**The Definition of the Determinant**

For our discussion of the determinant I am going to be using a slightly different definition for the determinant than the author uses in the text. The reason I want to do this is because I think that this definition will give you a little more insight into how the determinant works and may make some of the proofs we have to do a little easier. I will also show you that my definition is equivalent to the author's definition.

My definition starts with the concept of a *permutation*. A permutation of a list of numbers is a particular re-ordering of the list. For example, here are all of the permutations of the list {1,2,3}.

1,2,3

1,3,2

2,3,1

2,1,3

3,1,2

3,2,1

I will use the greek letter $\sigma$ to stand for a particular permutation. If I want to talk about a particular number in that permutation I will use the notation $\sigma(j)$ to stand for element $j$ in the permutation $\sigma$. For example, if $\sigma$ is the third permutation in the list above, $\sigma(2) = 3$.

The *sign* of a permutation is defined as $(-1)^{i(\sigma)}$, where $i(\sigma)$, the *inversion count* of $\sigma$, is defined as the number of cases where $i < j$ while $\sigma(i) > \sigma(j)$. Closely related to the inversion count is the *swap count*, $s(\sigma)$, which counts the number of swaps needed to restore the permutation to the original list order. It is possible to prove that $(-1)^{i(\sigma)} = (-1)^{s(\sigma)}$, so most of the time I will be using the swap count in place of the inversion count because it is easier to understand. For example, the sign of the permutation $\sigma$ = 3,2,1 is -1 because we can restore 3,2,1 to 1,2,3 in a single swap and $(-1)^1 = -1$.

With these concepts defined we can now define the determinant.

The *determinant* of an $n$ by $n$ matrix $A$ is defined to be

$$\sum_{\sigma} (-1)^{i(\sigma)} a_{1,\sigma(1)}\, a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} \tag{1}$$

The sum is carried out over all permutations of the list $\{1,\ldots,n\}$.

For example, using the list of permutations above we see that the determinant of the 3 by 3 matrix

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

is

$$a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{2,3}a_{3,2} + a_{1,2}a_{2,3}a_{3,1} - a_{1,2}a_{2,1}a_{3,3} + a_{1,3}a_{2,1}a_{3,2} - a_{1,3}a_{2,2}a_{3,1}$$

**The textbook definition**

By the definition above, to compute the determinant of an $n$ by $n$ matrix you have to start by writing down all the permutations of the list $\{1,2,\ldots,n\}$. After computing the sign of each permutation you can then use the formula to write out the determinant. In practice, this is a difficult thing to do, mostly because writing out the list of permutations is very tedious and error prone.

The author has a slightly different definition for the determinant. The author's definition effectively takes the definition shown above as a starting point and folds in a clever algorithm for computing permutations and their signs.

I am going to state the author's definition first, and then show you why it is equivalent to the definition above.

The author's definition begins with some preliminary definitions:

1. The *minor* of the entry $a_{i,j}$ in the matrix $A$, denoted $M_{i,j}$, is the matrix obtained by deleting row $i$ and column $j$ from the matrix $A$.
2. The *cofactor* of an entry $a_{i,j}$ in the matrix $A$, denoted $A_{i,j}$, is $(-1)^{i+j}\,|M_{i,j}|$.
3. The determinant of a 1 by 1 matrix is simply the value of its sole entry.

Finally, here is the author's definition of the determinant:

The determinant $|A|$ of an $n$ by $n$ matrix $A$ is defined to be

$$|A| = \sum_{j=1}^{n} a_{i,j}\,A_{i,j} = \sum_{j=1}^{n} (-1)^{i+j}\,a_{i,j}\,|M_{i,j}| \qquad (2)$$

for any $i$. (This formula is commonly referred to as *expansion in cofactors about the row $i$*.)

Since this formula is supposed to give the same result regardless of which row we choose to expand about, I will commonly do my expansions about row 1 for simplicity.

To begin to convince you that the author's definition is equivalent to mine, let's go through the exercise of using the author's definition to compute the determinant of

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

Since the author's definition defines the determinant of an $n$ by $n$ matrix in terms of combinations of determinants of minors, which are $(n-1)$ by $(n-1)$ matrices, we will find it useful to first determine the determinant of a 2 by 2 matrix. Let

$$B = \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix}$$

By formula (2) we have

$$|B| = b_{1,1} \, |M_{1,1}| - b_{1,2} \, |M_{1,2}| = b_{1,1} \, b_{2,2} - b_{1,2} \, b_{2,1}$$

(This is the well-known *cross multiplication formula* for the determinant of a 2 by 2 matrix.)

Now back to the formula for the determinant of $A$. By formula (2) expanding about row 1 we have

$$|A| = a_{1,1} \, |M_{1,1}| - a_{1,2} \, |M_{1,2}| + a_{1,3} \, |M_{1,3}|$$

$$|M_{1,1}| = \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} = a_{2,2} \, a_{3,3} - a_{2,3} \, a_{3,2}$$

$$|M_{1,2}| = \begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} = a_{2,1} \, a_{3,3} - a_{2,3} \, a_{3,1}$$

$$|M_{1,3}| = \begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} = a_{2,1} \, a_{3,2} - a_{2,2} \, a_{3,1}$$

$$|A| = a_{1,1} \left(a_{2,2} \, a_{3,3} - a_{2,3} \, a_{3,2}\right) - a_{1,2} \left(a_{2,1} \, a_{3,3} - a_{2,3} \, a_{3,1}\right) + a_{1,3} \left(a_{2,1} \, a_{3,2} - a_{2,2} \, a_{3,1}\right)$$

Expanding the right hand side out fully shows the exact same formula as above.

**Why are the definitions equivalent?**

The author's definition is based on a clever inductive algorithm for generating permutations. To explain how this works, let me introduce one more bit of notation. Let $perm(\{1,\ldots n\})$ stand for the list of all permutations of the list $\{1,\ldots n\}$. Let $\{1,\ldots,n\} - \{j\}$ stand for the list $1,2,\ldots n$ with element $j$ removed. With this notation, the list of permutations for a list $\{1,2,\ldots,n\}$ can be defined recursively as

$$1, perm(\{1,\ldots,n\} - \{1\})$$

$$2, perm(\{1,\ldots,n\} - \{2\})$$

$$\vdots$$

$$n, perm(\{1,\ldots,n\} - \{n\})$$

The author's formula effectively starts with the full permutation formula

$$\sum_{\sigma} (-1)^{i(\sigma)} a_{1,\sigma(1)} \, a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

and organizes the permutations according to the scheme above. All of the $\sigma$'s with $\sigma(1) = 1$ come first, followed by the permutations with $\sigma(1) = 2$, and so on. Next, the common first terms get factored out of each group - those common factors become the $a_{1,j}$ terms in the formula. The remaining terms become the cofactors, $A_{1,j}$.

How does the author's method get the signs right? Well, if you want to form a permutation that starts with an integer $k$ and has all the other numbers in order after that you can move the $k$ to the front from its starting position by doing $k$-1 swaps of $k$ with its left neighbor. Those swaps give the permutation starting with $k$ a sign of $(-1)^{k-1} = (-1)^{k+1}$. This accounts for the signs in the author's formula.

**A Practical Method for Computing a Determinant**

Leaving behind for the moment some of the complications caused by the two definitions above, let's take a look at a much more practical and straightforward method for computing the determinant of a matrix A. This practical method comes from the following three useful theorems. I will postpone the proof of these theorems until later.

**Theorem 1** Let $A$ be an upper triangular matrix. The determinant of $A$ is the product of the entries running down the diagonal.

**Theorem 2** Let matrix $B$ be the matrix that results when we interchange rows $j$ and $k$ in matrix $A$ ($j \neq k$). The determinants of these two matrices are related by

$$|B| = -|A|$$

**Theorem 3** Let matrix $B$ be the matrix that results when we multiply every entry of a single row in $A$ by a constant $c$. The determinants of these two matrices are related by

$$|B| = c\,|A|$$

or equivalently

$$|A| = \frac{1}{c}\,|B|$$

**Theorem 4** Let matrix $B$ be the matrix that results when we add a multiple of row $j$ to row $k$ in matrix $A$. The determinants of these two matrices are related by

$$|B| = |A|$$

The fact that these last three theorems involve the three elementary operations in Gauss-Jordan elimination should come as no surprise to you. Elimination is slowly emerging as the universal solution for just about every problem we run across.

Let's use this method to compute the determinant of a 4 by 4 matrix.

$$A = \begin{bmatrix} 2 & 0 & 1 & 3 \\ 1 & -1 & 2 & 0 \\ -1 & 1 & 1 & -1 \\ 0 & 1 & 3 & 2 \end{bmatrix}$$

To start with, we would like to switch rows 1 and 2 and then eliminate down the first column. Only the switch affects the value of the determinant. To compensate for the effect of the switch, we throw in a minus sign.

$$\begin{vmatrix} 2 & 0 & 1 & 3 \\ 1 & -1 & 2 & 0 \\ -1 & 1 & 1 & -1 \\ 0 & 1 & 3 & 2 \end{vmatrix} = - \begin{vmatrix} 1 & -1 & 2 & 0 \\ 0 & 2 & -3 & 3 \\ 0 & 0 & 3 & -1 \\ 0 & 1 & 3 & 2 \end{vmatrix}$$

Next, we switch the second and fourth rows at the cost of introducing another minus sign and continue eliminating down the second column.

$$\begin{vmatrix} 2 & 0 & 1 & 3 \\ 1 & -1 & 2 & 0 \\ -1 & 1 & 1 & -1 \\ 0 & 1 & 3 & 2 \end{vmatrix} = - \begin{vmatrix} 1 & -1 & 2 & 0 \\ 0 & 2 & -3 & 3 \\ 0 & 0 & 3 & -1 \\ 0 & 1 & 3 & 2 \end{vmatrix} = - - \begin{vmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 3 & -1 \\ 0 & 2 & -3 & 3 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & -9 & -1 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & -4 \end{vmatrix}$$

We can now use theorem 1 to compute the determinant of the last matrix.

$$\begin{vmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & -4 \end{vmatrix} = (1)(1)(3)(-4) = -12$$

**Comparing the Two Methods**

We now have two methods for computing the determinant of an *n* by *n* matrix, the author's expansion by cofactor method and the method based on elimination. Which is better?

In a few cases, the cofactor method is better. Certainly, for a 2 by 2 matrix the cofactor method leads to a simple formula that is easy to remember:

$$\begin{vmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{vmatrix} = b_{1,1}\, b_{2,2} - b_{1,2}\, b_{2,1}$$

Another case in which the cofactor formula (2) is better is the case of the *sparse matrix*. A sparse matrix is a matrix containing a lot of zeros. In many cases, a judicious combination of theorem 2 with formula (2) makes quick work of a sparse matrix. Consider this 5 by 5 example.

$$\begin{bmatrix} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 & 0 \end{bmatrix}$$

Notice row 3 - it has a lot of zeros in it. If we switch row 3 with row 1 and apply formula (2) we get

$$\begin{vmatrix} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 & 0 \end{vmatrix} = - \begin{vmatrix} 0 & 0 & 0 & 3 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 & -1 \\ 0 & 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 & 0 \end{vmatrix} = -\left((-1)^{1+4}\, 3\, |M_{1,4}|\right)$$

$$|M_{1,4}| = \begin{vmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{vmatrix} = -\left( (-1)^{1+1} \; 1 \begin{vmatrix} 0 & 0 & -1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} \right)$$

$$\begin{vmatrix} 0 & 0 & -1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = (-1)^{1+3}(-1) \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix} = -(2-0) = -2$$

Putting all of this together gives

$$\begin{vmatrix} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 & 0 \end{vmatrix} = -\left( (-1)^{1+4} \; 3 \left( -\left( (-1)^{1+1} \; 1 \; (-2) \right) \right) \right) = 6$$

For most other problems involving determinants the method based on elimination is far superior. The difference is the number of calculations needed to carry out each of the two methods. When we do a Gauss-Jordan elimination we usually need to eliminate something like $n^2/2$ elements. Eliminating a single element is done by computing the linear combination of two rows - a step that requires $n$ multiplications and $n$ additions. Thus Gauss-Jordan elimination takes about $n^3/2$ multiplications and $n^3/2$ additions to carry out. Computing the determinant by either the expansion by minors method or the equivalent permutation definition requires that we compute a product of $n$ factors for each possible permutation of $\{1,\ldots,n\}$. Unfortunately, the number of possible permutations is huge: $n!$ for a list with $n$ elements. Thus the expansion method requires $n!$ additions and $(n-1)\,(n!)$ multiplications. The difference starts to be quite dramatic after $n=5$.

| $n$ | $n^3$ | $n\,n!$ |
|---|---|---|
| 3 | 27 | 18 |
| 4 | 64 | 96 |
| 5 | 125 | 600 |
| 6 | 216 | 4320 |
| 7 | 343 | 35280 |

**Proving the Theorems**

We start by proving the most difficult of the four theorems.

**Theorem 2** Let matrix $B$ be the matrix that results when we interchange rows $j$ and $k$ in matrix $A$ ($j \neq k$). The determinants of these two matrices are related by

$$|B| = -\,|A|$$

**Proof** From the permutation definition we have that

$$|B| = \sum_{\sigma} (-1)^{i(\sigma)} b_{1,\sigma(1)} \, b_{2,\sigma(2)} \ldots b_{n,\sigma(n)}$$

For $l$ not equal to $j$ or $k$,

$$b_{l,\sigma(l)} = a_{l,\sigma(l)}$$

because the elements in row $l$ of $B$ match the elements in row $l$ of $A$. For $l$ equal to $j$ we have

$$b_{j,\sigma(j)} = a_{k,\sigma(j)}$$

because the elements in row $j$ of $B$ match the elements in row $k$ of $A$. Likewise,

$$b_{k,\sigma(k)} = a_{j,\sigma(k)}$$

Thus

$$|B| = \sum_{\sigma} (-1)^{i(\sigma)} a_{1,\sigma(1)} \cdots a_{k,\sigma(j)} \ldots a_{j,\sigma(k)} \ldots a_{n,\sigma(n)}$$

Next, let $\sigma'$ be a permutation with the property that

$$\sigma'(l) = \begin{cases} \sigma(k) & \text{for } l{=}j \\ \sigma(j) & \text{for } l{=}k \\ \sigma(l) & \textit{otherwise} \end{cases}$$

Replacing $\sigma$ with $\sigma'$ gives

$$|B| = \sum_{\sigma} (-1)^{i(\sigma)} a_{1,\sigma(1)} \cdots a_{k,\sigma(j)} \ldots a_{j,\sigma(k)} \ldots a_{n,\sigma(n)} = \sum_{\sigma} (-1)^{i(\sigma)} a_{1,\sigma'(1)} \cdots a_{k,\sigma'(k)} \ldots a_{j,\sigma'(j)} \ldots a_{n,\sigma'(n)}$$

Finally, since $\sigma'$ differs from $\sigma$ by a single interchange, $i(\sigma') = i(\sigma) + 1$ and

$$|B| = \sum_{\sigma} (-1)^{i(\sigma')-1} a_{1,\sigma'(1)} \cdots a_{k,\sigma'(k)} \ldots a_{j,\sigma'(j)} \ldots a_{n,\sigma'(n)}$$

$$= -\sum_{\sigma'} (-1)^{i(\sigma')} a_{1,\sigma'(1)} \cdots a_{k,\sigma'(k)} \ldots a_{j,\sigma'(j)} \ldots a_{n,\sigma'(n)}$$

$$= -|A|$$

Theorem 2 leads to immediately to two useful corollaries.

**Corollary 1** Let $A$ be a matrix with a row of all zeros. The determinant of $A$ is 0.

**Proof** Suppose the row of zeros is row one. From the expansion by minors formula

$$|A| = \sum_{j=1}^{n} a_{1,j} A_{1,j} = \sum_{j=1}^{n} 0\, A_{1,j} = 0$$

because all the entries in row one are 0. If the row of zeros is not row one, simply switch the row of zeros with row one at the cost of introducing a factor of -1. Since (-1) 0 = 0, the determinant is still 0.

**Corollary 2** Let $A$ be a matrix with two identical rows. The determinant of $A$ is 0.

**Proof** Let B be the matrix that results from swapping the two identical rows. Since the rows are identical, $A = B$

and $|A| = |B|$. Theorem 2 also says

$$|B| = - |A|$$

The only way to reconcile these two is to have $|A| = 0$.

After theorem 2, most of the remaining theorems are easy to prove.

**Theorem 1** Let $A$ be an upper triangular matrix. The determinant of $A$ is the product of the entries running down the diagonal.

**Proof** In an upper triangular matrix all entries $a_{i,j}$ with $i > j$ are 0. For every permutation $\sigma$ except for one we will have $j > \sigma(j)$ for at least one $j$. The only exception is the permutation for which $\sigma(j) = j$ for all $j$. For that particular permutation $i(\sigma) = 0$. Thus

$$|A| = \sum_{\sigma} (-1)^{i(\sigma)} a_{1,\sigma(1)}\, a_{2,\sigma(2)} \ldots a_{n,\sigma(n)} = a_{1,1}\, a_{2,2} \cdots a_{n,n}$$

which is the product of the entries down the diagonal.

**Theorem 3** Let matrix $B$ be the matrix that results when we multiply every entry of a single row in $A$ by a constant $c$. The determinants of these two matrices are related by

$$|B| = c\,|A|$$

**Proof** Let $k$ be the index of the row that has the additional factor of $c$. For elements on that row we have

$$b_{k,\sigma(k)} = c\, a_{k,\sigma(k)}$$

For all other rows we have

$$b_{j,\sigma(j)} = a_{j,\sigma(j)}$$

Thus

$$|B| = \sum_{\sigma} (-1)^{i(\sigma)} b_{1,\sigma(1)} \ldots b_{k,\sigma(k)} \ldots b_{n,\sigma(n)} = \sum_{\sigma} (-1)^{i(\sigma)} a_{1,\sigma(1)} \ldots c\, a_{k,\sigma(k)} \ldots a_{n,\sigma(n)} = c\,|A|$$

**Theorem 4** Let matrix $B$ be the matrix that results when we add a multiple of row $j$ to row $k$ in matrix $A$. The determinants of these two matrices are related by

$$|B| = |A|$$

**Proof** For elements in row $k$ of $B$ we have

$$b_{k,\sigma(k)} = a_{k,\sigma(k)} + c\, a_{j,\sigma(k)}$$

For all other rows

$$b_{j,\sigma(j)} = a_{j,\sigma(j)}$$

Thus we have

$$|B| = \sum_{\sigma} (-1)^{i(\sigma)} b_{1,\sigma(1)} \dots b_{k,\sigma(k)} \dots b_{n,\sigma(n)}$$

$$= \sum_{\sigma} (-1)^{i(\sigma)} a_{1,\sigma(1)} \dots \left(a_{k,\sigma(k)} + c\, a_{j,\sigma(k)}\right) \dots a_{n,\sigma(n)}$$

$$= \sum_{\sigma} (-1)^{i(\sigma)} a_{1,\sigma(1)} \dots a_{k,\sigma(k)} \dots a_{n,\sigma(n)} + c \sum_{\sigma} (-1)^{i(\sigma)} a_{1,\sigma(1)} \dots a_{j,\sigma(j)} \dots a_{j,\sigma(k)} \dots a_{n,\sigma(n)}$$

The second sum in the line above looks like the determinant of a matrix whose jth and kth rows are identical. By corollary 2 above that determinant is 0. The first sum in the last line above is the determinant of $A$. Thus the determinant of $B$ is equal to the determinant of $A$.