Determinants

> 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 2×2 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 2×2 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 $M = [\vec{u} \quad \vec{v}]$. To examine how M transforms areas, we look at the action of M on $\vec{e}_1$ and $\vec{e}_2$. $M\vec{e}_1 = \vec{u}$ 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 $\vec{v}$.

We wish to find the area of the parallelogram determined by $\vec{u}$ and $\vec{v}$.

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

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

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* 2×2 *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 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 below.

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 $D(A) = D(\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n)$ where $\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n$ 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 $D(\vec{e}_1, \vec{e}_2, \ldots \vec{e}_n) = 1$ since $\vec{e}_1, \vec{e}_2, \ldots \vec{e}_n$ 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 $D(\vec{a}_1, \vec{a}_2, \ldots \vec{a}_n) = 0$ *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 $\vec{z} = \vec{v} + \vec{w}$.

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 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 2×2 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 2×2 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 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 first formula provided will in fact be derived *directly* from the properties (1)-(3) that we've required the determinant to possess.

To fix ideas consider first the case $n = 2$. Let

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

Applying Property (3) on column 1 holding column 2 fixed,

$$D\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = D\left(\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}\right) + D\left(\begin{bmatrix} 0 & b \\ c & d \end{bmatrix}\right).$$

Again using Property (3), now on column 2 holding column 1 fixed in each of the two terms,

$$D\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = D\left(\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}\right) + D\left(\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}\right) + D\left(\begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix}\right) + D\left(\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}\right).$$

Now the columns in the first and last terms are linearly dependent so by Property (5),

$$D\left(\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}\right) = 0 \text{ and } D\left(\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}\right) = 0.$$

Together,

$$D\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a \cdot d \cdot D\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) + b \cdot c \cdot D\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right). \tag{1}$$

Let's leave this expression as it stands and continue on to $n = 3$, with

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

Notice that applying Property (3) on each column will yield $3 \cdot 3 \cdot 3 = 27$ terms. However, the terms that have entries in different columns but the same row are terms whose columns are linearly dependent and hence have determinant zero by Property (5). For example,

$$D\left(\begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & 0 \\ a_{31} & 0 & 0 \end{bmatrix}\right) = 0.$$

Eliminating all such terms, the only surviving six terms are

$$D\left(\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}\right), \quad D\left(\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{bmatrix}\right), \quad D\left(\begin{bmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{bmatrix}\right),$$

$$D\left(\begin{bmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{bmatrix}\right),\quad D\left(\begin{bmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{bmatrix}\right),\quad D\left(\begin{bmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & 0 & 0 \end{bmatrix}\right).$$

So,

$$\begin{aligned} D(A) \;=\; & a_{11} \cdot a_{22} \cdot a_{33} \cdot D\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) + a_{11} \cdot a_{32} \cdot a_{23} \cdot D\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\right) \\ & + a_{21} \cdot a_{12} \cdot a_{33} \cdot D\left(\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) + a_{21} \cdot a_{32} \cdot a_{13} \cdot D\left(\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\right) \\ & + a_{31} \cdot a_{12} \cdot a_{23} \cdot D\left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\right) + a_{31} \cdot a_{22} \cdot a_{13} \cdot D\left(\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\right)). \end{aligned} \tag{2}$$

To generalize to $n > 3$, it will be convenient to introduce the concept of **permutations**.

4 Permutations

Definition 2 *A* ***permutation*** *of the set* $\{1, \ldots, n\}$ *is a bijective function that maps the set* $\{1, \ldots, n\}$ *onto itself. Denote by* S_n *the set of all permutations of* $\{1, \ldots, n\}$*. Then the number of elements in* S_n *is* $n!$.

To see that the number of elements in S_n is $n!$, consider what $\sigma \in S_n$ looks like. $\sigma(1)$ can be any of the n integers. $\sigma(2)$ can be any one of the elements of this set except $\sigma(1)$ since σ is one-to-one. There are $n-1$ choices for $\sigma(2)$. Similarly, $\sigma(3)$ can be any of the elements except $\sigma(1)$ and $\sigma(2)$ and so there are $n-2$ choices for $\sigma(3)$. This process continues so the total number of permutations of $\{1, \ldots, n\}$ is $n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1 = n!$.

In what follows, we will denote a permutation σ as an ordered list : $(\sigma(1), \sigma(2), \ldots, \sigma(n))$. The elements of S_3, for example, are $(1,2,3)$, $(1,3,2)$, $(2,1,3)$, $(2,3,1)$, $(3,1,2)$, and $(3,2,1)$. The simplest permutation is the **identity** permutation $\epsilon = (1, 2, \ldots, n)$. Another class of simple permutations are **transpositions**, which exchange two elements in $\{1, \ldots, n\}$ and fix the remaining elements. The transposition that exchanges elements i and j will be denoted $\tau_{i,j}$. In S_5, $\tau_{2,4} = (1, \boxed{4}, 3, \boxed{2}, 5)$ is an example of a transposition, which exchanges 2 and 4.

Permutations can be composed and their composition is a permutation. We will be looking especially at the composition of a permutation with a transposition. Take a transposition $\tau_{i,j}$ and an arbitrary permutation σ. Then the composition $\tau_{i,j} \circ \sigma$ results in the permutation that exchanges elements i and j in the representation of σ. For example, take the permutation $\sigma = (2,4,3,1)$ and the transposition $\tau_{2,3} = (1,3,2,4)$.

Then $\tau_{2,3} \circ \sigma = (3, 4, 2, 1)$ as seen below.

$$\begin{array}{ccccc} & \sigma & & \tau_{2,3} & \\ 1 & \rightarrow & 2 & \rightarrow & 3 \\ 2 & \rightarrow & 4 & \rightarrow & 4 \\ 3 & \rightarrow & 3 & \rightarrow & 2 \\ 4 & \rightarrow & 1 & \rightarrow & 1 \end{array}$$

It is not hard to see that for any permutation, one can find a sequence of transpositions which will transform that permutation to the identity. For the same σ above, one possible sequence of transpositions that results in the identity is the following:

$$\begin{aligned} \sigma &= (2, 4, 3, 1) \\ \tau_{1,2} \circ \sigma &= (1, 4, 3, 2) \\ \tau_{2,3} \circ \tau_{1,2} \circ \sigma &= (1, 4, 2, 3) \\ \tau_{2,4} \circ \tau_{2,3} \circ \tau_{1,2} \circ \sigma &= (1, 2, 4, 3) \\ \tau_{3,4} \circ \tau_{2,4} \circ \tau_{2,3} \circ \tau_{1,2} \circ \sigma &= (1, 2, 3, 4) \end{aligned}$$

Notice that the sequence (and the number) of compositions that achieve this task is not unique. Instead of the first sequence, we can devise a different sequence by performing the following sequence of exchanges:

$$\begin{aligned} \sigma &= (2, 4, 3, 1) \\ \tau_{1,2} \circ \sigma &= (1, 4, 3, 2) \\ \tau_{2,3} \circ \tau_{1,2} \circ \sigma &= (1, 4, 2, 3) \\ \tau_{2,4} \circ \tau_{2,3} \circ \tau_{1,2} \circ \sigma &= (1, 2, 4, 3) \\ \tau_{2,4} \circ \tau_{2,4} \circ \tau_{2,3} \circ \tau_{1,2} \circ \sigma &= (1, 4, 2, 3) \\ \tau_{1,4} \circ \tau_{3,4} \circ \tau_{2,4} \circ \tau_{2,3} \circ \tau_{1,2} \circ \sigma &= (4, 1, 2, 3) \\ \tau_{3,4} \circ \tau_{1,4} \circ \tau_{3,4} \circ \tau_{2,4} \circ \tau_{2,3} \circ \tau_{1,2} \circ \sigma &= (3, 1, 2, 4) \\ \tau_{1,2} \circ \tau_{3,4} \circ \tau_{1,4} \circ \tau_{3,4} \circ \tau_{2,4} \circ \tau_{2,3} \circ \tau_{1,2} \circ \sigma &= (3, 2, 1, 4) \\ \tau_{1,3} \circ \tau_{1,2} \circ \tau_{3,4} \circ \tau_{1,4} \circ \tau_{3,4} \circ \tau_{2,4} \circ \tau_{2,3} \circ \tau_{1,2} \circ \sigma &= (1, 2, 3, 4) \end{aligned}$$

The *number* of transpositions necessary to transform a permutation to the identity is not unique: 4 are needed in the first example and 8 in the second. However, the *parity* of the number of exchanges turns out to be unique. In this example, the number of exchanges is always even. We now state this fundamental property.

Theorem 2 *For any $\sigma \in S_n$, the parity of the number of transpositions which transforms the permutation to the identity is well-defined.*

This means that for certain permutations, this number will always be even and for others, this number will always be odd. This theorem will not be proved here.

Definition 3 *The **sign** of a permutation is a function* sign: $S_n \rightarrow \{+1, -1\}$ *assigning* $+1$ *if an even number of transpositions is needed to transform the permutation to the identity and* -1 *if an odd number is needed.*

One last notation needs to be introduced. Associate with any $\sigma \in S_n$ the matrix E_σ that has a 1 in the entries $(\sigma(1), 1), (\sigma(2), 2) \ldots, (\sigma(n), n)$ and 0 elsewhere. Given $\sigma = (2, 4, 3, 1)$, the associated matrix E_σ is given by

$$E_\sigma = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

We wish to show that

Property 7 *$D(E_\sigma) = sign(\sigma)$.*

Proof. Note the relation between the columns of E_σ and σ:

$$E_\sigma = \left[\begin{array}{cccc} \vec{e}_{\sigma(1)} & \vec{e}_{\sigma(2)} & \cdots & \vec{e}_{\sigma(n)} \end{array} \right].$$

A sequence of transpositions which transform σ to the identity can be viewed as a sequence of column exchanges which transforms E_σ to I. By Property (4), each column exchange results in a factor of -1. $sign(\sigma)$ keeps track of the parity of the exchanges needed to transform σ to the identity permutation which is the parity of the column exchanges needed to transform E_σ to the identity matrix. So, $D(E_\sigma) = sign(\sigma)D(I) = sign(\sigma)$.

5 A Formula for the Determinant–Revisited

Glance back to the formula that we derived for the special cases $n = 2$, $n = 3$, given by Equations 1 and 2. The matrices involved in both are exactly E_σ, cycling through the permutations σ in S_2 and S_3, respectively. This observation, together with Property (7), allows us to write down a general expression for $D(A)$:

$$D(A) = \sum_{\sigma \in S_n} D(E_\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} = \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}.$$

We are now in a position to define the determinant of a general $n \times n$ matrix A.

Definition 4 *For any $n \times n$ matrix A, define the* ***determinant*** *of A to be the function $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$, given by*

$$\det(A) := \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}.$$

Remark. From this formula, we have that $\det(E_{\tilde{\sigma}}) = sign(\tilde{\sigma})$ for any $\tilde{\sigma} \in S_n$. This is because $(E_{\tilde{\sigma}})_{ij} = 1$, if $i = \tilde{\sigma}(j)$ and 0 otherwise. Therefore, $\prod_{i=1}^{n} (E_{\tilde{\sigma}})_{\sigma(i)i}$ will be 0 whenever $\sigma \neq \tilde{\sigma}$ and is 1 when $\sigma = \tilde{\sigma}$. So, $\det(E_{\tilde{\sigma}}) = sign(\tilde{\sigma})$.

This is the most transparent form from which to verify that Properties (1), (2) and (3) are satisfied. It should not be surprising that det satisfies these properties since we derived this formula from the properties themselves.

Property 1 Check that $det(I) = 1$. This is a special case of the remark, for $\epsilon = (1, 2, \ldots, n)$.

$$\det(I) = \det(E_\epsilon) = sign(\epsilon) = 1.$$

Property 2 Check that $\det(A) = 0$ if two columns are the same.
Let column $\vec{a}_i = \vec{a}_j$, for some $i \neq j$. For $\sigma \in S_n$, notice that the factors corresponding to any two permutations σ and $\sigma \circ \tau_{i,j}$ are the same.

$$\begin{array}{rcccccccc} \sigma & = & (\sigma(1), & \ldots, & \sigma(i), & \ldots, & \sigma(j), & \ldots, & \sigma(n)) \\ & & & & \uparrow & & \uparrow & & \\ & & & & i & & j & & \\ & & & & \downarrow & & \downarrow & & \\ \tilde{\sigma} = \sigma \circ \tau_{i,j} & = & (\sigma(1), & \ldots, & \sigma(j), & \ldots, & \sigma(i), & \ldots, & \sigma(n)) \end{array}$$

Since $sign(\tilde{\sigma}) = -sign(\sigma)$, these pairs cancel. It is possible to pair permutations this way; therefore, $\det(A) = 0$.

Property 3 Check that if $n-1$ columns of A are fixed then det is a linear function of the remaining column. Consider the j^{th} column $\vec{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$. Collect together all terms that involve a_{1j} and write these terms as $a_{1j}C_{1j}$. Similarly, collect together all terms involving a_{ij} and write these terms as $a_{ij}C_{ij}$. Then

$$\det(A) \quad = \quad a_{1j}C_{1j} + \ldots + a_{ij}C_{ij} + \ldots a_{nj}C_{nj}. \tag{3}$$

Since each C_{ij} contains no entries from column j, $\det(A)$ depends linearly on the j-th column holding the remaining $n-1$ columns fixed.

Since the function det satisfies Properties (1)-(3), it follows immediately that det also satisfies Properties (4)-(7) as well as Theorem 1. An important observation is that the derivation of this formula allows us to conclude that there is a *unique* function which satisfies Properties (1)-(3). One additional property can be shown using this formula.

Property 8 $\det(A^T) = \det(A)$.

Proof. Since $\sigma \in S_n$ is bijective, σ^{-1} exists and is also in S_n. Moreover, $sign(\sigma) = sign(\sigma^{-1})$.

$$\det(A^T) \quad = \quad \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} (A^T)_{\sigma(i)i}$$

$$
\begin{aligned}
&= \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \\
&= \sum_{\sigma \in S_n} sign(\sigma^{-1}) \prod_{i=1}^{n} a_{\sigma^{-1}(i)i} \\
&= \sum_{\tilde{\sigma}^{-1} \in S_n} sign(\tilde{\sigma}) \prod_{i=1}^{n} a_{\tilde{\sigma}(i)i} \\
&= \sum_{\tilde{\sigma} \in S_n} sign(\tilde{\sigma}) \prod_{i=1}^{n} a_{\tilde{\sigma}(i)i} \\
&= \det(A).
\end{aligned}
$$

□

From Property (8), 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 $\vec{r}_1^T, \ldots, \vec{r}_n^T$ denote the rows of A.

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

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

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

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

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 both for proving further results and to make the process of computing the determinant easier. (Computing the determinant of an $n \times n$ matrix directly takes $n!$ calculations. In order to reduce the number of calculations, computer programs are designed to row reduce the matrix to echelon form using the above rules before computing the determinant. We will see that the triangular form of a matrix in echelon form makes the determinant easy to compute.)

In light of these comments we can formulate a proposition in terms of elementary row operations performed on a matrix A. These are simply restatements of the properties already established.

Proposition 1 *(a) If B is the matrix obtained by adding a multiple of one row of A to another row of A, then* $\det(B) = \det(A)$.

(b) If B is the matrix obtained by interchanging two rows of A, then $\det(B) = -\det(A)$.

(c) If B is the matrix obtained by multiplying a row of A by a constant c, then $\det(B) = c\det(A)$.

6 Another Formula for the Determinant

In this section we seek a formula which will make it easier for us to get our hands dirty computing determinants. It is derived directly from the definition of determinant via permutations.

Let A be an $n \times n$ matrix. Recall equation (3) which wrote the formula for the determinant as a linear combination of the entries of the j^{th} column of A by grouping together all the terms of the summation corresponding to each entry:

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \ldots + a_{nj}C_{nj}.$$

The quantity C_{ij} is called the **$(\mathbf{i},\mathbf{j})$-cofactor** of the matrix A. We wish to find a specific formula for each $C_{i,j}$ so that we can evaluate the determinant using this expansion. We introduce the notation that A_{ij} is 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

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

Proposition 2 *Consider a matrix A whose first column is $a_{11} \cdot \vec{e}_1$. Then*

$$\det(A) = a_{11} \det(A_{11}).$$

Proof. Using the definition of determinant, we write

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} sign(\sigma) \\ &= a_{11} \sum_{\substack{\sigma \in S_n \\ \sigma(1)=1}} a_{\sigma(2)2} \cdots a_{\sigma(n)n} sign(\sigma) \end{aligned} \tag{4}$$

where the sum is simplified since any permutation which assigns $\sigma(1) \neq 1$ contributes zero to the sum. But a permutation σ on n symbols with $\sigma(1) = 1$ is really just a permutation $\tilde{\sigma}$ on the $n - 1$ symbols 2 through n. And the sign of σ is equal to the sign of $\tilde{\sigma}$ since both require the same number of transpositions to bring them back to their respective identities.

Let A_{11} denote the matrix obtained by deleting the first row and column of A. and let b_{ij} be the entries of A_{11}. From the above discussion, it is clear that

$$\sum_{\substack{\sigma \in S_n \\ \sigma(1)=1}} a_{\sigma(2)2} \cdots a_{\sigma(n)n} sign(\sigma) = \sum_{\sigma \in S_{n-1}} b_{\tilde{\sigma}(1)1} \cdots b_{\tilde{\sigma}(n-1)n-1} sign(\tilde{\sigma}) = \det(A_{11}).$$

Then equation (4) becomes

$$\det(A) = a_{11}\det(A_{11}).$$

Using this result, it is easy to see what the cofactors C_{ij} must be. Consider a 3×3 matrix A. Using the linearity of Property (3) on the first column, we write,

$$\begin{aligned}
\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
&= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} - \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ 0 & a_{12} & a_{13} \\ 0 & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{31} & a_{32} & a_{33} \\ 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \end{vmatrix}. \qquad (5)
\end{aligned}$$

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 (5), the 2×2 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 2 to evaluate each determinant, we obtain

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

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 Proposition 2 to evaluate the determinant of each matrix. In this way we arrive at the following formula for the determinant of a matrix.

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

This is called the **cofactor expansion along the first column** of A. 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 the formulas which are the goal of this section. 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" is justified because we see from the definition of the cofactors that $C_{ij} = (-1)^{i+j} \det(A_{ij})$.

Note that if A is 2×2, 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 $C = \begin{bmatrix} 0 & 1 & 3 & -3 \\ -2 & -3 & -5 & 2 \\ 2 & 0 & 0 & 1 \\ 4 & -4 & 4 & -6 \end{bmatrix}$. We expand along the third row.

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

7 Further Properties of the Determinant

In this section, we use the formula for the determinant in terms of the cofactor expansion to derive some useful results which can simplify the evaluation of a determinant in certain cases. We also use the determinant to establish an invertibility criterion for square matrices. The results of this section are simple corollaries of the following theorem.

Theorem 3 *The determinant of a triangular matrix is equal to the product of the entries on the main diagonal.*

Proof. Since the first column of A is $a_{11} \cdot \vec{e}_1$, we use Proposition 2 to conclude that

$$\det(A) = a_{11}\det(A_{11}).$$

Notice that A_{11} is upper triangular as well so that its first column is a constant times $\vec{e}_1$. We let $B_1 = A_{11}$ and apply Proposition 2 to obtain

$$\det(A) = a_{11}\det(B_1) = a_{11}a_{22}\det(B_2),$$

where B_2 is the $(n-2) \times (n-2)$ matrix obtained by deleting the first two rows and columns of A. Again, B_2 is easily seen to be upper triangular. In fact, if B_j is the matrix obtained by deleting the first j rows and columns of A, then B_j will be upper triangular. We simply apply Proposition 2 to each B_j to obtain

$$\det(B_{j-1}) = a_{jj}\det(B_j).$$

This process ends when we arrive at B_{n-1} which is simply the 1×1 matrix $[a_{nn}]$. Clearly $\det(B_{n-1}) = a_{nn}$ and so

$$\det(A) = a_{11} \cdot a_{22} \cdots a_{nn}$$

which is what we wanted to prove.

If A is lower triangular, A^T is upper triangular and so using Property (8)

$$\det(A) = \det(A^T) = a_{11} \cdot a_{22} \cdots a_{nn}.$$

□

In addition to simplifying the computation of determinants of triangular matrices, this result has two immediate consequences for non-triangular matrices: the first is another formula for the determinant of a matrix; the second is a criterion for the invertibility of a matrix.

Now recall Proposition 1 of Section 5 which stated how the determinant of a matrix is affected as we perform row operations on the matrix. Notice that any matrix can be brought to echelon form without using any scalings: only row interchanges and row replacements (i.e., replacing a row by the sum of itself and a multiple of another row.)

Corollary 1 *If a square matrix A is reduced to echelon form using only row interchanges and row replacements (but no scalings) then*

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

Proof. Let E be an echelon form of A obtained by performing only row interchanges and row replacements. E is an upper triangular matrix so by Theorem 3,

$$\det(E) = (\text{product of diagonal entries}) = (\text{product of pivots}).$$

E was obtained from A by row replacements, which do not change the determinant, and by row interchanges, which only change the sign of the determinant. So $\det(A)$ must agree with $\det(E)$ up to its sign. □

Corollary 1 immediately gives us a criterion for invertibility. If a matrix A is invertible, any echelon form of A has no zero pivots and so $\det(A)$ cannot be zero by Corollary 1. On the other hand, if A is not invertible, then it has at least one zero pivot and so $\det(A) = 0$. (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 1 establishes is the fact that an invertible matrix has a non-zero determinant.)

Corollary 2 *A square matrix A is invertible if and only if* $\det(A) \neq 0$.