

# Determinants

## §1. Prerequisites

1) Every row operation can be achieved by pre-multiplying (left-multiplying) by an invertible matrix  $E$  called the *elementary matrix* for that operation. The matrix  $E$  is obtained by applying the row operation to the appropriate identity matrix. The matrix  $E$  is also denoted by  $A_{ij}(c)$ ,  $M_i(c)$ , or  $P_{ij}$ , respectively, for the operations add  $c$  times row  $j$  to  $i$  operation, multiply row  $i$  by  $c$ , and permute rows  $i$  and  $j$ , respectively.

2) The Row Reduction Theorem asserts that every matrix  $A$  can be row reduced to a unique row echelon reduced matrix  $R$ . In matrix form: There is a unique row reduced matrix  $R$  and some elementary  $E_i$  with  $E_p \cdots E_1 A = R$ , or equivalently,  $A = F_1 \cdots F_p R$  where  $F_i = E_i^{-1}$  are also elementary.

3) A matrix  $A$  determines a linear transformation: It takes vectors  $\mathbf{x}$  and gives vectors  $A\mathbf{x}$ .

## §2. Restrictions

All matrices must be square. Determinants are not defined for non-square matrices.

## §3. Motivation

Determinants determine whether a matrix has an inverse. They give areas and play a crucial role in the change of variables formula in multivariable calculus.

Let's compute the area of the parallelogram determined by vectors  $(a, b)$  and  $(c, d)$ . See Figure 1.

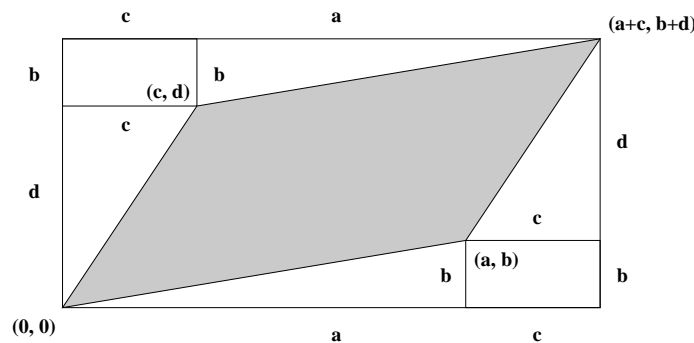


Figure 1.

The area is  $(a+c)(b+d) - 2(\frac{1}{2}ab) - 2(\frac{1}{2}cd) - 2bc = ab + ad + cb + cd - ab - cd - 2bc = ad - bc$ .

**Tentative definition:** The determinant of a 2 by 2 matrix is

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

which is the “signed” area of the parallelogram with sides determined by the rows of the matrix.

A similar argument for the signed volume of the parallelepiped (box with parallel sides) whose sides are determined by vectors  $(a, b, c)$ ,  $(d, e, f)$  and  $(g, h, i)$  shows

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - gec - hfa - idb.$$

We would like to extend these arguments to define determinants for  $n$  by  $n$  matrices. Unfortunately, this approach has problems. It is unclear what the  $n$ -dimensional analogue is and how signed  $n$ -volume should be defined. Further the above formulas do not easily generalize.

#### §4. Area and Row Operations

Area satisfies a few properties which have a connection to row operations.

Area of the unit square is 1. Area is invariant under shearing (see top of Fig. 2.), which corresponds to an add operation on the matrix. Area changes in proportion to a scaling (see middle of Fig. 2.), which corresponds to a multiply operation.

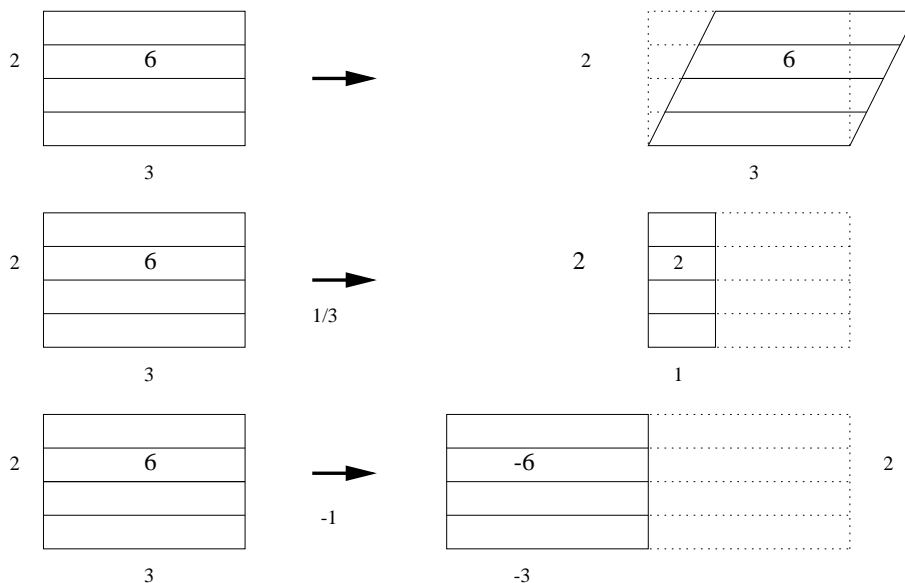


Figure 2.

Scaling by  $-1$  negates the area (see bottom of Fig 2.) since by convention, the area is positive if the angle from the *first* to the *second* vector is *counter-clockwise*. In this case, the vectors in the given order are said to form a *right-handed* system.

The above properties generalize easily and are closely enough connected to row operations that they form a good starting point for defining determinants.

**Definition:** The *determinant*,  $\det(A)$  or  $|A|$ , of a square  $n$  by  $n$  matrix  $A$  is the number satisfying axioms:

**Add Axiom** — adding one row to another does not change the determinant

$$\begin{vmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i + \mathbf{a}_j \\ \vdots \\ \mathbf{a}_n \end{vmatrix} = \begin{vmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_n \end{vmatrix}, \quad i \neq j.$$

**Multiply Axiom** — a constant can be pulled out of a single row

$$\begin{vmatrix} \mathbf{a}_1 \\ \vdots \\ c\mathbf{a}_i \\ \vdots \\ \mathbf{a}_n \end{vmatrix} = c \begin{vmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_n \end{vmatrix}, \quad \text{with } c = 0 \text{ allowed.}$$

**Identity Axiom** — identity has determinant 1

$$\begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{vmatrix} = 1.$$

**Remark:** The determinant as defined above is actually a function from the set of all square matrices to the set of numbers. There are two potential problems with every axiomatic approach like this. First, the axioms may not be restrictive enough — there could be several different determinant functions satisfying the axioms — for instance, without the identity axiom this would be the case. Second, the axioms may be too restrictive (“contradictory axioms”) — perhaps no function satisfies them all — for instance, with additional axiom  $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 0$  this would be the case.

Once we have deduced enough properties from the axioms, we show that there is no problem — the determinant function is unique and it exists. If we relied on geometry, because  $n$ -volume satisfies the above axioms, we would have the existence already. However, we do not rely on geometry, and in fact, define  $n$ -volume using determinants.

**Definition:** The  $n$ -volume of the  $n$ -dimensional parallelepiped with sides determined by vectors  $\mathbf{a}_1, \dots, \mathbf{a}_n$  which form the rows of a matrix  $A$  is  $\det(A)$ . The vectors are said to form a *right-handed* or *left-handed* system as  $\det(A)$  is positive or negative, respectively.  $n$ -volume is sometimes also called  $n$ -area. 1-volume is length, 2-volume is area and 3-volume is volume.

**Proposition:** General add row operations,  $A_{ij}(c)$ , do not change the determinant of a matrix.

*Proof:*  $A_{ij}(c) = M_j(1/c)A_{ij}(1)M_j(c)$  if  $c$  is non-zero.

**Proposition:** Each permute row operation changes the sign of the determinant.

*Proof:* Every permute operation can be achieved with three add operations and three multiply by  $-1$  operations. In matrix terms:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

**Example:**  $\begin{vmatrix} 3 & 4 \\ 5 & 6 \end{vmatrix} \stackrel{\text{Add}}{=} \begin{vmatrix} 3 & 4 \\ 0 & -2/3 \end{vmatrix} \stackrel{\text{Add}}{=} \begin{vmatrix} 3 & 0 \\ 0 & -2/3 \end{vmatrix} \stackrel{\text{Multi. } 3}{=} \begin{vmatrix} 1 & 0 \\ 0 & -2/3 \end{vmatrix} \stackrel{\text{Multi. } 3 \cdot -2/3}{=} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \stackrel{\text{Identity}}{=} -2.$

**Example:** Using a permute operation on previous, we see  $\begin{vmatrix} 5 & 6 \\ 3 & 4 \end{vmatrix} = 2.$

**Example:**  $\begin{vmatrix} 3 & 4 \\ 6 & 8 \end{vmatrix} \stackrel{\text{Add}}{=} \begin{vmatrix} 3 & 4 \\ 0 & 0 \end{vmatrix} = \begin{vmatrix} 3 & 4 \\ 0 \cdot 0 & 0 \cdot 0 \end{vmatrix} \stackrel{\text{Multi. } 0}{=} 0 \begin{vmatrix} 3 & 4 \\ 0 & 0 \end{vmatrix} = 0.$

The argument in the last example generalizes; if  $A$  reduces to a matrix  $R$  with a row of zeros then  $|A| = 0$ . Notice that, a **square** matrix  $R$  in row echelon form either has a row of zeros or is the identity. In the case where  $A$  reduces to  $I$ , there can never be a row of zeros at any point in the reduction, so  $|A| \neq 0$ : only the multiply by  $c = 0$  operation can generate a 0 determinant, but then there is a row of zeros.

This shows:

**Theorem (Determinants and Invertibility):**  $|A| = 0$  if and only if  $A$  does not reduce to  $I$   
 if and only if  $A$  reduces to a form with a row of zeros if and only if  $A^{-1}$  does not exist.

**Application:**  $(A - \lambda I)\mathbf{x} = \mathbf{0}$  has a non-trivial solution in  $\mathbf{x}$  if and only if  $|A - \lambda I| = 0$ .

## §5. Determinants via Row Operations

**Proposition (Upper Triangular Determinants):**

$$\begin{vmatrix} c_1 & & * \\ & \ddots & \\ 0 & & c_n \end{vmatrix} = c_1 c_2 \cdots c_n$$

*Proof:* If any  $c_i = 0$  then the matrix does not reduce to the identity and its determinant is 0, while if all  $c_i$  are nonzero then  $c_i$  can be pulled out giving

$$c_1 \cdots c_n \begin{vmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{vmatrix} \stackrel{\text{Add}}{=} c_1 \cdots c_n |I| = c_1 c_2 \cdots c_n.$$

Similarly for lower triangular matrices.

**Row Method:** One of the computationally fast ways to compute a determinant is with row operations. Use add operations to zero out the entries below the diagonal. Sometimes a few permute operations are also needed, each contributing a  $-1$  factor to the determinant. This leaves an upper triangular matrix, whose determinant is the product of the diagonal entries.

**Uniqueness:** The row method shows that there can be at most one determinant function. Using row operations, the value  $\det(A)$  can be determined for every  $A$ .

**Note:**  $\det(a) = a$  for 1 by 1 matrices.

**Caution:**  $|a|$  usually means the absolute value, not the determinant so use  $\det(a)$  for 1 by 1 matrices.

**Warning:**  $|cA| = c^n |A|$  for  $n$  by  $n$  matrices as there is one factor of  $c$  for each row. Geometrically, this is an obvious fact. If you take a cube and double the length of all the sides then the volume is  $c^n = 2^3 = 8$  times as much.

## §6. Determinants of Elementary Matrices and Products

$$(1) \quad |E| = \begin{cases} 1 & \text{if } E \text{ is an add matrix,} \\ c & \text{if } E = M_i(c), \\ -1 & \text{if } E \text{ is a permute matrix.} \end{cases}$$

The add and multiply elementary matrices are triangular so the result is easy. It follows that for elementary  $E$ :

$$(2) \quad |EA| = |E||A|$$

because for the three cases in equation (1) we get  $|EA|$  is respectively  $|A|$ ,  $c|A|$  and  $-|A|$  by using the axioms and their immediate consequences.

This observation leads to:

**Theorem:** For  $n$  by  $n$  matrices:  $|AB| = |A||B|$

*Proof:* Write  $A$  in reduced form  $A = F_1 \cdots F_p R$ .

If  $R = I$ , then

$$|AB| = |F_1 \cdots F_p B| \stackrel{\text{eq(2)}}{=} |F_1| \cdots |F_p| |B| \stackrel{\text{eq(2)}}{=} |F_1 \cdots F_p| |B| = |A||B|.$$

If  $R \neq I$ , then  $R$  has a row of zeros so  $|A| = 0$ . So it suffices to show  $|AB| = 0$ . Now  $AB = F_1 \cdots F_p RB$  and  $RB$  has a row of zeros since  $R$  does. So  $AB$  cannot reduce to  $I$ , and must therefore have determinant 0.

**Corollary:**  $|A^{-1}| = |A|^{-1}$

*Proof:* Take  $B = A^{-1}$  in the theorem obtaining  $1 = |A||A^{-1}|$ .

**Example:** Use  $|A| = 2$ ,  $|B| = 3$  and  $|C| = 5$  to determine  $|ABCB^{-1}A^2|$ .

Solution:  $|ABCB^{-1}A^2| = |A||B||C||B|^{-1}|A|^2$ . Although the original matrices cannot be rearranged, since determinants are just numbers we can rearrange the determinants obtaining  $|A|^3|C| = 2^3 \cdot 5 = 40$ .

**Example:** Under what condition is  $A^{-1}BAC$  invertible? The determinant, which is  $|A|^{-1}|B||A||C| = |B||C|$  must be nonzero. Thus  $B$  and  $C$  must have nonzero determinant. The condition is that both  $B$  and  $C$  must be invertible (and evidently also  $A$  otherwise  $A^{-1}BAC$  would be undefined).

Notice that although usually  $AB \neq BA$ , it is nevertheless true that  $|AB| = |BA|$  since both are equal to  $|A||B|$ .

**Warning:**  $|A + B| \neq |A| + |B|$  for most matrices. For example, consider  $A = B = I$  for 2 by 2 matrices.

## §7. Transposes and Column Operations

**Theorem:**  $|A^T| = |A|$

*Proof:* Write  $A$  in reduced form  $A = F_1 \cdots F_p R$ . Then  $|A^T| = |R^T F_p^T \cdots F_1^T| = |R^T| |F_p^T| \cdots |F_1^T| = |F_1^T| \cdots |F_p^T| |R^T|$ . We can drop all the transposes: the transpose of an add op. matrix is still an add op.; the transpose of a multiply op., of a permute op. or of  $R = I$  is the same matrix; the transpose of  $R \neq I$  is a matrix with a column of zeros and it must have determinant the same as  $|R| = 0$  since a matrix with a column of zeros can never reduce to  $I$  as the column never changes. Thus  $|F_1^T| \cdots |F_p^T| |R^T| = |F_1| \cdots |F_p| |R| = |F_1 \cdots F_p R| = |A|$ .

As a consequence,

**Meta Rule:** Any rule for determinants that holds for rows holds also for columns.

**Example:**  $\begin{vmatrix} 1 & 20 \\ 3 & 40 \end{vmatrix} = 10 \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$ .

**Example:** The area of the parallelogram determined by  $(1, 20)$  and  $(3, 40)$  is the same as that determined by  $(1, 3)$  and  $(20, 40)$  — it doesn't matter whether the vectors are written as rows or columns.

**Example:**  $\det[(A^T B)^{-1}]^T = (\det A \cdot \det B)^{-1}$ .

### §8. Additivity and the Laplace Expansion

**Additivity:**  $|\mathbf{a}_1 \dots \mathbf{a}_{n-1}(\mathbf{b} + \mathbf{c})| = |\mathbf{a}_1 \dots \mathbf{a}_{n-1}\mathbf{b}| + |\mathbf{a}_1 \dots \mathbf{a}_{n-1}\mathbf{c}|$  where the vectors are column vectors. By permuting and transposing, a similar result holds for any row or column.

*Proof (for addition on last column):* Apply the same row operations simultaneously to all three matrices reducing them to upper triangular form. Note that  $E(\mathbf{b} + \mathbf{c}) = E\mathbf{b} + E\mathbf{c}$  so the last column of the first matrix continues to be the sum of the last columns of the other two. If  $r_1, \dots, r_{n-1}, b'_n$  and  $r_1, \dots, r_{n-1}, c'_n$  are the diagonal entries for the second and third matrix, then  $r_1, \dots, r_{n-1}, b'_n + c'_n$  are the diagonal entries for the first matrix. Both sides of the equation equal  $r_1 \dots r_{n-1}(b'_n + c'_n)$ . This result is fairly obvious geometrically by shearing  $\mathbf{b} + \mathbf{c}$  into  $\mathbf{b}$  and  $\mathbf{c}$ . See Figure 3.

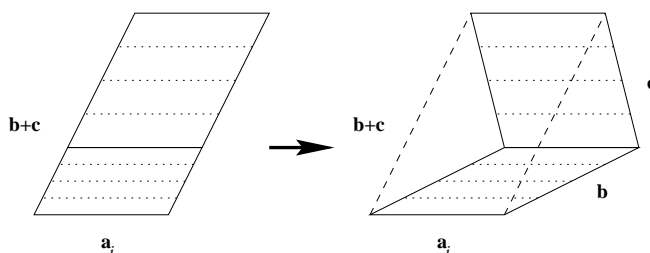


Figure 3.

**Example:**

$$\begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 4 & 5 \end{vmatrix} + \begin{vmatrix} 0 & 3 \\ 4 & 5 \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 4 & 5 \end{vmatrix} - \begin{vmatrix} 4 & 5 \\ 0 & 3 \end{vmatrix} = 2 \cdot 5 - 4 \cdot 3$$

The method used in the example can be extended. We expand a 3 by 3 determinant using additivity on the first column. The second equality is via column multiply ops and equality holds even if  $a = 0, d = 0$  or  $g = 0$ . The third equality is obtained via permute ops. The last equality holds since each matrix can be reduced to upper triangular form and the extra 1's in the previous matrices do not change the product of diagonal elements.

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= \begin{vmatrix} a & b & c \\ 0 & e & f \\ 0 & h & i \end{vmatrix} + \begin{vmatrix} 0 & b & c \\ d & e & f \\ 0 & h & i \end{vmatrix} + \begin{vmatrix} 0 & b & c \\ 0 & e & f \\ g & h & i \end{vmatrix} \\ &= a \begin{vmatrix} 1 & b & c \\ 0 & e & f \\ 0 & h & i \end{vmatrix} + d \begin{vmatrix} 0 & b & c \\ 1 & e & f \\ 0 & h & i \end{vmatrix} + g \begin{vmatrix} 0 & b & c \\ 0 & e & f \\ 1 & h & i \end{vmatrix} \\ &= a \begin{vmatrix} 1 & b & c \\ 0 & e & f \\ 0 & h & i \end{vmatrix} - d \begin{vmatrix} 1 & e & f \\ 0 & b & c \\ 0 & h & i \end{vmatrix} + g \begin{vmatrix} 1 & h & i \\ 0 & b & c \\ 0 & e & f \end{vmatrix} \\ &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} b & c \\ h & i \end{vmatrix} + g \begin{vmatrix} b & c \\ e & f \end{vmatrix} \end{aligned}$$

The method above, called Laplace expansion, extends easily to  $n$  by  $n$  matrices. By permuting and transposing it extends to any row or column.

**Laplace Expansion:** Given  $A$  square, let  $A_{ij}$  be the matrix obtained by deleting the  $i$ -th row and  $j$ -th column of  $A$ .  $A_{ij}$  is called the  $ij$ -th minor matrix and its determinant is called the  $ij$ -th minor.

Consider the sign  $(-1)^{i+j}$ . These form a checkerboard pattern  $\begin{pmatrix} + & - & + & \dots \\ - & + & - & \dots \\ + & - & + & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$ . The  $ij$ -th

cofactor is  $c_{ij} = (-1)^{i+j} \det A_{ij}$ . Consider its product with  $a_{ij}$ :

$$(-1)^{i+j} a_{ij} \det A_{ij}$$

Then  $\det A$  is the sum of these terms over any one row or column. For the degenerate case of 1 by 1 matrices it is simply the  $a_{11}$  entry.

**Remarks:** Although the formula has a certain elegance and theoretical usefulness it is an awful formula for computations. For example, a random numerical 10 by 10 determinant takes about 330 multiplications or divisions via the row operation method, but over 3.6 million multiplications via the Laplace expansion!

*Some people use it to define the determinant, but it is also awful for this purpose because it is very difficult to motivate, and totally detached from geometry and row reductions. It's only saving grace is that no divisions are required to compute it, and so it can be used in situations where division could be a problem, for example, if you restrict the numbers to be integers — even then, however, row operations can be used by first introducing the “field of fractions”.*

**Example:** Use Laplace expansion on rows or columns with a lot of zeros:

$$\begin{vmatrix} 0+ & 2 & 0 & 4 \\ 0- & 3 & 0 & 0 \\ 0+ & 0 & 7 & 0 \\ 6- & 0+ & 0- & 0+ \end{vmatrix} \stackrel{\text{last row}}{=} -6 \begin{vmatrix} 2+ & 0- & 4+ \\ 3 & 0 & 0- \\ 0 & 7 & 0+ \end{vmatrix} \stackrel{\text{last column}}{=} -6 \cdot 4 \begin{vmatrix} 3 & 0 \\ 0 & 7 \end{vmatrix} = -6 \cdot 4 \cdot 3 \cdot 7.$$

**Existence Theorem:** There is a determinant function and the axioms for determinants are not self contradictory. Namely, Laplace expansion down the first column satisfies the axioms.

*Proof:* The axioms hold for 1 by 1 matrices where  $\det(a)$  is defined as  $a$ . The add axiom is not applicable. The multiply axiom is true because  $\det(ca) = ca = c \det(a)$ . The identity axiom is trivially true.

Next we suppose the theorem is true for  $n - 1$  by  $n - 1$  matrices and show it holds for  $n$  by  $n$  matrices. This proof technique is called *induction*. We only consider the  $n = 3$  case because the general case is very similar but notationally messy. So assume that there is a determinant function for 2 by 2 matrices. Then all the properties we have proved up to this point hold for 2 by 2 matrices. We need to show that 3 by 3 determinants defined by Laplace expansion down the first column satisfy the axioms.

Identity Axiom:

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} - 0 \begin{vmatrix} * & * \\ * & * \end{vmatrix} + 0 \begin{vmatrix} * & * \\ * & * \end{vmatrix} = 1$$

Multiply Axiom: Without loss of generality consider the “multiply the first row by  $k$ ” case. The second equality follows by the multiply property for 2 by 2 determinants.

$$\begin{aligned} \begin{vmatrix} ka & kb & kc \\ d & e & f \\ g & h & i \end{vmatrix} &= ka \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} kb & kc \\ h & i \end{vmatrix} + g \begin{vmatrix} kb & kc \\ e & f \end{vmatrix} \\ &= k \left( a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} b & c \\ h & i \end{vmatrix} + g \begin{vmatrix} b & c \\ e & f \end{vmatrix} \right) \\ &= k \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \end{aligned}$$

Add axiom: Consider the case of adding the second to the first row. For the second equality use additivity on the middle determinant and subtract the second row from the first in the last determinant. For the last equality notice the second and fourth determinants cancel.

$$\begin{aligned} \begin{vmatrix} a+d & b+e & c+f \\ d & e & f \\ g & h & i \end{vmatrix} &= (a+d) \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} b+e & c+f \\ h & i \end{vmatrix} + g \begin{vmatrix} b+e & c+f \\ e & f \end{vmatrix} \\ &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} + d \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} b & c \\ h & i \end{vmatrix} - d \begin{vmatrix} e & f \\ h & i \end{vmatrix} + g \begin{vmatrix} b & c \\ e & f \end{vmatrix} \\ &= \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \end{aligned}$$

## §9. Cramer’s Rule

**Cramer’s Rule:** Write  $A = [\mathbf{a}_1 \dots \mathbf{a}_n]$ . If  $A$  is invertible, then the solution to  $A\mathbf{x} = \mathbf{b}$  is given by

$$x_i = \frac{|\mathbf{a}_1 \dots \mathbf{b} \dots \mathbf{a}_n|}{|\mathbf{a}_1 \dots \mathbf{a}_i \dots \mathbf{a}_n|}.$$

**Example:** The solution for  $\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$  has  $x_2 = \frac{\begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}} = (6 - 12)/(5 - 8)$ . However, Cramer’s rule is an inefficient method for finding solutions.

*Proof of Cramer’s Rule:* If  $A\mathbf{x} = \mathbf{b}$  then  $x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n = \mathbf{b}$ . So  $|\mathbf{a}_1 \dots \mathbf{b} \dots \mathbf{a}_n| = |\mathbf{a}_1 \dots (x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n) \dots \mathbf{a}_n|$ . Subtract  $x_1$  times the first column from the  $i$ -th column. Repeat for all columns except for the  $i$ -th obtaining

$$|\mathbf{a}_1 \dots (x_i\mathbf{a}_i) \dots \mathbf{a}_n| = x_i |\mathbf{a}_1 \dots \mathbf{a}_i \dots \mathbf{a}_n|.$$

## §10. Adj formula

The matrix  $A$  has  $ij$ -th element  $a_{ij}$ . By definition the *cofactor* matrix is  $\text{Cofactor}(A)$  having  $ij$ -th element  $c_{ij} = (-1)^{i+j} \det A_{ij}$ . The *adjugate* or (*classical adjoint*) matrix is  $\text{Adj}(A) = (\text{Cofactor}(A))^T$ . Its  $ij$ -th element is  $d_{ij} = c_{ji}$ .



**Example:**

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{Cofactor}(A) = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}, \quad \text{Adj}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

**Theorem:**  $A \text{Adj}(A) = (\det A)I$

*Proof:* The  $ii$ -th element on the left is  $\sum_j a_{ij}d_{ji} = \sum_j a_{ij}c_{ij}$  but the latter is  $\det(A)$  by Laplace expansion on the  $i$ -th row of  $A$ . For off-diagonal  $il$ -elements, which we must show are zero, on the left we get  $\sum_j a_{ij}c_{lj}$  which amounts to a Laplace expansion on the  $l$ -th row of the matrix obtained from  $A$  by duplicating the  $i$ -th row of  $A$  in the  $l$ -th row. Since this matrix has equal rows its determinant is zero.

**Remark:** This formula can be used to compute  $A^{-1}$ , but it is inefficient.

**§11. Integral Matrices**

**Proposition:** If  $A$  is integral, then  $\det A$  is an integer.

*Proof:* Row operations do not help here. Instead note that Laplace expansion will involve just integers.

**Proposition:** If  $A$  is integral and has an integral inverse, then  $\det A = \pm 1$ .

*Proof:*  $|A^{-1}| = 1/|A|$  so  $|A|$  is an integer whose reciprocal is an integer. The only such integers are  $\pm 1$ .

**Proposition:** If  $A$  is integral and  $\det A = \pm 1$  then  $A^{-1}$  is integral.

*Proof:* By the previous theorem  $A \text{Adj}(A) = (\pm 1)I$ , so  $A^{-1} = \pm \text{Adj}(A)$ . Each entry  $c_{ij} = (-1)^{i+j}|A_{ij}|$  of  $\text{Adj}(A)$  is integral by the first proposition above.

**§12. Miscellaneous**

**Theorem:** If a linear transformation with  $n$  by  $n$  standard matrix  $A$  is applied on an object with signed  $n$ -volume  $V$ , then the signed  $n$ -volume of the resulting object is  $\det(A) \cdot V$ .

*Proof:* By slicing the object into a number of  $n$ -parallelepipeds whose  $n$ -volume approximates that of the object, it suffices to show the result for the case of a single  $n$ -parallelepiped whose sides are given by  $\mathbf{b}_1, \dots, \mathbf{b}_n$ . Then  $V = |\mathbf{b}_1 \dots \mathbf{b}_n| = |B|$ . The  $n$ -parallelepiped obtained by applying  $A$  has sides  $A\mathbf{b}_1, \dots, A\mathbf{b}_n$  and volume  $|A\mathbf{b}_1 \dots A\mathbf{b}_n| = |AB| = |A||B| = |A|V$ .

The volume of the  $n$ -parallelepiped in  $\mathbf{R}^m$  with sides determined by the columns of  $m$  by  $n$  matrix  $A$  is  $\sqrt{\det(A^T A)}$ . For example, the area of a parallelogram with sides  $(1, 0, 4)$  and  $(2, 4, 6)$  can be found this way. Prove this by extending the  $n$ -box for  $A$  to an  $m$ -box for  $[A|P]$  by adjoining perpendicular unit vectors (which form  $P$ ) and using previous results.

The *sign* of a permutation  $1 \mapsto i_1, 2 \mapsto i_2, \dots, n \mapsto i_n$ , where  $\{1, 2, \dots, n\} = \{i_1, i_2, \dots, i_n\}$ , is defined to be  $|\mathbf{e}_{i_1} \mathbf{e}_{i_2} \dots \mathbf{e}_{i_n}|$ . This amounts to  $(-1)^s$  where  $s$  is the number of permutations needed to rearrange the columns as  $\mathbf{e}_1 \dots \mathbf{e}_n$ . The sign is 1 if and only if  $\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, \mathbf{e}_{i_n}$  is righthanded.

The *permanent* of a matrix is defined in the same way as Laplace expansion except the signs  $(-1)^{i+j}$  are dropped. It would appear there is no good geometric interpretation for permanents.