}, \dots \vec{a_{c_n}}$.
- Suppose that $\vec{a_{c_i}}$ is equal to some sum of vectors $\vec{x} + \vec{y}$.
- Then, consider the two matrices created by replacing this i-th column with the vectors $\vec{x}, \vec{y}$ respectively:

$$A_x = \begin{bmatrix} \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vec{a_{c_1}} & \dots & \vec{a_{c_{i-1}}} & \vec{x} & \vec{a_{c_{i+1}}} & \dots & \vec{a_{c_n}} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \end{bmatrix},$$
$$A_y = \begin{bmatrix} \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vec{a_{c_1}} & \dots & \vec{a_{c_{i-1}}} & \vec{y} & \vec{a_{c_{i+1}}} & \dots & \vec{a_{c_n}} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \end{bmatrix},$$

A map is called **n-linear** if

$$T(A) = T(A_x) + T(A_y),$$

for any column $\vec{a_{c_i}}$ and pair of vectors $\vec{x}, \vec{y}$ such that $\vec{x} + \vec{y} = \vec{a_{c_i}}$.

Theorem. The determinant is **n-linear**.

Proof. We first note a quick example that shows why the positive determinant is not n-linear: simply observe that the positive determinant of

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

is 0, as the parallelogram spanned by $(0, 0), (0, 1)$ has zero area. However, the parallelograms spanned by

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

both have area 1! Therefore, because $0 \neq 1 + 1$, the positive determinant is not n-linear.

To see that the determinant is n-linear: take any matrix A, any column $\vec{a_{c_i}}$, and any pair of vectors $\vec{x}, \vec{y}$ such that $\vec{x} + \vec{y} = \vec{a_{c_i}}$.

Write both of the vectors $\vec{x}, \vec{y}$ as linear combinations

$$\vec{x} = \text{proj}(\vec{x} \text{ onto columns of } A) + \text{orth}(\vec{x} \text{ onto columns of } A),$$
$$\vec{y} = \text{proj}(\vec{y} \text{ onto columns of } A) + \text{orth}(\vec{y} \text{ onto columns of } A).$$

Notice that because $\mathrm{x} + \mathrm{y} = \vec{a_{c_i}}$, we have

$$\text{orth}(\vec{x} + \vec{y} \text{ onto columns of } A) = \text{orth}(\vec{a_{c_i}} \text{ onto columns of } A) = \vec{0},$$

because $\vec{a_{c_i}}$ is itself a column of A. Consequently, we have

$$\text{orth}(\vec{x} \text{ onto columns of } A) = -\text{orth}(\vec{y} \text{ onto columns of } A).$$

Now, notice that for the matrix A_x, we have

$$\det(A_x) = \det(A_x \cdot E_{\text{add } \lambda \text{ copies of entry } k \text{ to entry } l},$$

for **any** λ, k, l such that $k \neq l$! In particular, if we write

$$\begin{aligned} \text{proj}(\vec{x} \text{ onto columns of } A) &= x_1\vec{a_{c_1}} + \ldots + x_n\vec{a_{c_n}}, \\ \text{proj}(\vec{y} \text{ onto columns of } A) &= y_1\vec{a_{c_1}} + \ldots + y_n\vec{a_{c_n}}, \end{aligned}$$

we can use these matrices to see that

$$\det(A_x) = \det\left(A_x \cdot \overbrace{E_{\text{add } -x_1 \text{ copies of entry 1 to entry } i} \cdot \ldots \cdot E_{\text{add } -x_n \text{ copies of entry } n \text{ to entry } i}}^{\text{these matrices add } -x_j \text{ copies of each column } j \neq i \text{ to column } i}\right),$$

and similarly that

$$\det(A_y) = \det\left(A_y \cdot \overbrace{E_{\text{add } -y_1 \text{ copies of entry 1 to entry } i} \cdot \ldots \cdot E_{\text{add } -y_n \text{ copies of entry } n \text{ to entry } i}}^{\text{these matrices add } -y_j \text{ copies of each column } j \neq i \text{ to column } i}\right).$$

But

$$\begin{aligned} &A_x \cdot \overbrace{E_{\text{add } -x_1 \text{ copies of entry 1 to entry } i} \cdot \ldots \cdot E_{\text{add } -x_n \text{ copies of entry } n \text{ to entry } i}}^{\text{these matrices add } -x_j \text{ copies of each column } j \neq i \text{ to column } i} \\ &= \begin{bmatrix} \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vec{a_{c_1}} & \ldots & \vec{a_{c_{i-1}}} & \vec{x} - (\overbrace{x_1\vec{a_{c_1}} + \ldots + x_n\vec{a_{c_n}}}^{\text{all entries except} x_i\vec{a_{c_i}}}) & \vec{a_{c_{i+1}}} & \ldots & \vec{a_{c_n}} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \end{bmatrix} \\ &= \begin{bmatrix} \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vec{a_{c_1}} & \ldots & \vec{a_{c_{i-1}}} & \text{orth}(\vec{x} \text{ onto columns of } A) + x_i\vec{a_{c_i}} & \vec{a_{c_{i+1}}} & \ldots & \vec{a_{c_n}} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \end{bmatrix}, \end{aligned}$$

and similarly

$$\begin{aligned} &A_y \cdot \overbrace{E_{\text{add } -y_1 \text{ copies of entry 1 to entry } i} \cdot \ldots \cdot E_{\text{add } -y_n \text{ copies of entry } n \text{ to entry } i}}^{\text{these matrices add } -y_j \text{ copies of each column } j \neq i \text{ to column } i} \\ &= \begin{bmatrix} \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vec{a_{c_1}} & \ldots & \vec{a_{c_{i-1}}} & \text{orth}(\vec{y} \text{ onto columns of } A) + y_i\vec{a_{c_i}} & \vec{a_{c_{i+1}}} & \ldots & \vec{a_{c_n}} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \end{bmatrix}. \end{aligned}$$

Call these two matrices A'_x, A'_y.

There are now two possibilities.

1. The columns of A form a basis for $\mathbb{R}^n$. In this case, we have that the two orth components above are both zero, because there is nothing in $\mathbb{R}^n$ orthogonal to all of $\mathbb{R}^n$. Therefore, we have that

$$\begin{aligned}\det(A_x) + \det(A_y) =& \det(A \cdot E_{\text{multiply entry i by } x_i}) + \det(A \cdot E_{\text{multiply entry i by } y_i})\\ =& x_i \det(A) + y_i \det(A)\\ =& (x_i + y_i)\det(A).\end{aligned}$$

   So: what is $x_i + y_i$? On one hand, we know that

$$\vec{x} + \vec{y} = \vec{a_{c_i}} = x_1\vec{a_{c_1}} + \ldots + x_n\vec{a_{c_n}} + y_1\vec{a_{c_1}} + \ldots + y_n\vec{a_{c_n}},$$

   and therefore that

$$\vec{0} = (x_1 + y_1)\vec{a_{c_1}} + \ldots + (x_i + y_i - 1)\vec{a_{c_i}} + \ldots + (x_n + y_n)\vec{a_{c_n}}.$$

   This is a nontrivial linear combination of elements in a basis that equals 0; therefore, all of the coefficients above must be 0! As a result, we must have $x_i + y_i = 1$. This gives us

$$\det(A_x) + \det(A_y) = (x_i + y_i)\det(A) = \det(A),$$

   as requested.

2. Otherwise, the columns of A do not form a basis for $\mathbb{R}^n$. In this case, the columns of A are linearly dependent! Take a combination

$$b_1\vec{a_{c_1}} + \ldots + b_n\vec{a_{c_n}} = \vec{0}$$

   where not all of the b_k's are zero.

   If in this combination the coefficent b_i is zero, then there is a combination of the columns of A, not using the i-th column, that combines to zero! This means that for both of the matrices A_x, A_y, the columns of these matrices are also a linearly dependent set, because this combination does not use the i-th column. Therefore, we have that the determinants of these two matrices are zero, much like the determinant of A itself, because all three are matrices with linearly dependent columns (and therefore correspond to paralleletopes that live in a $n-1$ dimensional space, and thus have zero volume.)

   Otherwise, in this combination the b_i coefficient is nonzero. This gives us a way to express the i-th column of A as a linear combination of the other columns of A!

   Therefore, by using the $E_{\text{add } \lambda \text{ copies of entry } k \text{ to entry } l}$ in a similar way to before, we can subtract multiples of all of the other columns of A_x from the i-th column of A'_x, such that we get rid of the $x_i \vec{a_{c_i}}$ part, without changing the determinant! We can also do the same trick to the A'_y matrix; this gives us that $\det(A_x)$ is the determinant of the matrix that you get by replacing the i-th column of A with orth($\vec{x}$ onto columns of A), and similarly for $\det(A_y)$!

But, because orth($\vec{x}$ onto columns of A) $= -$orth($\vec{y}$ onto columns of A), we have that the determinants of these two matrices are the same, except one is the opposite sign of the other! Therefore, we have that $\det(A_x) = -\det(A_y)$, and thus

$$\det(A_x) + \det(A_y) = 0 = \det(A),$$

again because the columns of A are linearly dependent.

$\square$

Why on earth do we care? Well: the homework will tell you, but basically, it's because this property — **n-linearity** — lets us actually **calculate** the determinant with relative ease!

7 Appendix: A Proof That All Matrices Are Products Of Elementary Matrices

In case you were wondering.

Theorem. Let A be an arbitrary $n \times n$ matrix. Then we can write A as the product of elementary matrices.

Proof. To do this process, first do the following:

1. Take the collection R of all of A's rows.
2. If this set is linearly independent, you're done!
3. Otherwise, there is some row that shows up in this collection that is a combination of the other rows. Get rid of that row, and return to (2).

This creates a subset R' of A's rows that is linearly independent. Furthermore, it creates a subset from which we can create **any** of A's rows, even the ones we got rid of! This is because we only got rid of rows that were linearly dependent on the earlier ones; i.e. we only got rid of rows that we can make with the rows we kept!

So: all we need to do now is make B into a matrix that has all of the rows in this subset R'! If we can do this, then we can just do the following, using the elementary matrices that correspond to these row operations:

- Multiply all of the other rows in B by zero.
- Now, using each all-zero row as an empty slot, create each of the rows from A that we don't have by combining the rows from R'. We can do this because all of the remaining rows in A were linear combinations of the R' rows!

- Finally, rearrange the rows using swaps so that our matrix **is** A (and not just a matrix with the same rows, but in some different order.)

This is our plan! We execute the plan as below:

1. We start with B equal to the $n \times n$ identity matrix. Note that B's rows span all of $\mathbb{R}^n$
2. If all of the rows in R' currently occur as rows of B, stop!
3. Otherwise, there is a row $\vec{a_r}$ in R' that is not currently a row in B.
4. If the rows of B span $\mathbb{R}$, then specifically there is a combination of the rows of B that yields $\vec{a_r}$.
5. Furthermore, this vector is not just a combination of rows in R', because R' is a linearly independent set. Therefore, in any linear combination of B's rows that creates $\vec{a_r}$, there is some row of B that is not one of the R' rows that's used in creating $\vec{a_r}$.
6. So: take the linear combination

   $$a_1\vec{b_{r_1}} + \dots a_n\vec{b_{r_n}} = \vec{a_r},$$

   and let $\vec{b_{r_k}}$ denote the row that occurs above that's not one of the R' rows and that has $a_k \neq 0$.
7. Take B, and multiply it by

   $$\overbrace{\overbrace{\begin{bmatrix} & \vdots & \\ \dots & a_1 & \dots \\ & \vdots & \end{bmatrix}}^{\substack{\text{add } a_1 \text{ copies of} \\ r_1 \text{ to } r_k}} \cdot \overbrace{\begin{bmatrix} & \vdots & \\ \dots & a_2 & \dots \\ & \vdots & \end{bmatrix}}^{\substack{\text{add } a_2 \text{ copies of} \\ r_2 \text{ to } r_k}} \cdot \ldots \cdot \overbrace{\begin{bmatrix} & \vdots & \\ \dots & a_n & \dots \\ & \vdots & \end{bmatrix}}^{\substack{\text{add } a_n \text{ copies of} \\ r_2 \text{ to } r_k}}}^{\text{all of the values } a_i, \text{ with } i \neq k} \cdot \overbrace{\begin{bmatrix} & \vdots & \\ \dots & a_k & \dots \\ & \vdots & \end{bmatrix}}^{\substack{\text{multiply row } r_k \\ \text{by } a_k}} \cdot B$$

   This takes the k-th row of B and fills it with the linear combination that creates $\vec{a_r}$! So this means that the row $\vec{a_r}$ is now in B.
8. Also, notice that the rows of B all still span $\mathbb{R}^n$! This is because

   $$\begin{aligned} a_1\vec{b_{r_1}} + \dots a_n\vec{b_{r_n}} &= \vec{a_r} \\ \Rightarrow \vec{b_{r_k}} &= \frac{1}{a_k}\left(\underbrace{a_1\vec{b_{r_1}} + \dots a_n\vec{b_{r_n}}}_{\text{terms that aren't } a_k\vec{b_{r_k}}} + \vec{a_r} \right). \end{aligned}$$

   Therefore, we have that the old k-th row $\vec{b_{r_k}}$ is in the span of the new B's rows! As well, because none of the other rows changed, those rows are all still in the span as well. Therefore, because the new B's rows contain the old B's rows in their span, they must span $\mathbb{R}^n$!

9. Go to (2), and repeat this process!

The result of this process is a matrix B that contains all of the rows in R', which is what we wanted (because we can make A out of this!) So we're done. $\square$

To illustrate this argument, we run another example:

Example. Consider the matrix

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 4 & -1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

Write A as a product of elementary matrices.

Proof. We start, as directed in the proof, by finding a subset of A's rows that is linearly independent. We can tell at the start that the collection of all rows is not linearly independent, because

$$1(0,1,2) + 1(4,-1,0) - 2(2,0,1) = (0,0,0).$$

However, we also have that the pair

$$(0,1,2), (2,0,1)$$

is linearly independent, because

$$\alpha(0,1,2) + \beta(2,0,1) = (0,0,0) \Rightarrow \alpha, \beta = 0,$$

and that these two vectors contain the third in their span.

So the set R' from our discussion above is just these two vectors!

Set B equal to the 3×3 identity matrix. We start by picking a vector from R' – let's choose $\vec{a_r} = (0,1,2)$.

We want to multiply B by elementary matrices so that it has $(0,1,2)$ as one of its rows. To do this, we first write $(0,1,2)$ as a combination of B's rows:

$$0(1,0,0) + 1(0,1,0) + 2(0,0,1) = (0,1,2).$$

We now pick a row from B whose coefficient above is nonzero, and that isn't a row in R'. For example, the coefficient of the second row above is 1, and the second row $(0,1,0)$ is not in R': so we can pick the second row.

We now turn the second row into this $\vec{a_r} = (0,1,2)$, by using the linear combination we have for $(0,1,2)$ above:

$$\overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}^{\substack{\text{add 2 copies of} \\ r_3 \text{ to } r_2}} \cdot \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}^{\substack{\text{add 0 copies of} \\ r_1 \text{ to } r_2}} \cdot \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}^{\substack{\text{multiply row } r_2 \\ \text{by } 1}} \cdot \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}^{\text{the matrix } B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Success! We repeat this. We choose another row from R', specifically $\vec{a_r} = (2,0,1)$. We write $(2,0,1)$ as a combination of B's rows:

$$2(1,0,0) + 0(0,1,2) + 1(0,0,1) = (2,0,1).$$

We now pick a row from B whose coefficient above is nonzero, and that isn't a row in R'; for example, the first row works here.

We now turn the first row into this $\vec{a_r} = (2,0,1)$, by using the linear combination we have for $(2,0,1)$ above:

$$\overbrace{\begin{bmatrix}1&0&1\\0&1&0\\0&0&1\end{bmatrix}}^{\substack{\text{add 1 copies of}\\ r_3 \text{ to } r_1}} \cdot \overbrace{\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}}^{\substack{\text{add 0 copies of}\\ r_2 \text{ to } r_1}} \cdot \overbrace{\begin{bmatrix}2&0&0\\0&1&0\\0&0&1\end{bmatrix}}^{\substack{\text{multiply row } r_1\\ \text{by } 2}} \cdot \overbrace{\begin{bmatrix}1&0&0\\0&1&2\\0&0&1\end{bmatrix}}^{\text{the matrix } B} = \begin{bmatrix}2&0&1\\0&1&2\\0&0&1\end{bmatrix}.$$

We are now out of rows of R'! This brings us to the second stage of our proof: multiply all of the remaining rows that aren't R' rows by 0.

$$\overbrace{\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}}^{\substack{\text{multiply row } r_3\\ \text{by } 0}} \cdot \overbrace{\begin{bmatrix}2&0&1\\0&1&2\\0&0&1\end{bmatrix}}^{\text{the matrix } B} = \begin{bmatrix}2&0&1\\0&1&2\\0&0&0\end{bmatrix}.$$

Now we are at the last stage of our proof: combine the R' rows to create whatever rows in A are left, in these "blank" all-zero rows!

Specifically, we take the one row of A that's left: $(4,-1,0)$. As we noted before, we can write

$$(4,-1,0) = 2(2,0,1) - 1(0,1,2).$$

Therefore, we have

$$\overbrace{\begin{bmatrix}1&0&0\\0&1&0\\2&0&1\end{bmatrix}}^{\substack{\text{add 2 copies of}\\ r_1 \text{ to } r_3}} \cdot \overbrace{\begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix}}^{\substack{\text{add } -1 \text{ copies of}\\ r_2 \text{ to } r_3}} \cdot \overbrace{\begin{bmatrix}2&0&1\\0&1&2\\0&0&0\end{bmatrix}}^{\text{the matrix } B} = \begin{bmatrix}2&0&1\\0&1&2\\4&-1&0\end{bmatrix}.$$

So we have a matrix with the same rows as A! Finally, we just shuffle the rows of B to get A itself:

$$\overbrace{\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}}^{\substack{\text{switch rows}\\ r_3 \text{ and } r_2}} \cdot \overbrace{\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}}^{\substack{\text{switch rows}\\ r_2 \text{ and } r_1}} \cdot \overbrace{\begin{bmatrix}2&0&1\\0&1&2\\4&-1&0\end{bmatrix}}^{\text{the matrix } B} = \begin{bmatrix}0&1&2\\4&-1&0\\2&0&1\end{bmatrix} = A.$$

Win! $\square$