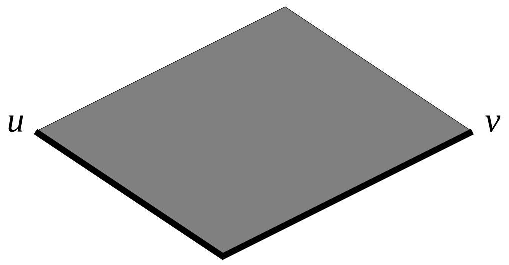
Mathematics 307—December 6, 1995

Determinants, areas, and volumes

If u and v are any two $2D$ vectors, they span a parallelogram. The geometric meaning of determinants is related to the area of this parallelogram.



To understand the relationship exactly, we must first understand what an orientation of the paralellogram is. An orientation is determined by the order in which the vectors u and v are specified. In the picture above, the parallelogram is oriented positively because the smallest angle necessary to rotate u towards v is positive. If it is negative the orientation is negative. If u and v lie on the same line, the parallelogram's orientation is undefined.

Proposition. *The area of the oriented parallelogram determined by u and v is equal to the determinant of the 2×2 matrix*

$$M_{u,v} = [\, u \quad v \,]$$

whose columns are the coordinates of the vectors u and v.

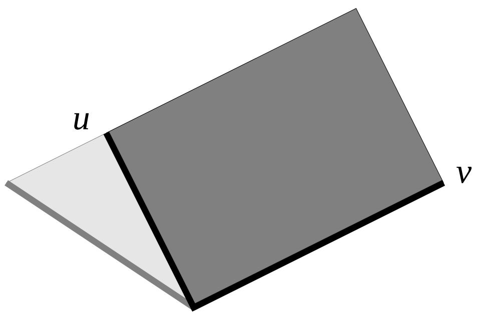
In particular, the determinant vanishes precisely when u and v lie in the same line, or equivalently when u and v are scalar multiples of each other.

The proof of the Proposition is essentially geometrical. We will start with the original pair u and v and transform them by a succession of elementary operations in which we can understand exactly how bothe the area and the determinant is modified. We begin with

Proposition. *If A and B are two 2×2 matrices then* $\det(AB) = \det(A)\det(B)$.

This is a messy but straightforward calculation.

Step (1) We shear the vector u along v so as to turn the parallelogram into a rectangle.



The area remains the same, since a shear does not change the height of the parallelogram spanned by u and v. The vector u is changed to $u_{\mathrm{new}} = u + cv$ for some constant c, and v_{new} isn't changed. So we change $M_{u,v}$ to

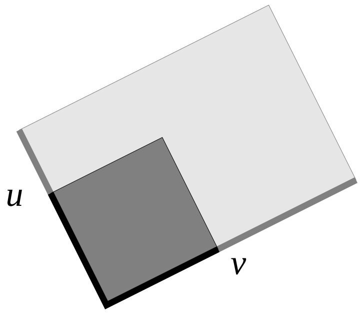
$$[\,u + cv \quad v\,] = [\,u \quad v\,] \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}$$

Because of the multiplication rule for determinants, the determinant doesn't change either.

We can find c explicitly:

$$(u + cv) \bullet v = 0, \quad c = \frac{u \bullet v}{v \bullet v}\,.$$

Step (2) We scale u and v to make them of unit length.

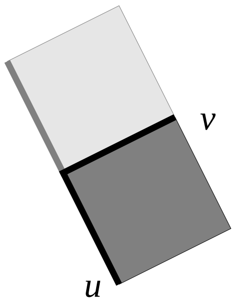


Areas get multiplied by the product of the two scaling factors a and b, which happen to be the inverse of the lengths of u and v. But the new matrix is

$$[\,au \quad bv\,] = [\,u \quad v\,] \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

so the determinant also gets multiplied by ab.

Step (3) We can scale u by -1 if necessary to get the orientation positive.

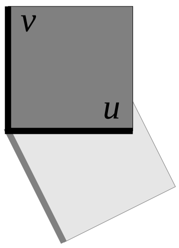


The area changes sign, and the matrix gets multiplied on the right by

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

These last two steps can by combined into a single scaling.

Step (4) We rotate u and v so v lies along the x-axis.



We don't change the area and the matrix gets multiplied by a rotation matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

so the determinant doesn't change either.

Step (4) At the end of this process, u and v span a unit square, and v lies along the x-axis. Tyhere are two possibilities: either $u = (0,1)$ or $v = (0,-1)$. Here both the oriented area and the determinant are both either ± 1, and in any event they agree.

Step (5) So we have changed the pair u, v into a pair where the area and the determinant agree, and we have done this in such a way that the area and the determinant chnage equally as we proceed. Therefore the area and the determinant were equal to begin with.

There is one simplification we might have made. We could have avoided the possibility of a negative orientation at the end if we had allowed a negative scale factor of u in Step (2). What the description above shows if we start with any matrix M we can find a vertical shear N, a diagonal matrix A, and a rotation matrix K such that $MNAK = I$. Multiplying by inverses we get

$$M = K^{-1}A^{-1}N^{-1} .$$

The inverse of a rotation is a rotation, etc. Therefore we have incidentally proved along the way:

Proposition. *If M is any matrix then we can write*

$$M = KAN$$

where N is the matrix of a vertical shear, A diagonal, K a rotation.

There is one problem we have overlooked. The procedure above works only if the vectors we start with are linearly independent. Otherwise we find ourselves trying to divide by 0. In case they are dependent, the Proposition remains valid, but a bit of fiddling is necessary in the argument to allow for vectors which are 0 after the initial shear.

What happens in $3D$

The whole process works also in $3D$, once we understand how to perform the original shear. So we are given u, v, w and we want to find a shear that takes these into three mutually perpendicular vectors u_*, v_*, w_*. We can do this in stages: (1) first of all u stays the same:

$$u_* = u$$

(2) Shear v by a multiple of u_* to get V_* perpendicular to u_*.

$$v_* = v - cu_*$$

where cu_* is the perpendicular projection of v_* onto the line through u_*:

$$c = \frac{v \bullet u_*}{u_* \bullet u_*}$$

(3) Shear w_* along the plane containing u_* and v_*:

$$w_* = w - au_* - bv_*$$

and now $au_* + bv_*$ is the perpendicular projection of w onto this plane so

$$a = \frac{w \bullet u_*}{u_* \bullet u_*}, \quad b = \frac{w \bullet v_*}{v_* \bullet v_*}$$

Thgis process works in any number of dimensions to change a given set of n linearly independent vectors to a set of mutually perpendicular ones. If we start with u_i for $i = 1$ to n we get $u_{*,i}$ where

$$\begin{aligned}
u_{*,1} &= u_1 \\
u_{*,2} &= u_2 - \frac{u_2 \bullet u_{*,1}}{u_{*,1} \bullet u_{*,1}}\, u_{*,1} \\
u_{*,3} &= u_3 - \frac{u_3 \bullet u_{*,1}}{u_{*,1} \bullet u_{*,1}}\, u_{*,1} - \frac{u_3 \bullet u_{*,2}}{u_{*,2} \bullet u_{*,2}}\, u_{*,2} \\
&\ldots \\
u_{*,n} &= u_n - \frac{u_n \bullet u_{*,1}}{u_{*,1} \bullet u_{*,1}}\, u_{*,1} - \frac{u_n \bullet u_{*,2}}{u_{*,2} \bullet u_{*,2}}\, u_{*,2} - \cdots - \frac{u_n \bullet u_{*,n-1}}{u_{*,n-1} \bullet u_{*,n-1}}\, u_{*,n-1}
\end{aligned}$$

This process is called **Gram-Schmid orthogonalization**. Again, it only works completely when the original vectors u_i are linearly independent.