17. Volume of a parallelepiped

17.1. **Criteria for uniqueness.**

Theorem 17.1. *Let A be an $n \times n$ square matrix. TFAE (the following are equivalent: meaning if anyone of them is true, then so are all of the others):*

(1) The equation $A\vec{x} = \vec{b}$ has a unique solution, for any $\vec{b}$.
(2) The homogeneous equation $A\vec{x} = \vec{0}$ has the unique solution $\vec{x} = 0$.
(3) A is invertible (that is, there is an inverse matrix).
(4) $\det A \neq 0$.

In fact all four statements are easily seen to be equivalent to saying that there are n pivots.

17.2. **Cramer's Rule.**

An unbelievably stupid way to solve a system of linear equations. Only use this rule under threat of death (or if it is assigned as a homework problem).

Let A be an $n \times n$ matrix. Let $\vec{b}$ be a column vector with n rows. Let $A_i(\vec{b})$ be the matrix whose ith column has been replaced by $\vec{b}$.

If A is an invertible matrix then the unique solution to the equation

$$A\vec{x} = \vec{b}$$

is the vector whose ith component is

$$x_i = \frac{\det A_i(\vec{b})}{\det A}.$$

Example 17.2. *Solve the linear system*

$$\begin{aligned} 2x + 3y &= 3 \\ x - y &= 2, \end{aligned}$$

using Cramer's rule.

Let

$$A = \begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix} \qquad \text{and} \qquad \vec{b} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Then

$$A_1(\vec{b}) = \begin{pmatrix} 3 & 3 \\ 2 & -1 \end{pmatrix} \qquad \text{and} \qquad A_2(\vec{b}) = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}.$$

So

$$\det A = -5 \qquad \det A_1 = -9 \qquad \text{and} \qquad \det A_2 = 1.$$

So the unique solution is

$$x_1 = \frac{-9}{-5} = \frac{9}{5} \qquad \text{and} \qquad x_2 = \frac{1}{-5} = -\frac{1}{5}.$$

Of course it is much easier to apply Gaussian elimination:

$$\left(\begin{array}{cc|c} 2 & 3 & 3 \\ 1 & -1 & 2 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -1 & 2 \\ 2 & 3 & 3 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -1 & 2 \\ 0 & 5 & -1 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -1 & 2 \\ 0 & 1 & -1/5 \end{array}\right)$$

Thus

$$x_2 = -\frac{1}{5},$$

and so

$$x_1 + \frac{1}{5} = 2 \qquad \text{so that} \qquad x_1 = \frac{9}{5}.$$

17.3. **Volume of a parallelepiped.** Given a 2×2 matrix A note that the columns of A determine a parallelogram. Similarly if A is a 3×3 matrix then the columns of A determine a parallelepiped.

Theorem 17.3. *If A is a 2×2 matrix the area of the parallelogram determined by the columns of A is $|\det A|$, the absolute value of the determinant.*

If A is a 3×3 matrix the volume of the parallelepiped determined by the columns of A is $|\det A|$, the absolute value of the determinant.

Proof. We just do the case when A is a 2×2 matrix.

Let's start with an easy case. Suppose that A is a diagonal matrix:

$$\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}.$$

Then the parallelogram determined by A is the rectangle with vertices $(0,0)$, $(a,0)$, $(0,d)$ and (a,d). This has sides of length $|a|$ and $|d|$ and the area is $|a||d| = |ad|$ which is the same as the absolute value of the determinant:

$$\begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} = ad.$$

Notice that this parallelogram is special as it is both a rectangle, one side is part of the x-axis and one side is part of the y-axis.

In general the area of a parallelogram is

$$\frac{1}{2} \times \text{base} \times \text{height}.$$

Imagine taking a parallelogram and first straightening one side so that it is part of the x-axis and then straightening the second side so that it is part of the y-axis. In the end you get a parallelogram as above, when the area and the absolute value of the determinant are the same.

How do the area and the determinant change when you make the second side so that it is part of the y-axis? Well the base of the rectangle and the parallelogram are the same and the height is unchanged, so that the area is unchanged. How does the determinant change? Straightening the parallelogram corresponds to applying a shear, a linear transformation corresponding to a matrix:

$$B = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$

where a is some scalar. The determinant of this matrix is 1 and so

$$\det(BA) = \det B \det A = \det A.$$

Thus the determinant is unchanged as well.

As we have equality after the shear we must have equality before. □

Example 17.4. *What is the volume of the parallelepiped with vertices at* $(1, 1, 1)$*,* $(2, 1, 1)$*,* $(3, 3, 1)$ *and* $(-1, -1, -1)$*?*

This the same as the volume of the parallelepiped with sides

$$\vec{c}_1 = (1, 0, 0) \qquad \vec{c}_2 = (2, 2, 0) \qquad \text{and} \qquad \vec{c}_3 = (-2, -2, -2).$$

Let A be the matrix with these columns:

$$\begin{pmatrix} 1 & 2 & -2 \\ 0 & 2 & -2 \\ 0 & 0 & -2 \end{pmatrix}$$

The determinant of A is

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

so that the volume is 4.

If

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

is a linear transformation it sends cubes to parallelepipeds. The volume of the cube is multiplied by the same scaling factor, the absolute value of the determinant of the corresponding matrix. Since any volume can be computed arbitrarily accurately by dicing into cubes, the linear transformation rescales volumes by the same factor, the absolute value of the determinant of the corresponding matrix.

17.4. **Efficiency.** Consider how long it takes to solve a system of equations. Assume that there are as many variables as unknowns. Then solving the system of equations is equivalent to solving the matrix equation

$$A\vec{x} = \vec{b},$$

where A is an $n \times n$ matrix. Obviously as n grows, the time it takes to solve these equations grows as well. Typically n can be quite large $n = 10^6$ is commonplace; $n = 10^9$ is also quite common. Given this, efficiency in computation is a serious issue. There is no point in devising elegant methods of solving equations, if it takes too long to implement them.

To figure out how long it will take to solve a system of equations, we have to decide on what we should consider a basic operation. A reasonable choice is to consider an operation on one entry of the matrix as one unit of time (in other words, we won't worry about issues of how big the numbers we are adding). Thus operating on one row of length n takes n units of time. Clearly the time it takes for the computer to perform one operation depends on the computer's power. Your laptop/desktop won't be nearly as fast as a commercial high speed computer (it is not that slow either). Let us somewhat arbitrarily say that we can perform 10^9 operations in one second.

How long does it take to perform Gaussian elimination? Ignoring permuting the rows, multiplying rows by a scalar to create ones and performing backwards substitution (all of which take a negliglible time), the point is that we need to eliminate a lot of matrix entries, that is turn them into zeroes. Each time we create a zero it takes n units of time.

To warm up, let us consider the case of a 4×4 matrix. In the first column we need to create three zeroes (the three entries below the top entry). In the second row we need to create two zeroes, in the third row one, and in the last row none. In total we need to create

$$3 + 2 + 1 = 6,$$

zeroes. Each time we do this it takes 4 units of time, and so it will take $6 \cdot 4 = 24$ units of time.

Now suppose that we have an $n \times n$ matrix. In the first row we need to create $n - 1$ zeroes, in the second row $n - 2$ and so on. In other words, we need to create

$$(n - 1) + (n - 2) + (n - 3) + \dots 3 + 2 + 1,$$

zeroes. It is easy to sum this if we reverse the order add the two sums together and divide by two,

$$n(n-1)/2.$$

Not surprisingly these are called triangular numbers. In fact the trick of reversing the order is the same as saying these two triangles, when put together makes a rectangle with sides $n-1$ and n. Now

$$n(n-1)/2 = n^2/2 - n/2.$$

In fact, since we are only really worried what happens when n is large, we can ignore the second term, since if n is large, then n is negligible in terms of n^2 (compare for example $n = 10^3$ with $n^2 = 10^6$). Put differently a $n-1 \times n$ rectangle is very close to a $n \times n$ square. Since each operation costs n units of time, it takes

$$n^3/2,$$

units of time to solve a system of linear equations using Gaussian elimination. If $n = 10^3$, then this is about

$$10^9/2$$

operations, which takes about

$$1/2$$

a second. If $n = 10^6$ however, then this is about

$$10^{18}/2$$

operations, which takes about

$$10^9/2$$

seconds, which is quite long, just under 32 years! Let us consider what happens if we use Gauss-Jordan elimination. The answer is easy. We need to create twice as many zeroes (those above the main diagonal). It would take then

$$n^3.$$

operations. Around about 64 years, if $n = 10^6$.

How about computing the inverse and then using the formula,

$$\vec{x} = A^{-1}\vec{b}?$$

The time taken to multiply matrices is negligible, and so it is only a question of how to compute the inverse. Well the best method to compute the inverse is to use Gauss-Jordan elimination. The only difference is that we operate on a matrix whose rows are of length $2n$ not just n. Thus it takes

$$2n^3,$$

operations. About 128 years, if $n = 10^6$.

Finally how about Cramer's rule? Well part of the rule says to compute the determinant. We could compute the determinant by using Gaussian elimination, but if we did that, why use Cramer's rule at all? If we use the method of cofactors then we have to make $n!$ operations, which is simply enormous. For example if $n = 10^3$, then $n!$ is approximately 10^{2549} seconds. To get a sense of how big this number is, the number of atoms in the universe is estimated to be no more than 10^{80}. The age of the universe is at most 10 billion years, about 10^{17} seconds.