Dot Products, Cross Products, Determinants, and Volume

June 2025

1 Introduction

In courses in vector calculus, a lot of statements about areas and volumes in terms of dot products and cross products are stated, but not proven. This article proves all of these assertions. These facts can be summarized by using the concept of k-volume in $\mathbf{R}^n$. The quantity $\det(A^T A)$ plays an important role. The singular value decomposition leads to a better understanding of where this quantity comes from, in terms of the singular values of the matrix.

2 A Review of Dot Products

The *dot product* of two vectors $\mathbf{a} = \langle a_1, a_2 \rangle$ and $\mathbf{b} = \langle b_1, b_2 \rangle$ in $\mathbf{R}^2$ is defined to be the number $a_1 b_1 + a_2 b_2$. Similarly, the *dot product* of two vectors $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$ in $\mathbf{R}^3$ is defined to be the number $a_1 b_1 + a_2 b_2 + a_3 b_3$. The *length*, or magnitude, of a vector $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ is defined to be $\sqrt{a_1^2 + a_2^2 + a_3^2}$. It is written as $|\mathbf{a}|$. This is the distance from the origin to the point (a_1, a_2, a_3). The length of this vector can be expressed in terms of the dot product. This is because

$$\mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 + a_3^2,$$

so

$$\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2.$$

There is a useful geometric interpretation of the dot product. Let θ be the smaller of the two angles between $\mathbf{a}$ and $\mathbf{b}$. Then

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta.$$

There is a simple proof using the Law of Cosines. Form a triangle with vertices $(0,0,0)$, (a_1, a_2, a_3), and (b_1, b_2, b_3). The Law of Cosines says that

$$|\mathbf{b} - \mathbf{a}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2|\mathbf{a}||\mathbf{b}| \cos \theta.$$

We also have $|\mathbf{b} - \mathbf{a}|^2 = (\mathbf{b} - \mathbf{a}) \cdot (\mathbf{b} - \mathbf{a}) = \mathbf{b} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{a} - 2\mathbf{b} \cdot \mathbf{a}$, as well as $|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a}$ and $|\mathbf{b}|^2 = \mathbf{b} \cdot \mathbf{b}$. Substituting all this in, we get

$$|\mathbf{a}|^2 + |\mathbf{b}|^2 - 2\mathbf{a}\mathbf{b} = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2|\mathbf{a}||\mathbf{b}| \cos \theta.$$

This simplifies to

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta.$$

As a consequence, we see that $\mathbf{a} \cdot \mathbf{b} = 0$ if and only if $\theta = \pi/2$, that is, if $\mathbf{a}$ and $\mathbf{b}$ are perpendicular. Such vectors are also said to be *orthogonal*.

3 A Review of Determinants

There is a general theory of n by n determinants. However, 1 by 1, 2 by 2, and 3 by 3 determinants are defined by relatively simple formulas. The determinant of the 1 by 1 matrix

$$A = [a]$$

is defined to be a. The determinant of the 2 by 2 matrix

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

is defined to be the number $ad - bc$. The determinant of the 3 by 3 matrix

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

is defined to be the number

$$a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1).$$

The connection between these determinants is that the 3 by 3 determinant is equal to

$$a_1 \det \begin{bmatrix} b_2 & b_3 \\ c_2 & c_3 \end{bmatrix} - a_2 \det \begin{bmatrix} b_1 & b_3 \\ c_1 & c_3 \end{bmatrix} + a_3 \det \begin{bmatrix} b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}$$

So the 3 by 3 determinant is defined using 2 by 2 determinants.

Determinants have a number of useful properties. One of these is that swapping two rows reverses the sign of the determinant. This is not too hard to see for a 3 by 3 determinant. For example, swapping rows 1 and 3 leads to the matrix

$$A' = \begin{bmatrix} c_1 & c_2 & c_3 \\ b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \end{bmatrix}$$

and the determinant of A' is

$$c_1(b_2b_3 - b_3a_2) - c_2(b_1a_3 - b_3a_1) + c_3(b_1a_2 - b_2a_1),$$

and it is easy to check that this equals $-\det A$.

Another useful property is that the determinant of a matrix is equal to the determinant of the transpose of the matrix.

Another useful property is that the determinant is multiplicative in that

$$\det(AB) = \det(A) \cdot \det(B).$$

Another useful property is that the determinant can be calculated using cofactor expansion across any row. Since $\det(A^T) = \det(A)$, this implies that the determinant can be calculated using cofactor expansion down any column.

We will learn how the absolute value of the determinant of a matrix is related to areas and volumes, and we will use all of these properties of determinants.

4 Review of the Cross Product

Unlike the dot product, the cross product is only defined for vectors in $\mathbf{R}^3$. Suppose that $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$ are vectors in $\mathbf{R}^3$. The cross product of $\mathbf{a}$ and $\mathbf{b}$, denoted by $\mathbf{a} \times \mathbf{b}$, is defined to be the vector

$$\langle a_2 b_3 - a_3 b_2, -(a_1 b_3 - a_3 b_1), a_1 b_2 - a_2 b_1 \rangle.$$

The cross product has some interesting properties. First of all, $\mathbf{a} \times \mathbf{a} = \mathbf{0}$. This is immediately seen by letting $b_1 = a_1$, $b_2 = a_2$, and $b_3 = a_3$, and observing that each component is 0. Another property is the anti-commutativity property: $-\mathbf{v} \times \mathbf{u} = \mathbf{u} \times \mathbf{v}$. The cross product also satisfies a distributive property: $\mathbf{u} \times (c\mathbf{v} + d\mathbf{w}) = c(\mathbf{u} \times \mathbf{v}) + d(\mathbf{u} \times \mathbf{w})$. These properties are easy to check by explicitly calculating both sides and checking that they are equal.

One very useful fact about the cross product is that $\mathbf{u} \times \mathbf{v}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$. This follows from a direct computation:

$$\begin{aligned} \mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) &= u_1(u_2 v_3 - u_3 v_2) - u_2(u_1 v_3 - u_3 v_1) + u_3(u_1 v_2 - u_2 v_1) \\ &= u_1 u_2 v_3 - u_2 u_1 v_3 - u_1 u_3 v_2 + u_3 u_1 v_2 + u_2 u_3 v_1 - u_3 u_2 v_1 = 0. \end{aligned}$$

A similar computation shows that $\mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0$. Alternatively, we have

$$\mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot (-\mathbf{v} \times \mathbf{u}) = -\mathbf{v} \cdot (\mathbf{v} \times \mathbf{u}) = 0.$$

The equality $\mathbf{v} \cdot (\mathbf{v} \times \mathbf{u}) = 0$ comes from the equality $\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v})$ by replacing $\mathbf{u}$ by $\mathbf{v}$ and $\mathbf{v}$ by $\mathbf{u}$.

5 Area of a Parallelogram in $\mathbf{R}^2$

Suppose that $\mathbf{u} = \langle a, b \rangle$ and $\mathbf{v} = \langle c, d \rangle$ are two vectors in $\mathbf{R}^2$. Consider the parallelogram with these vectors as edges. There is a simple formula for the area of the parallelogram. To see this, recall that the area of a parallelogram is the base of the parallelogram times the height. We can use the line segment from 0 to (a, b) as the base. To find the height, drop a perpendicular from $\mathbf{v}$ to the line spanned by $\mathbf{u}$. The height of the parallelogram is equal to $|\mathbf{v}||\sin\theta|$. So the area of the parallelogram is equal to

$$|\mathbf{u}||\mathbf{v}||\sin\theta|.$$

We can express this in terms of dot products. The square of the area is $|\mathbf{u}|^2|\mathbf{v}|^2\sin^2\theta$. Since $\sin^2\theta = 1 - \cos^2\theta$, the square of the area is equal to $|\mathbf{u}|^2|\mathbf{v}|^2(1-\cos^2\theta) = |\mathbf{u}|^2|\mathbf{v}|^2 - |\mathbf{u}|^2|\mathbf{v}|^2\cos^2\theta$. Since $(\mathbf{u}\cdot\mathbf{v})^2 = |\mathbf{u}|^2|\mathbf{v}|^2\cos^2\theta$, the square of the area of the parallelogram is equal to

$$\begin{aligned}
|\mathbf{u}|^2|\mathbf{v}|^2 - (\mathbf{u}\cdot\mathbf{v})^2 &= \det\begin{bmatrix}\mathbf{u}\cdot\mathbf{u} & \mathbf{v}\cdot\mathbf{v}\\ \mathbf{v}\cdot\mathbf{u} & \mathbf{v}\cdot\mathbf{v}\end{bmatrix}\\
&= (a^2+b^2)(c^2+d^2) - (ac+bd)^2\\
&= a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 - (a^2c^2 + b^2d^2 + 2abcd)\\
&= a^2d^2 - 2abcd + b^2c^2\\
&= (ad-bc)^2.
\end{aligned}$$

Therefore, the area of the parallelogram is $|ad-bc|$. This can be interpreted in terms of the determinant. Let A be the matrix whose columns are $\mathbf{u}$ and $\mathbf{v}$:

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

Then $\det A = ad - bc$. So the area of the parallelogram with edges $\mathbf{u}$ and $\mathbf{v}$ is the absolute value of the determinant of the matrix whose columns are $\mathbf{u}$ and $\mathbf{v}$.

6 Area of a Parallelogram in $\mathbf{R}^3$

Suppose that $\mathbf{a}$ and $\mathbf{b}$ are two vectors in $\mathbf{R}^3$. Consider the parallelogram with these vectors as edges. There is a simple formula for the area of the parallelogram. To see this, recall that the area of a parallelogram is the base of the parallelogram times the height. We can use the line segment from 0 to (a_1, a_2, a_3) as the base. To find the height, drop a perpendicular from $\mathbf{b}$ to the line spanned by $\mathbf{a}$. Then the height is equal to $|\mathbf{b}||\sin\theta|$. So the area of the parallelogram is equal to

$$|\mathbf{a}||\mathbf{b}||\sin\theta|.$$

We can express this in terms of dot products. The square of the area is $|\mathbf{a}|^2|\mathbf{b}|^2\sin^2(\theta)$. Now $(\mathbf{a}\cdot\mathbf{b})^2 = |\mathbf{a}|^2|\mathbf{b}|^2\cos^2\theta$, so the square of the area of the parallelogram is equal to

$$|\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a}\cdot\mathbf{b})^2 = \det\begin{bmatrix}\mathbf{a}\cdot\mathbf{a} & \mathbf{a}\cdot\mathbf{b}\\ \mathbf{b}\cdot\mathbf{a} & \mathbf{b}\cdot\mathbf{b}\end{bmatrix}$$

One can also check that

$$|\mathbf{a}|^2|\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2 = |\mathbf{a} \times \mathbf{b}|^2.$$

Since the area of a parallelogram is non-negative, we see that the area of the parallelogram is equal to the magnitude of $\mathbf{a} \times \mathbf{b}$. This gives an interpretation for the magnitude of the cross product of two vectors.

7 The Volume of a Parallelepiped

The 3-dimensional analogue of a parallelogram is a parallelepiped. Given three vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbf{R}^3$, the parallelopiped with edges $\mathbf{u}, \mathbf{v}, \mathbf{w}$ is the set

$$\{\mathbf{x} = c_1\mathbf{u} + c_2\mathbf{v} + c_3\mathbf{w} \mid 0 \leq c_1, c_2, c_3 \leq 1\}$$

Just as we have found a formula for the area of a parallelogram, we want a formula for the volume of the parallelopiped. The volume of this is equal to the area of the base, times the height. Let's use the base formed by $\mathbf{u}$ and $\mathbf{v}$. We know that the area of this parallelogram is equal to $|\mathbf{u} \times \mathbf{v}|$. We just need the height. For this, drop an altitude from $\mathbf{w}$ to the plane spanned by $\mathbf{u}$ and $\mathbf{v}$. The vector $\mathbf{w}$ makes an angle θ with $\mathbf{u} \times \mathbf{v}$. The height is $|\mathbf{w}||\cos\theta|$. Therefore, the volume of the parallelopiped is

$$|\mathbf{u} \times \mathbf{v}||\mathbf{w}||\cos\theta|,$$

which is equal to

$$|(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}|.$$

The number $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$ is sometimes called the *scalar triple product*. So the volume is the absolute value of this scalar triple product.

This has a useful interpretation in terms of determinants. It is equal to the determinant of the matrix whose rows are $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$. This can be seen by directly computing both quantities. We have

$$\mathbf{u} \times \mathbf{v} = \langle u_2v_3 - u_3v_2, -(u_1v_3 - u_3v_1), u_1v_2 - u_2v_1\rangle,$$

so

$$(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = w_1(u_2v_3 - u_3v_2) - w_2(u_1v_3 - u_3v_1) + w_3(u_1v_2 - u_2v_1).$$

This is exactly what is obtained by calculating the determinant

$$\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

via cofactor expansion across the third row. This shows that the volume of the parallelepiped is equal to the absolute value of the determinant of the matrix whose rows are $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$. Since the determinant of a matrix is equal to the determinant of the transpose of the matrix, we can also say that the volume of the parallelepiped is equal to the absolute value of the determinant of the matrix whose columns are $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$.

8 Orthogonal Matrices

A square matrix Q is said to be *orthogonal* if it is invertible and its inverse is equal to its transpose: $Q^T = Q^{-1}$. Multiplying both sides by Q on the right, we see that $Q^TQ = I$. The matrix transformation arising from an orthogonal matrix is said to be an *orthogonal transformation*. Orthogonal transformations are special in many respects. First, they don't distort length: $Q\mathbf{v}$ has the same length as $\mathbf{v}$. The easiest way to see this is to observe that

$$|Q\mathbf{v}|^2 = (Q\mathbf{v}) \cdot (Q\mathbf{v}) = (Q\mathbf{v})^T(Q\mathbf{v}) = \mathbf{v}^TQ^TQ\mathbf{v} = \mathbf{v}^T\mathbf{v} = |\mathbf{v}|^2$$

Taking square roots of the equation $|Q\mathbf{v}|^2 = |\mathbf{v}|^2$ gives $|Q\mathbf{v}| = |\mathbf{v}|$.

More generally, orthogonal matrices preserve the dot product. This is a vague way of saying the following: The dot product of $Q\mathbf{v}$ and $Q\mathbf{w}$ is equal to the dot product of $\mathbf{v}$ and $\mathbf{w}$. This is essentially the computation we did above:

$$(Q\mathbf{v}) \cdot (Q\mathbf{w}) = (Q\mathbf{v})^T(Q\mathbf{w}) = \mathbf{v}^TQ^TQ\mathbf{w} = \mathbf{v}^T\mathbf{w} = \mathbf{v} \cdot \mathbf{w}.$$

As a consequence of this, orthogonal transformations also preserve angles, meaning that if $\mathbf{v}$ and $\mathbf{w}$ make an angle of θ, then so do $Q\mathbf{v}$ and $Q\mathbf{w}$. Suppose that the angle made by $Q\mathbf{v}$ and $Q\mathbf{w}$ is α. We will show that $\theta = \alpha$ We have

$$\theta = \cos^{-1}\left(\frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}||\mathbf{w}|}\right)$$

and

$$\alpha = \cos^{-1}\left(\frac{Q\mathbf{v}\cdot Q\mathbf{w}}{|Q\mathbf{v}||Q\mathbf{w}|}\right)$$

Since $Q\mathbf{v}\cdot Q\mathbf{w} = \mathbf{v}\cdot\mathbf{w}$, $|Q\mathbf{v}| = |\mathbf{v}|$, and $|Q\mathbf{w}| = |\mathbf{w}|$,

$$\frac{Q\mathbf{v}\cdot Q\mathbf{w}}{|Q\mathbf{v}||Q\mathbf{w}|} = \frac{\mathbf{v}\cdot\mathbf{w}}{|\mathbf{v}||\mathbf{w}|}$$

Taking inverse cosines of both sides, we see that $\theta = \alpha$.

Orthogonal matrices have determinant either 1 or -1. This can be seen by calculating:

$$1 = \det(I) = \det(Q^TQ) = \det(Q^T)\det(Q) = \det(Q)^2.$$

An important fact about 2 by 2 and 3 by 3 orthogonal matrices with determinant 1 is that they are rotation matrices. Every 2 by 2 matrix rotates the plane by a certain angle, and every 3 by 3 matrix rotates the space by a certain angle about a certain axis.

9 Which way does the cross product point?

Suppose that $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbf{R}^3$, and let $\mathbf{w} = \mathbf{u}\times\mathbf{v}$. We know that $\mathbf{w}$ is orthogonal to $\mathbf{u}$ and $\mathbf{v}$, but this still doesn't determine which way $\mathbf{w}$ points. For example, suppose $\mathbf{u} = \mathbf{e}_1$ and $\mathbf{v} = \mathbf{e}_2$. Does $\mathbf{w}$ point along $\mathbf{e}_3$ or in the opposite direction of $\mathbf{e}_3$. It is easy to check that in this case $\mathbf{w} = \mathbf{e}_3$, but what happens in general? This is what we will answer in this section.

Let A be the matrix whose rows are $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ (in this order). Then $\det A \geq 0$. This can be seen by using the scalar triple product:

$$\det A = (\mathbf{u}\times\mathbf{v})\cdot(\mathbf{u}\times\mathbf{v})$$

which is non-negative because it is the dot product of a vector with itself.

We will determine the direction of $\mathbf{w} = \mathbf{u}\times\mathbf{v}$. First suppose that $\mathbf{u}$ and $\mathbf{v}$ are orthonormal. Then $\mathbf{w}$ has length 1 so the matrix whose columns are $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ is orthogonal. Furthermore, its determinant is positive. Therefore, it has determinant 1 and since it is a 3 by 3 matrix, it is a rotation about some axis.

What if $\mathbf{u}$ and $\mathbf{v}$ are orthogonal but not necessarily unit length? Then rotate space so that $\mathbf{e}_1$ and $\mathbf{e}_2$ point in the same direction as $\mathbf{u}$ and $\mathbf{v}$. The cross product $\mathbf{u}\times\mathbf{v}$ points in the same direction as $\mathbf{e}_3$.

Now we handle the general case where we are given two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{R}^3$ that are not necessarily orthogonal. The vector $\mathbf{w}_2 = \mathbf{v} - \mathrm{proj}_{\mathbf{u}}(\mathbf{v})$ is orthogonal to $\mathbf{u}$.

We have

$$\mathbf{u} \times \mathbf{w}_2 = \mathbf{u} \times (\mathbf{v} - \frac{\mathbf{v} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u}) = \mathbf{u} \times \mathbf{v} - \frac{\mathbf{v} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}(\mathbf{u} \times \mathbf{u}) = \mathbf{u} \times \mathbf{v}.$$

We arrive at how to tell the direction of $\mathbf{u} \times \mathbf{v}$. Project $\mathbf{v}$ onto the line along $\mathbf{u}$ and let $\mathbf{w}_2 = \mathbf{v} - \mathrm{proj}_{\mathbf{u}}\mathbf{v}$. Rotate space so that $\mathbf{e}_1$ points in the direction of $\mathbf{u}$ and $\mathbf{e}_2$ points in the direction of $\mathbf{w}_2$. Then $\mathbf{e}_3$ points in the direction of $\mathbf{u} \times \mathbf{v}$.

10 k-Volume in $\mathbf{R}^n$

We have seen that the formulas for the areas of parallelograms in $\mathbf{R}^2$ and volumes of parallelepipeds in $\mathbf{R}^3$ are both given by putting the edge vectors as rows or columns of a matrix and finding the absolute value of the determinant of the matrix. We have found a formula for the area of a parallelogram in $\mathbf{R}^3$, but this is not as simple as calculating a determinant, because the edges will not form a square matrix. Going back to parallelograms in $\mathbf{R}^3$, let A be the matrix whose columns are $\mathbf{a}$ and $\mathbf{b}$. This is a 3 by 2 matrix:

$$A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}$$

Then

$$\begin{aligned} A^T A &= \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix} \\ &= \begin{bmatrix} \sum a_i^2 & \sum a_i b_i \\ \sum b_i a_i & \sum b_i^2 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{a} \cdot \mathbf{a} & \mathbf{a} \cdot \mathbf{b} \\ \mathbf{b} \cdot \mathbf{a} & \mathbf{b} \cdot \mathbf{b} \end{bmatrix} \end{aligned}$$

Therefore, the area of the parallelogram is equal to $\sqrt{\det(A^T A)}$.

What about if A is a 2 by 2 or 3 by 3 matrix? Then A^T is also a square matrix, and by properties of determinants,

$$\det(A^T A) = \det(A^T)\det(A) = \det(A) \cdot \det(A) = \det(A)^2.$$

Therefore,

$$\sqrt{\det(A^T A)} = \sqrt{\det(A)^2} = |\det(A)|.$$

So in all three cases, the resulting volume was equal to $\sqrt{\det(A^T A)}$.

This even works for a 1 by 3 or 1 by 2 matrix. Suppose that $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and let A be the matrix with a single column which is the vector $\mathbf{a}$:

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

Then $A^T A$ is the 1 by 1 matrix $\begin{bmatrix} a_1^2 + a_2^2 + a_3^2 \end{bmatrix}$ and $\sqrt{\det(A^T A)} = \sqrt{a_1^2 + a_2^2 + a_3^2}$, which is the length of $\mathbf{a}$!

To summarize:

- The length of a vector $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ is equal to $\sqrt{\det A^T A}$ where A is the 3 by 1 matrix whose column is $\mathbf{a}$.

- The area of the parallelogram in $\mathbf{R}^2$ with edges $\mathbf{a}$ and $\mathbf{b}$ is equal to $\sqrt{\det A^T A}$ where A is the 2 by 2 matrix whose columns are $\mathbf{a}$ and $\mathbf{b}$.

- The area of the parallelogram with edges $\mathbf{a}$ and $\mathbf{b}$ is equal to $\sqrt{\det A^T A}$ where A is the 3 by 2 matrix whose columns are $\mathbf{a}$ and $\mathbf{b}$.

- The volume of the parallelepiped with edges $\mathbf{u}, \mathbf{v}, \mathbf{w}$ is equal to $\sqrt{\det A^T A}$ where A is the 3 by 3 matrix whose columns are $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

One way to summarize what we have done is to first define the concept of "k-parallelepiped" and then the concept of "k-volume" in $\mathbf{R}^n$ for an integer $1 \leq k \leq n$. Let $\mathbf{v}_1, \ldots \mathbf{v}_k$ be vectors in $\mathbf{R}^n$. The *k-parallelepiped* formed by $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is defined to be

$$\mathcal{P}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \{\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \mid 0 \leq c_i \leq 1, i = 1, \ldots, k\}.$$

For example, the 1-parallelepiped formed by a single vector $\mathbf{v} = \langle v_1, \ldots, v_n \rangle$ in $\mathbf{R}^n$ is the line segment from the origin to the point $(v_1, \ldots, v_n)$. Now we define the k-volume of these k-parallelepipeds. The 1-*volume* of a 1-parallelepiped in $\mathbf{R}^n$ is defined to be its *length.* The 2-*volume* of a 2-parallelepiped is its *area.* The 3-*volume* of a 3-parallelepiped is its *volume.*

We have found that the k-volume of the k-parallelepiped with edges $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in $\mathbf{R}^n$ is $\sqrt{\det A^T A}$ where A is the n by k matrix whose columns are $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. But where does this quantity come from? It isn't very intuitive. The most transparent explanation that I can think of comes from the singular value decomposition, which we discuss next.

11 The Singular Value Decomposition

We will only work with matrices A that are n by k where $k \leq n$. The singular value decomposition theorem says that for any n by k matrix A, there is an orthogonal n by n matrix U, an orthogonal k by k matrix V, and a "diagonal" n by k matrix Σ such that

$$A = U \Sigma V^T.$$

The word "diagonal" is in quotation marks because Σ is not a square matrix if $k \neq n$. What is meant is that the submatrix obtained from the first k rows is a diagonal matrix, and the remaining $n-k$ rows at the bottom are all zero rows. The (i,i)-entries from $i = 1$ to $i = k$ are called the *singular values* of A and are usually denoted by σ_i:

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \sigma_k \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix}$$

This is an advanced theorem in linear algebra, but we can still use it. Using a singular value decomposition as above, we calculate

$$A^T A = (U \Sigma V^T)^T (U \Sigma V^T) = V \Sigma^T U^T U \Sigma V^T = V \Sigma^T \Sigma V^T.$$

Now $\Sigma^T\Sigma$ is the k by k matrix whose entries are the squares of the singular values:

$$\Sigma^T\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_n^2 \end{bmatrix}$$

Therefore,

$$\begin{aligned} \det(A^TA) &= \det(V\Sigma^T\Sigma V^T) \\ &= \det(V)\det(\Sigma^T\Sigma)\det(V^{-1}) \\ &= \det(\Sigma^T\Sigma) \\ &= \sigma_1^2\cdots\sigma_k^2 \end{aligned}$$

Since the σ_i's are all non-negative,

$$\sqrt{\det(A^TA)} = \sqrt{\sigma_1^2\cdots\sigma_n^2} = \sigma_1\cdots\sigma_k.$$

So, $\sqrt{\det(A^TA)}$ can be interpreted as the product of the singular values of A.

The SVD implies that orthogonal transformations don't affect k-volume. More precisely, suppose that $\mathbf{v}_1,\dots,\mathbf{v}_k$ are vectors in $\mathbf{R}^n$ and let Q be an orthogonal matrix. Then the k-volume of $\mathcal{P}(\mathbf{v}_1,\dots,\mathbf{v}_k)$ is equal to the k-volume of $\mathcal{P}(Q\mathbf{v}_1,\dots,Q\mathbf{v}_k)$. To prove this, let A be the matrix whose jth column is $\mathbf{v}_j$. Then if $A = U\Sigma V^T$ is an SVD of A, the k-volume is the product of the diagonal entries of Σ. The matrix whose jth column is $Q\mathbf{v}_j$ is just QA, which has an SVD of $(QU)\Sigma V^T$. The k-volume is the product of the diagonal entries of Σ. Therefore, Q does not affect the k-volume.

We can also see this using the characterization of k-volume in terms of the determinant. We have

$$(QA)^T(QA) = A^TQ^TQA = A^TA.$$

Therefore, $\sqrt{\det((QA)^T(QA))} = \sqrt{\det(A^TA)}$.

*****

Here is an example. Suppose that $\mathbf{v}_1 = \langle 4, 11, 14\rangle$ and $\mathbf{v}_2 = \langle 8, 7, -2\rangle$. One way to find the area of the parallelogram with these edges is to calculate the cross product:

$$\mathbf{u} \times \mathbf{v} = \langle -120, 120, -60\rangle$$

The magnitude of this vector is $\sqrt{(-120)^2 + 120^2 + (-60)^2} = \sqrt{32400} = 180$. Another way is to form the matrix whose columns are $\mathbf{v}_1$ and $\mathbf{v}_2$,

$$A = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix}$$

and then calculate a singular value decomposition

$$A = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix}$$

The matrix A takes $\mathbf{e}_1$ to $\mathbf{v}_1$ and $\mathbf{e}_2$ to $\mathbf{v}_2$. The SVD breaks down the process into three steps, one for each matrix in the decomposition. The first matrix, V^T, takes $\mathbf{e}_1$ to $\langle 3\sqrt{10}, 1/\sqrt{10}\rangle$ and $\mathbf{e}_2$ to $\langle 1/\sqrt{10}, -3/\sqrt{10}\rangle$. To understand the effect of Σ on these two vectors, it may help to decompose Σ further into

$$\Sigma = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Call the matrices on the right side of the equality S and T, so that $\Sigma = ST$. The matrix T takes $\langle 3\sqrt{10}, 1/\sqrt{10}\rangle$ to $\langle 3\sqrt{10}, 1/\sqrt{10}, 0\rangle$ and $\langle 1/\sqrt{10}, -3/\sqrt{10}\rangle$ to $\langle 1/\sqrt{10}, -3/\sqrt{10}, 0\rangle$. The matrix embeds $\mathbf{R}^2$ to the plane $z = 0$ in $\mathbf{R}^3$. T does not distort volumes.

However, the matrix S does. It stretches the x-axis by a factor of $6\sqrt{10}$ and stretches the y-axis by a factor of $3\sqrt{10}$, while keeping the z-axis fixed. This matrix sends $\langle 3\sqrt{10}, 1/\sqrt{10}, 0\rangle$ to $\langle 18, 3, 0\rangle$ and $\langle 6, -9, 0, \rangle$, respectively. S does distort angles and volumes.

The final matrix, U, is an orthogonal transformation, taking $\langle 18, 3, 0\rangle$ to $\mathbf{v}_1$ and $\langle 6, -9, 0\rangle$ to $\mathbf{v}_2$.

The product of the singular values of A is $(6\sqrt{10})(3\sqrt{10}) = 180$, as expected. Throughout the entire process, the only distortion occurs at the step where the plane's axes are scaled by factors of $\sigma_1 = 6\sqrt{10}$ and $\sigma_2 = 3\sqrt{10}$, under the transformation from the matrix S, causing the area of the parallelogram to be multiplied by a factor of $\sigma_1\sigma_2$.