# A Derivation of Determinants

Mark Demers
Linear Algebra MA 435

March 25, 2019

> According to the precepts of elementary geometry, the concept of volume depends on the notions of length and angle and, in particular, perpendicularity... Nevertheless, it turns out that volume is independent of all these things, except for an arbitrary multiplicative constant that can be fixed by specifying that the unit cube have volume one.
>
> *Peter Lax*

We will adopt an approach to the determinant motivated by our intuitive notions of volume; however, the determinant of a matrix tells us much more. We list here some of its principal uses.

1. The determinant of a matrix gives the *signed volume* of the parallelepiped generated by its columns.

2. The determinant gives a criterion for invertibility. A matrix $A$ is invertible if and only if $\det(A) \neq 0$.

3. A formula for $A^{-1}$ can be given in terms of determinants; in addition, the entries of $x$ in the inverse equation $x = A^{-1}b$ can be expressed in terms of determinants. This is known as *Cramer's Rule*.

## 1 The Determinant of a $2 \times 2$ Matrix.

Viewing a square matrix $M$ as a linear transformation from $\mathbb{R}^n$ to itself leads us to ask the question: How does this transformation change volumes? In the case of a $2\times 2$ matrix, it is possible to compute the answer explicitly using some familiar facts from geometry and trigonometry.

Let $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and let $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. Define $M$ to be the matrix $M = [\vec{u}\ \ \vec{v}]$. To examine how $M$ transforms areas, we look at the action of $M$ on $\vec{e}_1$ and $\vec{e}_2$ (see Figure 1). $M\vec{e}_1 = \vec{u}$ and $M\vec{e}_2 = \vec{v}$ so that $M$ transforms the unit square determined by $\vec{e}_1$ and $\vec{e}_2$ into the parallelogram determined by $\vec{u}$ and $\vec{v}$.

Figure 2 shows the parallelogram determined by $\vec{u}$ and $\vec{v}$. We wish to find its area. The area of the parallelogram is given by $Area = base \times height = \|\vec{u}\|h$ where $\|\vec{u}\| = \sqrt{(u_1)^2 + (u_2)^2}$ is the length of the vector $\vec{u}$. Define

$$\begin{aligned} \theta &= \text{angle formed by } \vec{u} \text{ and } \vec{v} \text{ at the origin,} \\ \theta_u &= \text{angle formed by } \vec{u} \text{ and the positive } x\text{-axis,} \\ \theta_v &= \text{angle formed by } \vec{v} \text{ and the positive } x\text{-axis.} \end{aligned}$$

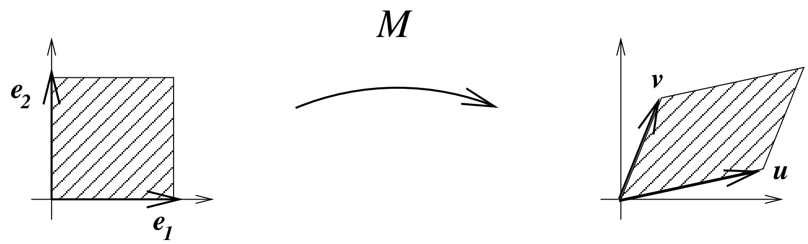



Figure 1: The action of $M$ on the unit square.

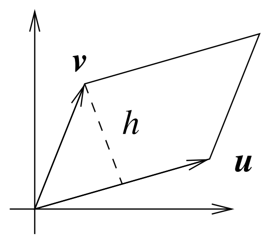



Figure 2: The parallelogram determined by $\vec{u}$ and $\vec{v}$.

Note that $\theta = \theta_v - \theta_u$. Now we use some simple trigonometry. Recall that

$$\sin(A - B) = \sin A \cos B - \sin B \cos A$$

and that in a right triangle

$$\sin A = \frac{\text{opposite}}{\text{hypotenuse}} \qquad \text{and} \qquad \cos A = \frac{\text{adjacent}}{\text{hypotenuse}}.$$

Therefore,

$$\sin\theta = \frac{h}{\|\vec{v}\|}, \quad \sin\theta_u = \frac{u_2}{\|\vec{u}\|}, \quad \cos\theta_u = \frac{u_1}{\|\vec{u}\|}, \quad \sin\theta_v = \frac{v_2}{\|\vec{v}\|} \quad \text{and} \quad \cos\theta_v = \frac{v_1}{\|\vec{v}\|}.$$

We can now express the area of the parallelogram in terms of the entries of $\vec{u}$ and $\vec{v}$.

$$\begin{aligned} Area = \|\vec{u}\|h &= \|\vec{u}\|\|\vec{v}\|\sin\theta = \|\vec{u}\|\|\vec{v}\|\sin(\theta_v - \theta_u) \\ &= \|\vec{u}\|\|\vec{v}\|(\sin\theta_v\cos\theta_u - \sin\theta_u\cos\theta_v) \\ &= \|\vec{u}\|\|\vec{v}\|\left(\frac{v_2}{\|\vec{v}\|}\frac{u_1}{\|\vec{u}\|} - \frac{u_2}{\|\vec{u}\|}\frac{v_1}{\|\vec{v}\|}\right) \\ &= u_1v_2 - v_1u_2 \end{aligned}$$

This geometric derivation motivates the following definition.

**Definition 1.** *Given a* $2 \times 2$ *matrix* $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *we define the* **determinant** *of* $M$*, denoted* $\det(M)$*, as*

$$\det(M) = ad - bc.$$

In the example above, the determinant of the matrix is equal to the area of the parallelogram formed by the columns of the matrix. This is always the case up to a negative sign. Take for

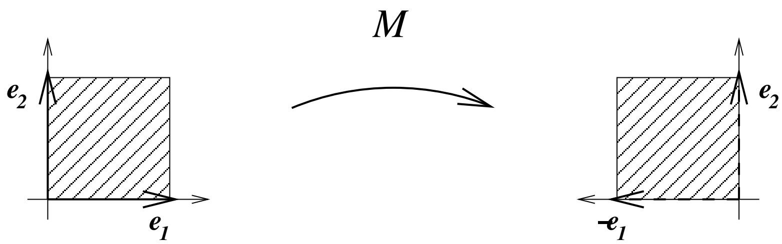



Figure 3: The action of $M$ on the unit square reverses orientation.

example $M = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. $M\vec{e}_1 = -\vec{e}_1$ and $M\vec{e}_2 = \vec{e}_2$. The action of $M$ on the unit square is depicted in Figure 3.
The area of the region is still clearly 1, but $\det(M) = -1(1) - 0(0) = -1$. This is because the determinant reflects the fact that the region has been "flipped", i.e. the orientation of the vectors describing the original parallelogram has been reversed in the image. Generally, we have $\det(M) = \pm Area$, where the determinant is positive if orientation is preserved and negative if it is reversed. Thus $\det(M)$ represents the *signed volume* of the parallelogram formed by the columns of $M$.

## 2 Properties of the Determinant

The convenience of the determinant of an $n \times n$ matrix is not so much in its formula as in the properties it possesses. In fact, the formula for $n > 2$ is quite complicated and any attempt to calculate it as we did for $n = 2$ from geometric principles is cumbersome. Rather than focus on the formula, we instead define the determinant in terms of three intuitive properties that we would like volume to have. It is an amazing fact that these three properties alone are enough to uniquely define the determinant.

### 2.1 Defining the Determinant in Terms of its Properties

We seek a function $D : \mathbb{R}^{n \times n} \to \mathbb{R}$ which assigns to each $n \times n$ matrix a single number. We adopt a flexible notation: $D$ is a function of a matrix so we write $D(A)$ to represent the number that $D$ assigns to the matrix $A$. However, it is also convenient to think of $D$ as a function of the columns of $A$ and so we write $D(A) = D(\vec{a}_1, \vec{a}_2, \dots \vec{a}_n)$ where $\vec{a}_1, \vec{a}_2, \dots \vec{a}_n$ are the columns of the matrix $A$.

Motivated by our intuitive ideas of volume, we require that the function $D$ have the following three properties:

**Property 1.** $D(I) = 1$.

This can also be written as $D(\vec{e}_1, \vec{e}_2, \dots \vec{e}_n) = 1$ since $\vec{e}_1, \vec{e}_2, \dots \vec{e}_n$ are the columns of the identity matrix $I$. These vectors describe the unit cube in $\mathbb{R}^n$ which should have volume 1.

**Property 2.** $D(\vec{a}_1, \vec{a}_2, \dots \vec{a}_n) = 0$ *if* $\vec{a}_i = \vec{a}_j$ *for some* $i \neq j$.

This condition says that if two edges of the parallelepiped are the same, then the parallelepiped is degenerate (i.e. "flat" in $\mathbb{R}^n$) and so should have volume zero.

**Property 3.** *If* $n - 1$ *columns are held fixed, then* $D$ *is a linear function of the remaining entry.*

Stated in terms of the $j^{th}$ column of the matrix, this property says that

$$\begin{aligned} D(\vec{a}_1, \ldots \vec{a}_{j-1}, \vec{u} + c\vec{v}, \vec{a}_{j+1}, \ldots \vec{a}_n) &= D(\vec{a}_1, \ldots \vec{a}_{j-1}, \vec{u}, \vec{a}_{j+1}, \ldots \vec{a}_n) \\ &\quad + cD(\vec{a}_1, \ldots \vec{a}_{j-1}, \vec{v}, \vec{a}_{j+1}, \ldots \vec{a}_n) \end{aligned}$$

so that $D$ is a linear function of the $j^{th}$ column when the other columns are held fixed. Note, this does *not* mean that $D(A + B) = D(A) + D(B)$! This is false!

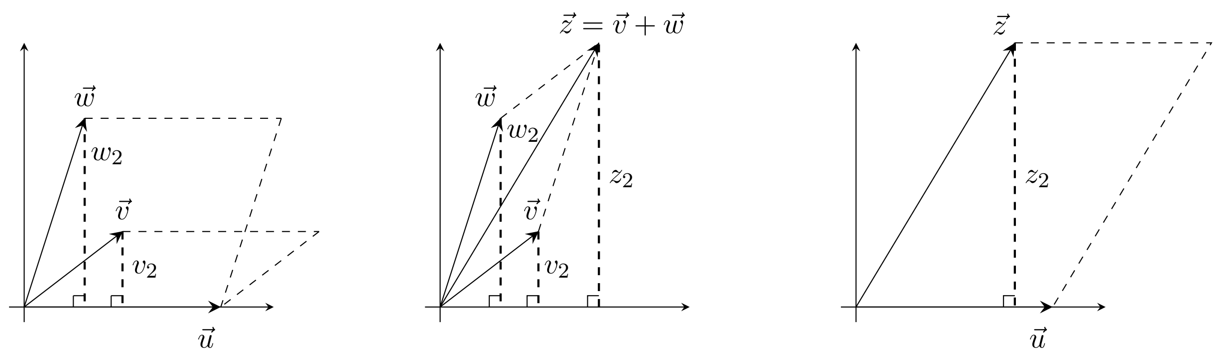
Property (3) reflects the way volumes add. This is best illustrated with a simple example. Let $\vec{u}$, $\vec{v}$ and $\vec{w}$ be vectors in $\mathbb{R}^2$ and let $A_{\vec{x},\vec{y}}$ denote the area of the parallelogram generated by $\vec{x}$ and $\vec{y}$. According to Property (3),

$$D\left(\begin{bmatrix} u_1 & v_1 + w_1 \\ u_2 & v_2 + w_2 \end{bmatrix}\right) = D\left(\begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix}\right) + D\left(\begin{bmatrix} u_1 & w_1 \\ u_2 & w_2 \end{bmatrix}\right).$$

In terms of areas, this would mean that

$$A_{\vec{u},\vec{v}+\vec{w}} = A_{\vec{u},\vec{v}} + A_{\vec{u},\vec{w}}.$$

To see that the areas actually behave in this way, we draw a diagram. Without loss of generality, we may assume that $\vec{u}$ lies along the positive $x$-axis. We let $\vec{z} = \vec{v} + \vec{w}$.




$A_{\vec{u},\vec{v}} = u_1 v_2$
$A_{\vec{u},\vec{w}} = u_1 w_2$

$z_2 = v_2 + w_2$

$A_{\vec{u},\vec{z}} = u_1 z_2 = u_1(v_2 + w_2)$

It is clear from the diagram that $A_{\vec{u},\vec{v}+\vec{w}} = A_{\vec{u},\vec{v}} + A_{\vec{u},\vec{w}}$: the bases of the parallelograms are the same and the altitude of the parallelogram formed by $\vec{u}$ and $\vec{z}$ is simply the sum of the altitudes of the parallelograms formed by $\vec{u}$ and $\vec{v}$ and by $\vec{u}$ and $\vec{w}$. Property (3) is a direct consequence of this observation about the additive properties of volume.

## 2.2 Additional Properties of the Determinant

Our goal is to show that the three properties stated in Section 2.1 actually determine a specific formula for $D$ in terms of the entries of a given matrix so that there can be only one function $D : \mathbb{R}^{n\times n} \to \mathbb{R}$ with these three properties. This function we will define as the determinant. In this section we formulate some of the consequences of Properties (1)-(3) as additional properties which will be crucial in deriving the formula for the determinant.

**Property 4.** *$D$ is an* **alternating** *function of the columns, i.e. if two columns are interchanged, the value of $D$ changes by a factor of -1.*

*Proof.* Let's say we interchange columns $\vec{a}_i$ and $\vec{a}_j$ of the matrix $A$. We keep the notation simple by writing $D(\vec{a}_i, \vec{a}_j)$ instead of $D(\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n)$ since these are the two entries we will be concerned with. The other entries remain constant.

$$\begin{aligned}
D(\vec{a}_i, \vec{a}_j) &= D(\vec{a}_i, \vec{a}_j) + D(\vec{a}_i, \vec{a}_i) && \text{by Property (2)} \\
&= D(\vec{a}_i, \vec{a}_j + \vec{a}_i) && \text{by Property (3)} \\
&= D(\vec{a}_i, \vec{a}_j + \vec{a}_i) - D(\vec{a}_j + \vec{a}_i, \vec{a}_j + \vec{a}_i) && \text{by Property (2)} \\
&= D(-\vec{a}_j, \vec{a}_j + \vec{a}_i) && \text{by Property (3)} \\
&= -D(\vec{a}_j, \vec{a}_j + \vec{a}_i) && \text{by Property (3)} \\
&= -D(\vec{a}_j, \vec{a}_j) - D(\vec{a}_j, \vec{a}_i) && \text{by Property (3)} \\
&= -D(\vec{a}_j, \vec{a}_i) && \text{by Property (2)}
\end{aligned}$$

□

**Property 5.** *If $\{\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n\}$ is a linearly dependent set of vectors, then $D(\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n) = 0$.*

*Proof.* If the vectors are linearly dependent then one of them can be written as a linear combination of the others. Without loss of generality, let's say that vector is $\vec{a}_1$.

$$\vec{a}_1 = c_2\vec{a}_2 + \ldots + c_n\vec{a}_n$$

Then using the fact that $D$ is a linear function of one column when the others are held fixed (Property (3)), we have

$$\begin{aligned}
D(\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n) &= D(c_2\vec{a}_2 + \ldots + c_n\vec{a}_n, \vec{a}_2, \ldots \vec{a}_n) \\
&= c_2 D(\vec{a}_2, \vec{a}_2, \ldots \vec{a}_n) + c_3 D(\vec{a}_3, \vec{a}_2, \ldots \vec{a}_n) + \ldots + c_n D(\vec{a}_n, \vec{a}_2, \ldots \vec{a}_n)
\end{aligned}$$

Note that every term in the last line is zero by Property (2). □

An immediate consequence of Property (5) is *the fact that a non-invertible matrix must have determinant equal to zero.* This is because the columns of a non-invertible matrix are linearly dependent and so $D$ is forced to be zero by Property (5).

**Property 6.** *Adding a multiple of one column to another does not change the determinant.*

*Proof.* Suppose the matrix $B$ is obtained from $A$ by adding $c$ times column $j$ to column $i$. Then

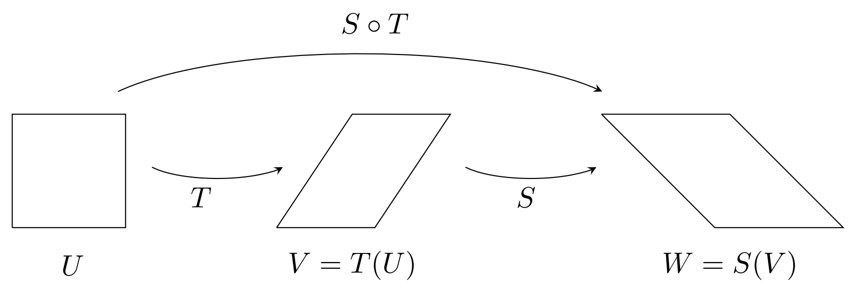
$$\begin{aligned}
D(B) &= D(\vec{a}_1, \ldots \vec{a}_{i-1}, \vec{a}_i + c\vec{a}_j, \vec{a}_{i+1}, \ldots \vec{a}_n) \\
&= D(\vec{a}_1, \ldots \vec{a}_{i-1}, \vec{a}_i, \vec{a}_{i+1}, \ldots \vec{a}_n) + cD(\vec{a}_1, \ldots \vec{a}_{i-1}, \vec{a}_j, \vec{a}_{i+1}, \ldots \vec{a}_n) \\
&= D(A) \qquad \text{since the second term is 0.}
\end{aligned}$$

□

There is one further property of determinants which is very convenient.

**Theorem 2.1.** *If $A$ and $B$ are $n \times n$ matrices, then $D(AB) = D(A)D(B)$.*

We could prove this formally, but it is more instructive to see why this is true by regarding $AB$ as the composition of two linear transformations. If $S$ is the linear transformation which multiplies vectors by $A$ and $T$ is the linear transformation which multiplies vectors by $B$, then the composition $S \circ T$ multiplies vectors by the matrix $AB$. To see what the determinant of $AB$ must be, we need only look at how $S \circ T$ changes volumes.




$T$ transforms volumes by a factor of $D(B)$ and $S$ transforms volumes by a factor of $D(A)$. This means that if $T(U) = V$, then $\text{vol}(V) = D(B)\text{vol}(U)$ where $\text{vol}(X)$ is the signed volume of the set $X$. Similarly, if $S(V) = W$, then $\text{vol}(W) = D(A)\text{vol}(V)$. Putting these together,

$$\text{vol}(W) = D(A)\text{vol}(V) = D(A)D(B)\text{vol}(U)$$

so that the transformation $S \circ T$ changes volumes by a factor of $D(A)D(B)$. This is precisely what $D(AB)$ represents. So $D(AB) = D(A)D(B)$.

### 2.3 Checking the 2 × 2 Determinant

Before deriving formulas for computing the determinant of an $n \times n$ matrix, let's check that the determinant of a $2 \times 2$ matrix we motivated geometrically in Section 1 satisfies the three properties we postulated in Section 2.1. Once we check that it has these three properties, we conclude that it also satisfies all the additional properties of Section 2.2 since these were proved on the basis of the first three.

Recall that we defined the determinant of a $2 \times 2$ matrix $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ by $\det(M) = ad - bc$. At this point we introduce the notation that the determinant of a matrix can also be expressed as the matrix array with absolute value bars instead of square brackets.

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)$$

**Property 1** Check that $\det(I) = 1$.

$$\det(I) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \cdot 1 - 0 \cdot 0 = 1$$

**Property 2** Check that $\det(M) = 0$ if two columns are the same.

$$\det(M) = \begin{vmatrix} a & a \\ b & b \end{vmatrix} = a \cdot b - a \cdot b = 0$$

**Property 3** Check that if 1 column is held fixed, then the determinant is a linear function of the remaining column.
Let's hold the second column fixed and put a linear combination of vectors in the first column.

Suppose the two columns of $M$ are $\vec{u} + c\vec{v}$ and $\vec{a}$.

$$\begin{aligned}
\det(\vec{u} + c\vec{v}, \vec{a}) &= \begin{vmatrix} u_1 + cv_1 & a_1 \\ u_2 + cv_2 & a_2 \end{vmatrix} = (u_1 + cv_1)a_2 - a_1(u_2 + cv_2) \\
&= u_1a_2 + cv_1a_2 - a_1u_2 - a_1cv_2 = u_1a_2 - a_1u_2 + c(v_1a_2 - a_1v_2) \\
&= \begin{vmatrix} u_1 & a_1 \\ u_2 & a_2 \end{vmatrix} + c \begin{vmatrix} v_1 & a_1 \\ v_2 & a_2 \end{vmatrix} \\
&= \det(\vec{u}, \vec{a}) + c\det(\vec{v}, \vec{a})
\end{aligned}$$

So we see that the determinant is a linear function of the first column when the second column is held fixed. The proof for the second column is entirely similar so we omit it.

# 3 Determinants of Special Matrices and a Criterion for Invertibility

In this section, we compute the determinants of certain types of matrices which we will then use to derive general formulas for the determinant of an $n \times n$ matrix. The formulas provided will in fact be derived *directly* from the properties (1)-(3) that we've required the determinant to possess.

**Proposition 3.1.** *The determinant of a triangular matrix is the product of its diagonal entries.*

*Proof.* To fix ideas consider first the case $n = 3$. Let $A$ be the upper triangular matrix given by

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}.$$

Applying Property (3) to column 2 while holding the other columns fixed yields

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix}.$$

Notice that the first term on the right is zero since the first and second columns are multiples of one another. Now we apply Property (3) to column 3 while holding the others fixed to obtain

$$\begin{vmatrix} a_{11} & 0 & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & a_{13} \\ 0 & a_{22} & 0 \\ 0 & 0 & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & a_{23} \\ 0 & 0 & 0 \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix}.$$

In this expression, the first two terms on the right are both zero, leaving only the third. Putting these expansions together yields

$$D(A) = \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = a_{11}a_{22}a_{33}$$

as desired.

It is clear that this method of expanding successive columns generalizes to any dimension by induction. Given an $n \times n$ upper triangular matrix, assume we have expanded the first $j-1$ columns as above and discarded the terms which are zero. Expanding the $j^{\text{th}}$ column, we see that every term whose non-zero entry in column $j$ lies in a row above the diagonal is a multiple of a previous column in the matrix, making that determinant zero. Only the matrix with the entry in the $j^{\text{th}}$ column and $j^{\text{th}}$ row survives.

The argument is similar for a lower triangular matrix. □

**Proposition 3.2.** *Let $E$ be an $n \times n$ elementary matrix. Then*

*(a) if $E$ adds $c$ times row $i$ to row $j$, then $D(E) = 1$;*

*(b) if $E$ scales row $i$ by a factor of $k$, then $D(E) = k$;*

*(c) if $E$ interchanges two rows, then $D(E) = -1$.*

*Proof.* If $E$ is of type (a) in the statement above, then $E$ is a triangular matrix with 1's on the diagonal. By Proposition 3.1, $D(E) = 1$.

If $E$ is of type (b) in the statement above, then $E$ is a triangular matrix with 1's on the diagonal except for row $i$ which has diagonal entry $k$. Again using Proposition 3.1, $D(E) = k$.

If $E$ is of type (c) in the statement above, then $E$ is the identity matrix $I$, except with rows $i$ and $j$ interchanged. This is equivalent to starting with I and interchanging columns $i$ and $j$, so we may use Property 4 to conclude that $D(E) = -1$. □

Note that for an elementary matrix $E$, $E^T$ is an elementary matrix of the same type. Proposition 3.2 implies that $D(E) = D(E^T)$. This yields the next imporant property of determinants.

**Property 7.** $D(A) = D(A^T)$.

*Proof.* If $A$ is not invertible, then $D(A) = 0$ and the echelon form of $A$ has a zero row, i.e. the rows of $A$ are linearly dependent. This implies that the columns of $A^T$ are linearly dependent, so $D(A^T) = 0$ by Property 5.

If $A$ is invertible, then $A$ can be written as a product of elementary matrices: $A = E_k \cdots E_2 E_1$. Then using Theorem 1 to break up the product, we have

$$\begin{aligned} D(A) &= D(E_k \cdots E1) = D(E_k) \cdots D(E_1) \\ &= D(E_k^T) \cdots D(E_1^T) = D(E_1^T) \cdots D(E_k^T) \\ &= D(E_1^T \cdots E_k^T) = D((E_k \cdots E_1)^T) \ = \ D(A^T)\,. \end{aligned}$$

□

The importance of Property 7 is that it allows us to conclude that all the properties that we have stated for columns also work for rows. We see this because the columns of $A^T$ are the rows of $A$. We highlight these additional properties, letting $\vec{r}_1, \ldots, \vec{r}_n$ denote the rows of $A$.

**Property 2′** If $\vec{r}_i = \vec{r}_j$ for some $i \neq j$, then $D(A) = 0$.

**Property 3′** $D$ depends linearly on each row $\vec{r}_i$ keeping the remaining $n-1$ rows fixed.

**Property 4′** If two rows are interchanged, then $D$ changes by a factor of $-1$.

**Property 5′** If $\{\vec{r}_1, \vec{r}_2, \ldots \vec{r}_n\}$ is a linearly dependent set, then $D(A) = 0$.

**Property 6′** Adding a multiple of one row to another does not change the determinant.

What these properties crystallize for us is the way in which elementary row operations affect the determinant of a matrix. This knowledge is important for both proving further results and shortening the computation of the determinant. (Computing the determinant of an $n \times n$ matrix directly takes $n!$ calculations. In order to reduce the number of calculations, computer programs first row reduce the matrix to echelon form using the above rules before computing the determinant. The triangular form of a matrix in echelon form makes the determinant easy to compute.)

These results have two immediate consequences for general $n \times n$ matrices: the first is a formula for the determinant of a matrix; the second is a criterion for the invertibility of a matrix.

**Corollary 3.3.** *Suppose a square matrix $A$ is reduced to echelon form using only two types of row operations: the interchange of rows and the addition of a constant multiple of one row to another (i.e. no scaling of rows is used). Then*

$$D(A) = \pm(\textit{product of the pivots in echelon form}).$$

*Proof.* Let $U$ be an echelon form of $A$ obtained by performing only the row operations stated above. $U$ is an upper triangular matrix so by Proposition 3.1,

$$D(U) = (\text{product of diagonal entries}) = (\text{product of pivots}).$$

$U$ was obtained from $A$ by additions of a constant multiple of one row to another, which do not change the determinant, and by row interchanges, which only change the sign of the determinant. So $D(A)$ must agree with $D(U)$ up to its sign. □

Corollary 3.3 immediately gives us a criterion for invertibility. If a matrix $A$ is invertible, any echelon form of $A$ has no zero pivots and so $D(A)$ cannot be zero by Corollary 3.3. On the other hand, if $A$ is not invertible, then it has at least one zero pivot and so $D(A) = 0$. (We already knew that the determinant of a noninvertible matrix is zero by Property (5) since the columns of the matrix are linearly dependent. What Corollary 3.3 establishes is the fact that an invertible matrix has a non-zero determinant.)

**Corollary 3.4.** *A square matrix $A$ is invertible if and only if $D(A) \neq 0$.*

## 4 A Formula for the Determinant

At this point we seek an explicit formula for the determinant of an $n \times n$ matrix for $n > 2$. The formula will make it easier for us to get our hands dirty computing determinants. It follows directly from the properties proved so far. First we establish two convenient facts.

Let $E$ be an $n \times n$ elementary matrix and let $E'$ be the $(n+1) \times (n+1)$ matrix given by

$$E' = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & E & \\ 0 & & & \end{bmatrix}.$$

**Lemma 4.1.** *Let $E$ and $E'$ be as above. Then $D(E) = D(E')$.*

*Proof.* The lemma follows from Proposition 3.2 and the observation that $E'$ is also an elementary matrix which represents a row operation of the same type as $E$. □

Before proceeding, we will need some new notation. Given an $n \times n$ matrix $A$, we denote by $A_{ij}$ the matrix obtained by deleting the $i^{th}$ row and $j^{th}$ column of $A$.

For example, if $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ then

$$A_{11} = \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} \quad A_{32} = \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix} \quad A_{13} = \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}.$$

**Lemma 4.2.** *Consider an $n \times n$ matrix $A$ whose first column is $a_{11} \cdot \vec{e}_1$. Then*

$$D(A) = a_{11} D(A_{11}).$$

*Proof.* First, using Property $3'$ we expand along the first row of $A$:

$$\begin{aligned} |A| &= \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A_{11} & \\ 0 & & & \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \dots & 0 \\ 0 & & & \\ \vdots & & A_{11} & \\ 0 & & & \end{vmatrix} + \dots + \begin{vmatrix} 0 & 0 & \dots 0 & a_{1n} \\ 0 & & & \\ \vdots & & A_{11} & \\ 0 & & & \end{vmatrix} \\ &= a_{11} \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A_{11} & \\ 0 & & & \end{vmatrix}. \end{aligned} \tag{1}$$

In the first line, every term but the first on the right hand side is zero since the first column of each matrix is the zero vector. Let $B$ be the matrix remaining on the right hand side of equation (1).

Suppose $A_{11}$ is not invertible. Then columns 2 through $n$ of $B$ are linearly dependent and therefore $D(A) = D(A_{11}) = 0$.

Now suppose $A_{11}$ is invertible. Then $A_{11} = E_k \cdots E_1 I_{n-1}$ for some sequence of elementary matrices. For each $j$, let $E'_j$ be the elementary matrix given by

$$E'_j = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & E_j & \\ 0 & & & \end{bmatrix}.$$

Since the first row and column of $B$ are $\vec{e}_1$ and $B_{11} = A_{11}$, the sequence of row operations analogous to that which transforms $I_{n-1}$ into $A_{11}$ will transform $I_n$ into $B$. Thus $B = E'_k \cdots E'_1 I_n$. Using Lemma 4.1 we conclude that

$$D(A) = a_{11} D(B) = a_{11} D(E'_k) \cdots D(E'_1) = a_{11} D(E_k) \cdots D(E_1) = a_{11} D(A_{11}).$$

□

Using this result, the formula for the determinant follows quickly. Consider a $3 \times 3$ matrix $A$. Using the linearity of Property (3) on the first column, we write

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} - \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ 0 & a_{12} & a_{13} \\ 0 & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{31} & a_{32} & a_{33} \\ 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \end{vmatrix}. \end{aligned} \tag{2}$$

In the second line we have brought the rows with the non-zero entries in the first column to the top row while maintaining the relative ordering of the other rows. Since the determinant is an alternating function of the rows, each exchange introduces a factor of $-1$. Notice that in the first term of equation (2), the $2 \times 2$ matrix obtained by deleting the first row and column of the matrix is $A_{11}$. In the second term, the matrix obtained by deleting the first row and column is $A_{21}$. In the third term, the matrix obtained by deleting the first row and column is $A_{31}$. Using Proposition 4.2 to evaluate each determinant, we obtain

$$D(A) = a_{11}D(A_{11}) - a_{21}D(A_{21}) + a_{31}D(A_{31}).$$

For a general $n \times n$ matrix $A$, we can do the same expansion: expand the determinant into the sum of $n$ determinants — the first column of the $i^{th}$ term will have $a_{i1}$ in the $i^{th}$ row and zeroes elsewhere; then we use $i-1$ transpositions to move the $i^{th}$ row to the first row while maintaining the relative ordering of the other rows; finally, we use Lemma 4.2 to evaluate the determinant of each matrix. In this way we arrive at the following formula for the determinant of a matrix.

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

This is called the **cofactor expansion along the first column** of $A$.

The importance of this formula is that it allows us to define the determinant of a matrix inductively: we define the determinant of an $n \times n$ matrix in terms of the determinants of $(n-1) \times (n-1)$ matrices. We begin by defining the determinant of a $1 \times 1$ matrix $A = [a]$ by $\det(A) = a$. Then we proceed to two dimensions and so on using the following definition.

**Definition 2.** *Let $A$ be an $n \times n$ matrix. Define the* **determinant of A** *to be*

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

In fact, we can do a similar expansion along any row or column of $A$, simply by following the procedure outlined above for expanding the determinant and then moving each position in the chosen row or column into the upper left corner of the matrix through a sequence of row and column exchanges. In order to move the entry in the $i^{th}$ row and $j^{th}$ column of the matrix into the upper left hand corner and preserve the order of the other rows and columns, we need $i-1$ row exchanges and $j-1$ column exchanges. Each exchange introduces a factor of $-1$ to the determinant of that term.

In this way we obtain additional formulas for the determinant. If we choose to expand along the $i^{th}$ row of $A$, we have

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j-2} a_{ij} \det(A_{ij}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$

This is called the **cofactor expansion along the $\mathbf{i^{th}}$ row** of $A$. If we choose to expand along the $j^{th}$ column of $A$, we have

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

This is called the **cofactor expansion along the $\mathbf{j^{th}}$ column** of $A$.

The term "cofactor expansion" stems from the the fact that the quantity $C_{ij} = (-1)^{i+j}\det(A_{ij})$ is called the **(i, j)-cofactor** of the matrix $A$.

Note that if $A$ is $2\times 2$, then this formula for the determinant agrees with the one we motivated geometrically in Section 1 from the area of a parallelogram. Setting $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, we see that $A_{11} = [a_{22}]$ and $A_{12} = [a_{21}]$ so expanding along the first row,

$$\begin{aligned}
\det(A) &= \sum_{j=1}^{2}(-1)^{1+j}a_{1j}\det(A_{1j}) \\
&= (-1)^{1+1}a_{11}\det(A_{11}) + (-1)^{1+2}a_{12}\det(A_{12}) \\
&= a_{11}\cdot a_{22} - a_{12}\cdot a_{21}.
\end{aligned}$$

Let's practice using the cofactor expansions to compute the determinants of some matrices.

**Example 1.** Let $A = \begin{bmatrix} 1 & 3 & -3 \\ -3 & -5 & 2 \\ -4 & 4 & -6 \end{bmatrix}$. We expand along the first row.

$$\begin{aligned}
\det(A) &= (-1)^{1+1}a_{11}\det(A_{11}) + (-1)^{1+2}a_{12}\det(A_{12}) + (-1)^{1+3}a_{13}\det(A_{13}) \\
&= (-1)^2(1)\begin{vmatrix} -5 & 2 \\ 4 & -6 \end{vmatrix} + (-1)^3(3)\begin{vmatrix} -3 & 2 \\ -4 & -6 \end{vmatrix} + (-1)^4(-3)\begin{vmatrix} -3 & -5 \\ -4 & 4 \end{vmatrix} \\
&= (30-8) - 3(18+8) + (-3)(-12-20) \\
&= 22 - 78 + 96 = 40
\end{aligned}$$

**Example 2.** Let $B = \begin{bmatrix} 0 & 1 & 3 \\ -2 & -3 & -5 \\ 4 & -4 & 4 \end{bmatrix}$. We expand along the first column.

$$\begin{aligned}
\det(B) &= (-1)^{1+1}b_{11}\det(B_{11}) + (-1)^{2+1}b_{21}\det(B_{21}) + (-1)^{3+1}b_{31}\det(B_{31}) \\
&= (-1)^2(0)\begin{vmatrix} -3 & -5 \\ -4 & 4 \end{vmatrix} + (-1)^3(-2)\begin{vmatrix} 1 & 3 \\ -4 & 4 \end{vmatrix} + (-1)^4(4)\begin{vmatrix} 1 & 3 \\ -3 & -5 \end{vmatrix} \\
&= 0 - (-2)(4+12) + (4)(-5+9) \\
&= 32 + 16 = 48
\end{aligned}$$

**Example 3.** Let $C = \begin{bmatrix} -2 & -3 & -5 & 2 \\ 0 & 1 & 3 & -3 \\ 2 & 0 & 0 & 1 \\ 4 & -4 & 4 & -6 \end{bmatrix}$. We expand along the third row.

$$\begin{aligned}
\det(C) &= (-1)^{3+1}c_{31}\det(C_{31}) + (-1)^{3+2}c_{32}\det(C_{32}) + (-1)^{3+3}c_{33}\det(C_{33}) \\
&\quad + (-1)^{3+4}c_{34}\det(C_{34}) \\
&= (-1)^4(2)\det(A) + (-1)^5(0)\det(C_{32}) + (-1)^6(0)\det(C_{33}) + (-1)^7(1)\det(B) \\
&= 2(40) + 0 + 0 - 48 = 32
\end{aligned}$$