Determinants

Christian Zhou-Zheng

Abstract

This paper attempts to introduce determinants and their properties, as well as Leibniz's formula and Cramer's rule.

1 Overview

1.1 Introduction to Determinants

In linear algebra, the determinant of a square matrix is a scalar value which characterizes certain properties of its corresponding matrix. Determinants can be used to find the eigenvalues of a matrix, show the matrix has an inverse, solve homogenous systems of linear equations, and more. I will denote the determinant of a matrix A as $|A|$ throughout this paper.

An initial geometric intuition of the determinant can be gained through looking at 2x2 matrices as linear operators from $\mathbb{C}^2$ to $\mathbb{C}^2$. The determinant of a matrix A is the absolute value of the area of the image of the unit square under A: if A purely rotates the unit square, its determinant is 1; if A reflects the unit square, its determinant is -1; if A stretches the unit square, its determinant is greater than 1; and if A shrinks the unit square, its determinant is less than 1.

This definition can be easily extended to higher dimensions; for instance, the determinant of a 3x3 matrix is the volume of the image of the unit cube under A. However, this is impractical to scale to dimensions beyond human intuition, and so we seek a more rigorous definition of the determinant.

1.2 Properties of Determinants

The simplest way to define the determinant is through its properties:

1. The determinant of the identity matrix is 1.

2. Whenever two columns of a matrix are identical, its determinant is 0.

3. The determinant is *multilinear*. If the jth column of a matrix A is written as a linear combination $r \cdot \vec{v} + \vec{w}$ for a scalar r and vectors $\vec{v}, \vec{w}$, the determinant can be separated across addition and scalar multiplication:

$$\begin{aligned} A &= |a_1, \dots, r \cdot \vec{v} + \vec{w}, \dots, a_n| \\ &= r \cdot |a_1, \dots, \vec{v}, \dots, a_n| + |a_1, \dots, \vec{w}, \dots, a_n| \end{aligned} \tag{1}$$

These three properties can be proven by examining Leibniz's formula, which will be discussed later. Corollaries of these properties include:

4. If a column of the matrix is all 0, the determinant is 0.

5. For a scalar r, $|rA| = r^n|A|$, where n is the dimension of A.

6. Any permutation of the columns multiplies the determinant by the sign of the permutation (which will be explained with Leibniz's formula).

7. If any column is linearly dependent on the others, the determinant is 0.

8. Adding a scalar multiple of one column to another does not change the determinant.

The word "column" is interchangeable with "row" in all of the above. Four more general properties of the determinant are as follows:

10. The determinant of a matrix transpose is equal to the determinant of the original matrix: $|A^T| = |A|$.

11. The determinant of a matrix's conjugate is equal to the conjugate of the determinant of the original matrix: $|\overline{A}| = \overline{|A|}$.

12. The determinant is distributive: $|AB| = |A||B|$.

13. The determinant of the inverse of a matrix is the inverse of the determinant of the original matrix: $|A^{-1}| = |A|^{-1}$.

Examples of these are shown in the next section. Armed with this foundation, we can now proceed to more interesting math with determinants. I will go over ways to calculate the determinant, as well as Cramer's rule.

2 Calculating Determinants

Calculating determinants is easy in low dimensions but quickly becomes not only computationally inefficient, but also difficult to wrap one's head around, in higher dimensions. I will attempt to introduce the mathematics behind determinant calculations, but will leave out computational implementations.

2.1 Basic Formulas

2.1.1 2x2 Matrices

The simplest matrices to calculate determinants of (apart from trivial 1x1 matrices) are 2x2 matrices. The determinant of a 2x2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is the difference in the products of the diagonals:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \tag{2}$$

This can be proven by using the properties of determinants as stated in §1.2. Firstly, the determinant of the identity matrix is 1. Then, using the fact that the determinant is multilinear, the determinant of a 2x2 matrix can be expanded as follows:

$$\begin{aligned} \begin{vmatrix} a & b \\ c & d \end{vmatrix} &= \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & d \end{vmatrix} \\ &= \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} \\ &= ad \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + bc \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \\ &= ad - bc \end{aligned} \tag{3}$$

The sign of the latter term is inverted because the sign of the permutation is -1, something which will be expanded on in §2.2.2.

2.1.2 Demonstrating Properties

I will display some of the aforementioned determinant properties here, using simple 2x2 matrices. While these are not formal proofs that these properties hold for higher dimensions, it is useful to have a visual representation of these properties.

1. The determinant of the identity matrix is 1.

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

2. Whenever two columns of a matrix are identical, its determinant is 0.

$$\begin{vmatrix} a & a \\ b & b \end{vmatrix} = ab - ab = 0$$

3. The determinant is multilinear (under addition and scalar multiplication).

$$\begin{vmatrix} a & b+b' \\ c & d+d' \end{vmatrix} = ad+ad'-bc-b'c = (ad-bc)+(ad'-b'c) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a & b' \\ c & d' \end{vmatrix}$$

$$\begin{vmatrix} ra & b \\ rc & d \end{vmatrix} = rad - rbc = r(ad-bc) = r\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

4. If a column of the matrix is all 0, the determinant is 0.

$$\begin{vmatrix} 0 & b \\ 0 & b \end{vmatrix} = 0 \cdot b - 0 \cdot b = 0$$

8. Adding a scalar multiple of one column to another does not change the determinant.

$$\begin{vmatrix} a & b+ra \\ c & d+rc \end{vmatrix} = a(d+rc)-c(b+ra) = ad+rac-rac-bc = ad-bc = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

2.1.3 Rule of Sarrus and 3x3 Matrices

The determinant of a 3x3 matrix can be calculated using the *rule of Sarrus.* To use the rule of Sarrus, write out the matrix, but rewrite the left two columns on the right of the matrix as follows:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \rightarrow \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \begin{matrix} a & b \\ d & e \\ g & h \end{matrix} \tag{4}$$

Now, starting from the top row in the original matrix, multiply out diagonally down and right and sum these. Then, starting from the bottom row in the original matrix, multiply out diagonally up and right and sum these. Subtract the second sum from the first, and this is the determinant of the matrix:

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= (aei + bfg + cdh) - (gec + hfa + idb) \\ &= aei + bfg + cdh - afh - bdi - ceg \end{aligned} \tag{5}$$

This does not extend to higher dimensions, but it serves usefully for 3x3 matrices as are often seen in physics.

2.1.4 Laplace's Expansion

Laplace's expansion is a generalized method of calculating the determinant of a matrix by breaking it into a sum of determinants of matrices of one dimension lower. These subdeterminants are called the *minors* of the original matrix. The formula for Laplace's expansion is as follows:

$$|A| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij} \tag{6}$$

for a row index i. $M_i j$ is the minor corresponding to row i and column j, formed from the original matrix by removing row i and column j. For instance, M_{11} is the minor formed by removing the first row and first column.

For example, expanding the determinant of a 3x3 matrix along the first row yields:

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ &= a(ei - fh) - b(di - fg) + c(dh - eg) \\ &= aei + bfg + cdh - afh - bdi - ceg \end{aligned} \tag{7}$$

This equation matches that returned by using the rule of Sarrus for 3x3 matrices. However, this method quickly becomes unwieldy for higher dimensions due to the need to continually recursively expand determinants, and so we seek a more efficient method of calculating determinants.

2.2 Leibniz's Formula

Leibniz's formula provides a way to calculate the determinant of a matrix of arbitrary dimension using *permutations*. The formal statement of Leibniz's formula is as follows:

$$|A| = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i} \tag{8}$$

This states that the determinant of A is the sum, for every permutation σ in the symmetric group of row indices S_n, of the sign of σ times the product of the elements of A corresponding to the permutation. Let's walk through this step by step.

2.2.1 Permutations, Notation, and Symmetric Groups

To understand the formula statement, an understanding of permutations is imperative. In a set $S = \{s_1, s_2, \ldots, s_n\}$, a *permutation* is a bijection from the set to itself: each element is mapped to exactly one element (not necessarily different), its image. Permutations are denoted with lowercase sigma, σ.

For instance, one permutation of the set $\{1, 2, 3\}$ is $\{3, 1, 2\}$, which maps 1 to 3, 2 to 1, and 3 to 2. This is written as $\sigma(1) = 3$, $\sigma(2) = 1$, and $\sigma(3) = 2$.

More commonly, either Cauchy's two-line notation or cycle notation is used to denote permutations. Cauchy's two-line notation is written as follows:

$$\sigma = \begin{Bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{Bmatrix} \tag{9}$$

where each element in the top row is mapped to the corresponding element in the bottom row. Cycle notation is written as follows:

$$\sigma = (132) \tag{10}$$

where each element is mapped to the next element in the cycle, and the last element is mapped to the first element. Individual cycles are separated by parentheses - in this case, all elements are contained within one cycle, but this is not always the case.

In fact, cycle notation is useful for writing permutations as a product of disjoint cycles - for instance, the permutation of $S = \{1, 2, 3, 4, 5, 6\}$ such that $\sigma = (132)(45)(6)$ maps 1 to 3, 3 to 2, 2 to 1, 4 to 5, 5 to 4, and 6 to 6. This can also be expanded into smaller cycles, which will be important in the next section: $\sigma = (13)(32)(45)(6)$ is equivalent notation.

The set of all permutations of a set S is called the *symmetric group* of S, denoted S_n. For instance, the symmetric group of $\{1, 2, 3\}$ can be written in set notation:

$$S_n = \{\{1, 2, 3\}, \{1, 3, 2\}, \{2, 1, 3\}, \{2, 3, 1\}, \{3, 1, 2\}, \{3, 2, 1\}\}$$

Or in cycle notation, in corresponding order:

$$S_n = \{(1)(2)(3), (1)(23), (12)(3), (132), (123), (13)(2)\}$$

2.2.2 Sign of a Permutation

The sign of a permutation, the $\text{sgn}(\sigma)$ term in Leibniz's formula, is defined using the parity of the number of inversions of σ: even parity returns $+1$, while odd parity returns -1.

An *inversion* is a pair of elements in a permutation such that two elements are swapped from their natural ordering. For instance, in the permutation $\sigma = (1)(23)$, 2 and 3 are swapped, so there is one inversion. In the permutation $\sigma = (1)(2)(3)$, there are no inversions.

This is much easier to discern using cycle notation, as mentioned in the previous section. When writing out a permutation in cycle notation, the permutation is odd iff there are an odd number of *even-length* disjoint cycles (hereafter referred to as just "cycles"), and even otherwise.

To explain why, examine 2-cycles: cycles that map only one element to one other, like (13) or (21). Each of these can be thought of as one inversion, even if the second element is not mapped back to the first. Intuitively then, if decomposed entirely into 2-cycles, a permutation with an odd number of 2-cycles would have odd parity, and an even number of 2-cycles implies even parity.

Since 1-cycles are trivial, they can be omitted from cycle notation and only implied; they contribute nothing to the number of inversions. All remaining cycles can be written as products of 2-cycles: specifically, cycles of odd length can be written as products of even amounts of 2-cycles, and cycles of even length can be written as products of odd amounts of 2-cycles.

Considering the parity of the permutation is dependent on the parity of the number of 2-cycles, it is visible that any odd cycles contribute nothing to the parity of the permutation (as adding an even integer does not change parity), and each even cycle flips the parity (as adding an odd integer changes parity). Thus, the parity of the permutation is the parity of the number of even cycles, regardless of length.

2.2.3 Explanation and Demonstration

Armed with this knowledge of permutations and their sign, we can now return to Leibniz's formula and explain it in detail, as well as provide examples. Reiterating, Leibniz's formula is:

$$|A| = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i} \tag{11}$$

For the set of row indices of a matrix of dimension n, $\{1, 2, \ldots, n\}$, the symmetric group S_n is the set of all permutations of these indices. The sum in Leibniz's formula iterates over every one of these permutations.

The $\text{sgn}(\sigma)$ term, as explained in the previous section, determines the sign of the next term of the sum, which is the product of the elements of A corresponding to the permutation: take one element from every column, with a row index corresponding to the image of that column index in the current permutation σ.

We demonstrate this with a 2x2 and 3x3 matrix. For the 2x2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the sum iterates over the two permutations of the set $\{1, 2\}$: $\sigma = (1)(2)$ and $\sigma = (12)$. The first permutation has even parity, and the second has odd parity, so the first term of the sum is positive and the second is negative. The first term is $a_{\sigma(1),1}a_{\sigma(2),2} = a_{1,1}a_{2,2} = ad$, and the second term is $a_{\sigma(1),2}a_{\sigma(2),1} = a_{1,2}a_{2,1} = bc$. Thus, the determinant of a 2x2 matrix is $ad - bc$.

Leibniz's formula is perhaps easier to visualize using a 3x3 matrix, but in the interest of space only the permutations such that $\sigma(1) = 1$ will be considered. Of the symmetric group of $\{1, 2, 3\}$, these are (1)(2)(3) and (1)(23), which have odd and even parity respectively. The first term is $a_{\sigma(1),1}a_{\sigma(2),2}a_{\sigma(3),3} = a_{1,1}a_{2,2}a_{3,3} = aei$, and the second term is $a_{\sigma(1),1}a_{\sigma(2),3}a_{\sigma(3),2} = a_{1,1}a_{2,3}a_{3,2} = afh$.

Looking at Laplace's expansion, it can be seen that $(aei - afh) = a(ei - fh)$ is the first term after decomposition in Equation (7)! Logically, similarities can be drawn between the diagonal/zigzaggy nature of Laplace's expansion (taken from the method of calculating the 2x2 determinant across the diagonals) and Leibniz's formula, but I am not good enough at articulating abstract thought to explain it here.

3 Cramer's Rule

3.1 Statement

Cramer's rule is a method of finding the *unique* solution of a system of linear equations using determinants, when there are as many unknowns as equations. The rule gives the solution for a system with n equations and n unknowns, written as $A\vec{x} = \vec{b}$ for the n x n coefficient matrix A and vector of variables $\vec{x}$, as long as the determinant of A is nonzero. Specifically, the formula is:

$$x_i = \frac{|A_i|}{|A|} \tag{12}$$

Where A_i is the matrix A with the ith column replaced with $\vec{b}$.

3.2 Example

Consider the system of equations:

$$\begin{aligned} x + y + z &= 2 \\ 2x + y + 3z &= 9 \\ x - 3y + z &= 10 \end{aligned}$$

This can be written as $A\vec{x} = \vec{b}$, where $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \\ 1 & -3 & 1 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} 2 \\ 9 \\ 10 \end{bmatrix}$. The determinant of A is 4, so Cramer's rule can be applied.

We'll go through each column in order, omitting the expanded determinant calculations in the interest of brevity: $A_1 = \begin{bmatrix} 2 & 1 & 1 \\ 9 & 1 & 3 \\ 10 & -3 & 1 \end{bmatrix}$, then $|A_1| = 4$ and $x_1 = x = \frac{4}{4} = 1$. $A_2 = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 9 & 3 \\ 1 & 10 & 1 \end{bmatrix}$, then $|A_2| = -8$ and $x_2 = y = \frac{-8}{4} = -2$. $A_3 = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 9 \\ 1 & -3 & 10 \end{bmatrix}$, then $|A_3| = 12$ and $x_3 = z = \frac{12}{4} = 3$.

Plugging these back into the original equations, the validity of the solution is confirmed:

$$\begin{aligned} 1 + (-2) + 3 &= 2 \\ 2 + (-2) + 9 &= 9 \\ 1 + 6 + 3 &= 10 \end{aligned}$$

3.3 Proof

Take a set of linear equations:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & b_n \end{array} \tag{13}$$

This is the expanded form of $A\vec{x} = \vec{b}$. Consider the determinant of A:

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \tag{14}$$

Now multiply $|A|$ by one of our variables x_i, and using the multilinearity of determinants (property 3 demonstrated in §2.1.2), also multiply the corresponding column inside the determinant by x_i. I will demonstrate with x_1:

$$x_1|A| = \begin{vmatrix} a_{11}x_1 & a_{12} & \dots & a_{1n} \\ a_{21}x_1 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}x_1 & a_{n2} & \dots & a_{nn} \end{vmatrix} \tag{15}$$

Now, using the property of determinants that adding a constant multiple of one column to another does not change the determinant (property 8 demonstrated in §2.1.2), add each of the other columns to the target column, multiplied by the corresponding coefficient x_i:

$$x_1|A| = \begin{vmatrix} a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n & a_{12} & \dots & a_{1n} \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{n2}x_n & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & a_{n2} & \dots & a_{nn} \end{vmatrix} \tag{16}$$

The target column is now the left side of the system of equations! Now, replace that column with the corresponding elements of $\vec{b}$:

$$x_1|A| = \begin{vmatrix} b_1 & a_{12} & \dots & a_{1n} \\ b_2 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n2} & \dots & a_{nn} \end{vmatrix} \tag{17}$$

If $\vec{b} = \vec{0}$, by property 4 demonstrated in §2.1.2, this becomes $x_i|A| = 0$ for all i, meaning there is either only the trivial solution ($\vec{x} = \vec{0}$, if $|A| \neq 0$) or a family of nontrivial solutions (if $|A| = 0$). If $\vec{b} \neq 0$ and $|A| = 0$, the system has no solution, as the following step is impossible; however, **when $\vec{b} \neq 0$ and $|A| \neq 0$**, we can divide both sides by $|A|$:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & \dots & a_{1n} \\ b_2 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n2} & \dots & a_{nn} \end{vmatrix}}{|A|} \tag{18}$$

Since the determinant in the numerator is the same as the determinant of A with the first column replaced by $\vec{b}$, this is the determinant of A_1 as defined earlier. Thus, the case of Cramer's Rule for x_1 is proven:

$$x_1 = \frac{|A_1|}{|A|} \tag{19}$$

Going back to the step in Equation (15), the previous steps can be done for any x_i and its corresponding column, so this is generalized:

$$x_i = \frac{|A_i|}{|A|} \tag{20}$$

This is Cramer's Rule. $\square$