This tutorial teaches you about 2×2 matrices: transforming vectors, composition, inversion, and their geometric meanings.

1 What Is A Matrix?

A *matrix* is a rectangular grid of numbers. An $n \times m$ matrix is one with n rows and m columns. For example, here is a specific 2×3 matrix:

$$M = \begin{bmatrix} 1 & -4.2 & \pi \\ 0 & 2 & 2.2 \end{bmatrix}.$$

In math, it is common to index the rows using $1 \leq i \leq n$ and the columns using $1 \leq j \leq m$. The entry of M in the ith row and jth column is denoted $M_{i,j}$ or simply M_{ij} if that's clear enough:

$$M = \begin{bmatrix} M_{1,1} & M_{1,2} & M_{1,3} \\ M_{2,1} & M_{2,2} & M_{2,3} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \end{bmatrix}.$$

In our course, because we're using C, we instead index from 0:

$$M = \begin{bmatrix} M_{00} & M_{01} & M_{02} \\ M_{10} & M_{11} & M_{12} \end{bmatrix}.$$

In an earlier tutorial, we learned about vectors. A n-dimensional vector (within a standardized coordinate system) is essentially the same thing as an $n \times 1$ matrix. Because there is only one column, the column index is always 0, and it is usually dropped, along with the "$\vec{}$":

$$\vec{v} = \begin{bmatrix} \vec{v}_{0,0} \\ \vec{v}_{1,0} \\ \vdots \\ \vec{v}_{n-1,0} \end{bmatrix} = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{n-1} \end{bmatrix}.$$

In this tutorial, we work exclusively with 2×2 matrices and 2×1 matrices. They are important to us because they form a convenient framework for 2-dimensional geometry: 2×1 matrices represent vectors and points, and 2×2 matrices represent transformations of those vectors and points.

2 Transforming Vectors

We now explain the way in which 2×2 matrices represent transformations of 2×1 matrices. The key concept here is matrix *multiplication*. The product of a 2×2 matrix M and a 2×1 matrix $\vec{v}$ is another 2×1 matrix $M\vec{v}$, defined as

$$M\vec{v} = \begin{bmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \end{bmatrix} = \begin{bmatrix} M_{00}v_0 + M_{01}v_1 \\ M_{10}v_0 + M_{11}v_1 \end{bmatrix}.$$

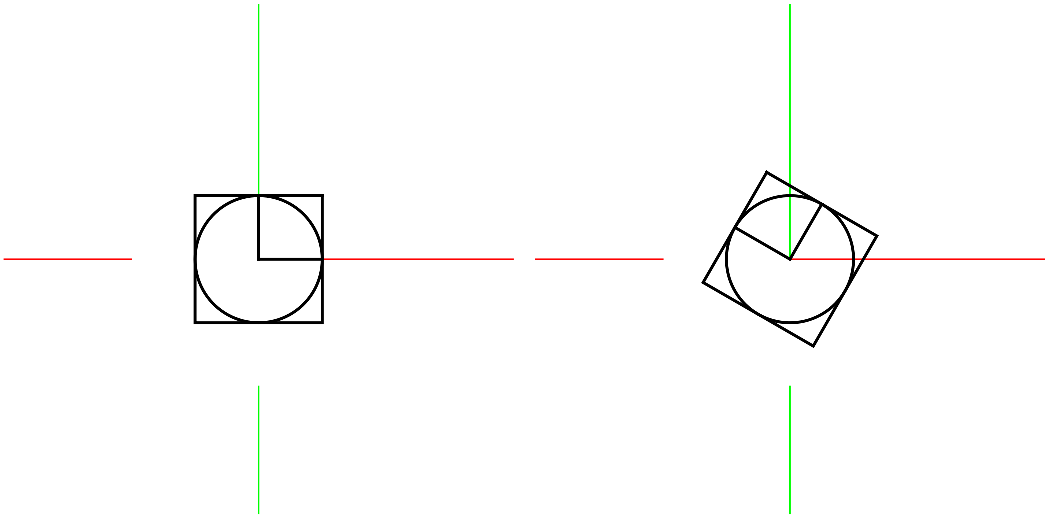

Figure 1: The left side is a shape in the plane. The right side is the same shape, after each of its points $\vec{v}$ has been rotated by the rotation matrix with $\theta = \pi/3 = 60°$.

To build your intuition, here are four important examples of what M could be.

First, for any angle θ, the matrix

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

represents counterclockwise rotation of the plane through θ. For example, if $\theta = \pi/3 = 60°$, then, for every vector $\vec{v}$,

$$M\vec{v} = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \end{bmatrix} = \begin{bmatrix} v_0/2 - v_1\sqrt{3}/2 \\ v_0\sqrt{3}/2 + v_1/2 \end{bmatrix}$$

is a vector of the same magnitude as $\vec{v}$, but directed $\pi/3$ radians or $60°$ counterclockwise from $\vec{v}$'s direction. See Figure 1.

Second, for any number $k > 0$,

$$M = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

represents a top-to-the-right shear. For example, if $k = 3$, then

$$M \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

See Figure 2. Similarly, if $k < 0$, then the shear is top-to-the-left.

Third, for any numbers k and ℓ,

$$M = \begin{bmatrix} k & 0 \\ 0 & \ell \end{bmatrix}$$

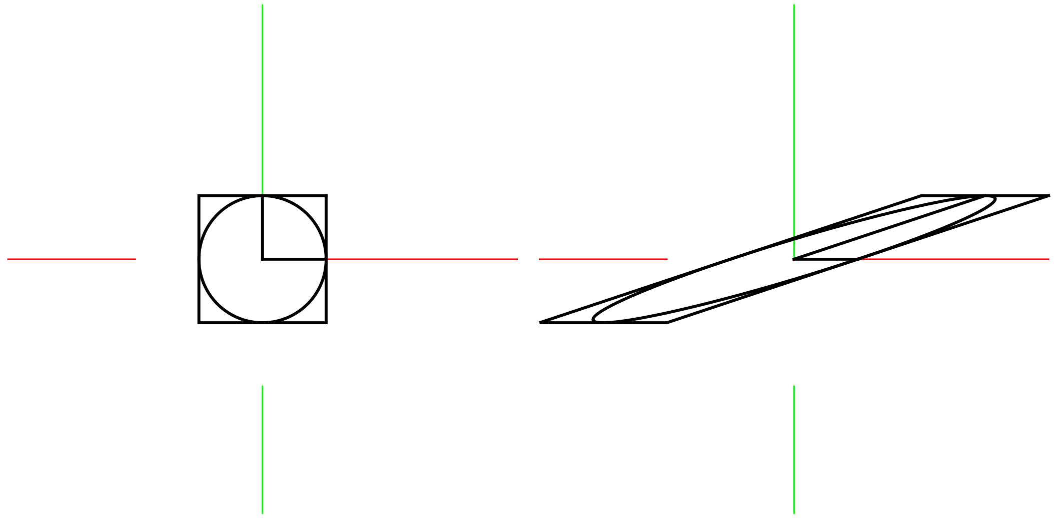

Figure 2: The left side is a shape in the plane. The right side is the same shape, after each of its points $\vec{v}$ has been sheared by the shear matrix with $k = 3$.

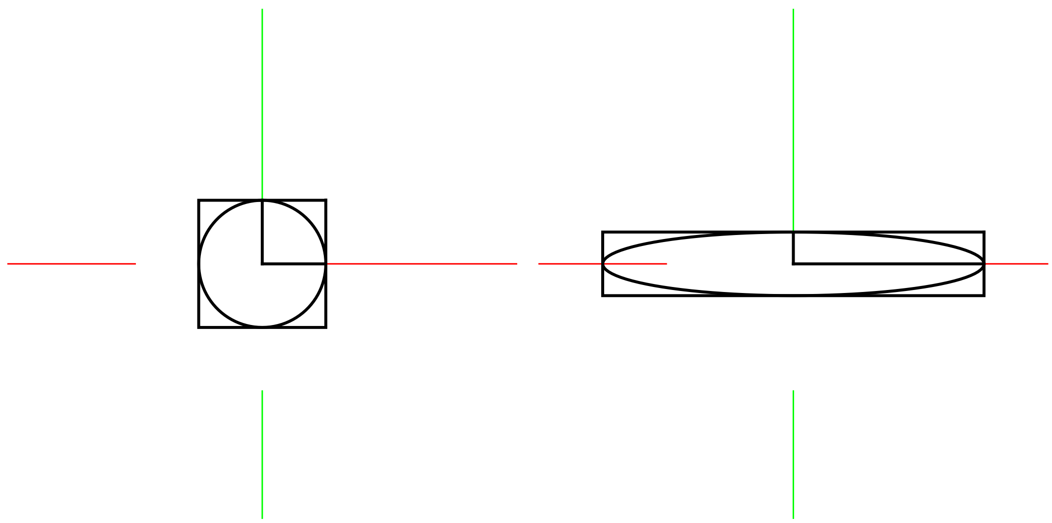

Figure 3: The left side is a shape in the plane. The right side is the same shape, after each of its points $\vec{v}$ has been distorted by the distortion matrix with $k = 3$ and $\ell = 1/2$.

represents a distortion that stretches the plane horizontally by a factor of k and vertically by a factor of ℓ. See Figure 3. Actually, "stretch" is a good descriptor only if $k, \ell > 1$. For example, if $\ell = 1/2$, then M compresses the plane in the vertical direction. If $k = -5$, then M's horizontal effect is to flip and stretch.

Fourth,

$$M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is the 2×2 *identity* matrix, usually denoted I. For any vector $\vec{v}$, $I\vec{v} = \vec{v}$. Geometrically, I represents the trivial transformation, that does nothing. Notice that I is a special case of all three preceding examples: It is rotation by $\theta = 0 = 0°$, shear by $k = 0$, and distortion by $k = \ell = 1$.

3 Composing Transformations

To *compose* transformations is to do one after another. For example, suppose that we want to shear the vector $\vec{v}$ and then rotate that sheared vector. Then we might compute

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \left(\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \vec{v} \right).$$

For each $\vec{v}$, we have to do two matrix multiplications: the inner one and then the outer one. When there are many vectors $\vec{v}$ to transform, we can approximately double the speed of the computation as follows.

Define the product of two 2×2 matrices like this:

$$MN = \begin{bmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{bmatrix} \begin{bmatrix} N_{00} & N_{01} \\ N_{10} & N_{11} \end{bmatrix} = \begin{bmatrix} M_{00}N_{00} + M_{01}N_{10} & M_{00}N_{01} + M_{01}N_{11} \\ M_{10}N_{00} + M_{11}N_{10} & M_{10}N_{01} + M_{11}N_{11} \end{bmatrix}.$$

It may help to notice that the left column of MN is M times the left column of N, and the right column of MN is M times the right column of N. Anyway, it turns out that matrix multiplication is associative, meaning that $(MN)P = M(NP)$ for any matrices. So we can rewrite the shear-then-rotate computation as

$$\left(\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \right) \vec{v} = \begin{bmatrix} \cos\theta & k\cos\theta - \sin\theta \\ \sin\theta & \cos\theta + k\sin\theta \end{bmatrix} \vec{v}.$$

Now transforming each $\vec{v}$ requires only one matrix multiplication.

You can string together three, four, or any number of transformations like this. The first transformation is on the right, and the last transformation is on the left. The order is important, because matrix multiplication is not commutative! For example, rotate-then-shear is

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta + k\sin\theta & k\cos\theta - \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

which, you can see, is different from shear-then-rotate.

The identity matrix I plays the role of 1 in the world of matrices, in that $IM = M = MI$ for all matrices M. This should make geometric sense: If we compose a transformation M with a transformation I that does nothing, then the overall effect is just M.

4 Determinant

The *determinant* of a 2×2 matrix M is a number defined as

$$\det M = \det \begin{bmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{bmatrix} = M_{00}M_{11} - M_{01}M_{10}.$$

The determinant captures a crucial geometric property of the transformation represented by M: its area distortion. For example, a rotation matrix has determinant

$$\det \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \cos^2\theta + \sin^2\theta = 1.$$

If you draw a square in the plane and then rotate it, the rotated square has the same area as the original square (Figure 1). Similarly, the shear matrix has determinant 1, indicating that it doesn't change the area of a square, even though it changes the shape (Figure 2).

The distortion matrix, on the other hand, has determinant $k\ell$ and changes the area of a square by that factor (Figure 3). If $\ell = 1/k$ then the determinant is 1 and there is no area change. If $k > 0$ and $\ell < 0$, then the determinant is negative, indicating that the plane has been flipped. The same is true if $k < 0$ and $\ell > 0$. If both k and ℓ are negative, then the determinant is positive. Intuitively, the transformation flips the plane twice, so it's really not flipped at all.

5 Inversion

In the real numbers, every non-zero number x has a multiplicative inverse $x^{-1} = 1/x$, meaning that $xx^{-1} = 1 = x^{-1}x$. Analogously, in the 2×2 matrices, every matrix M with non-zero determinant has a multiplicative *inverse* M^{-1} such that $MM^{-1} = I = M^{-1}M$. Intuitively, M^{-1} is the transformation that undoes M, because

$$M^{-1}M\vec{v} = I\vec{v} = \vec{v}$$

for all $\vec{v}$. Another way of looking at it is that M is the transformation that undoes M^{-1}:

$$MM^{-1}\vec{v} = I\vec{v} = \vec{v}.$$

A matrix M with determinant 0 does not have an inverse M^{-1}.

For 2×2 matrices, computing inverses is fairly easy, quick, and numerically stable. The inverse of an M with $\det M \neq 0$ is

$$M^{-1} = \begin{bmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{bmatrix}^{-1} = \frac{1}{\det M} \begin{bmatrix} M_{11} & -M_{01} \\ -M_{10} & M_{00} \end{bmatrix}.$$

On the right side of that equation, multiplying the matrix by the number $1/\det M$ means multiplying each entry of the matrix by that number — much like when you scale a vector by a number.

Here's a problem for you, to test your understanding. On paper, check that the expression for M^{-1} satisfies $MM^{-1} = I$. Also, compute the inverses of the four examples given above (rotation, shear, distortion, identity). Do the answers make geometric sense? Depending on your math background, that last question might not be easy. Talk to me about it.

6 Solving Linear Systems

Here's a common math problem: Given numbers a, b, c, d, g, h, find numbers x, y such that

$$\begin{aligned} ax + by &= g, \\ cx + dy &= h. \end{aligned}$$

Those two equations can be rewritten as one equation of 2×1 matrices,

$$\begin{bmatrix} ax + by \\ cx + dy \end{bmatrix} = \begin{bmatrix} g \\ h \end{bmatrix},$$

which can in turn be rewritten using matrix multiplication:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} g \\ h \end{bmatrix}.$$

Let M be that 2×2 matrix in there. So we're trying to solve

$$M \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} g \\ h \end{bmatrix}.$$

If M^{-1} exists, then we can multiply the equation, on the left side of both sides, by M^{-1}:

$$M^{-1} M \begin{bmatrix} x \\ y \end{bmatrix} = M^{-1} \begin{bmatrix} g \\ h \end{bmatrix}.$$

We do that because the left side of the equation simplifies down to the thing that we wanted:

$$\begin{bmatrix} x \\ y \end{bmatrix} = M^{-1} \begin{bmatrix} g \\ h \end{bmatrix}.$$

Now we know what x and y are. In this way, matrices and their inverses help us solve systems of linear equations.

(If M^{-1} does not exist, then there may be no solutions or, in rare cases, infinitely many. That issue won't arise much in this course.)