This tutorial teaches you about 2×2 matrices: multiplication, inversion, application to vectors, and their geometric meaning.

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 particular 2×3 matrix:

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

It is common to index the entries of an $n \times m$ matrix with subscripts $1 \leq i \leq n$ and $1 \leq j \leq m$. For example, the entry in row 5 and column 2 of a big matrix M is denoted M_{52}. Here is a general 2×3 matrix in that notation:

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

Here is a general 2×1 matrix. When writing $n \times 1$ and $1 \times m$ matrices, it is common to drop the index that never varies.

$$V = \begin{bmatrix} V_{11} \\ V_{21} \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}.$$

In this tutorial, we will 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 linear transformations of those vectors and points.

Frequently we will need to transcribe matrices into C, where rows and columns are indexed from 0 rather than 1. So let's switch to that indexing convention right now. Here are a general 2×2 matrix M and 2×1 matrix $\vec{v}$:

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

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

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

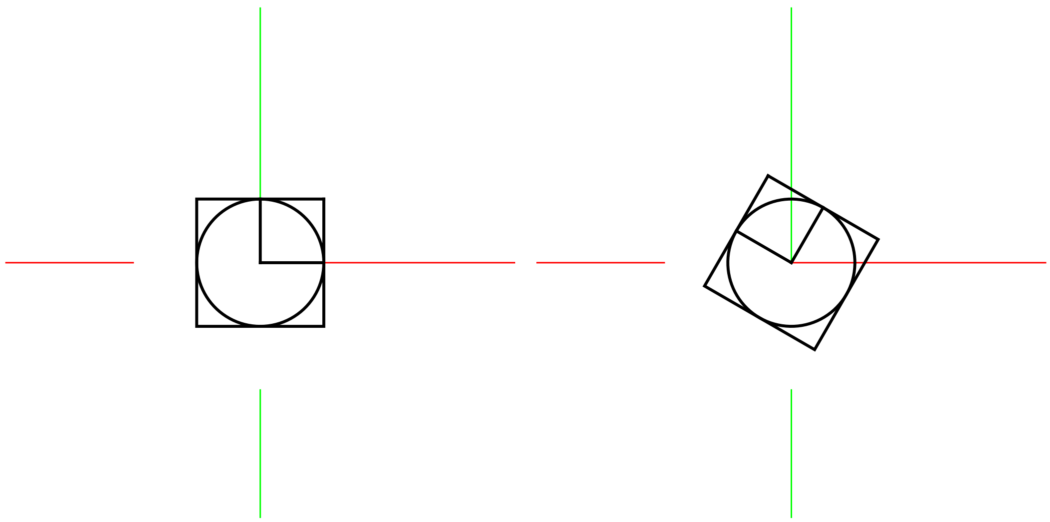

Figure 1: The left side is a shape in the plane. The right side is the same shape, after each of its points $\vec{v}$ has been rotated by the rotation matrix with $\theta = \pi/3 = 60°$.

Here are four important examples, to build your intuition.

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 $\vec{v}$,

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

is a vector of the same magnitude as $\vec{v}$, but directed $\pi/3$ radians or $60°$ counterclockwise from $\vec{v}$'s direction. See Fig. 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 Fig. 2. Similarly, if $k < 0$, then the shear is top-to-the-left.

Third, for any numbers k and ℓ,

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

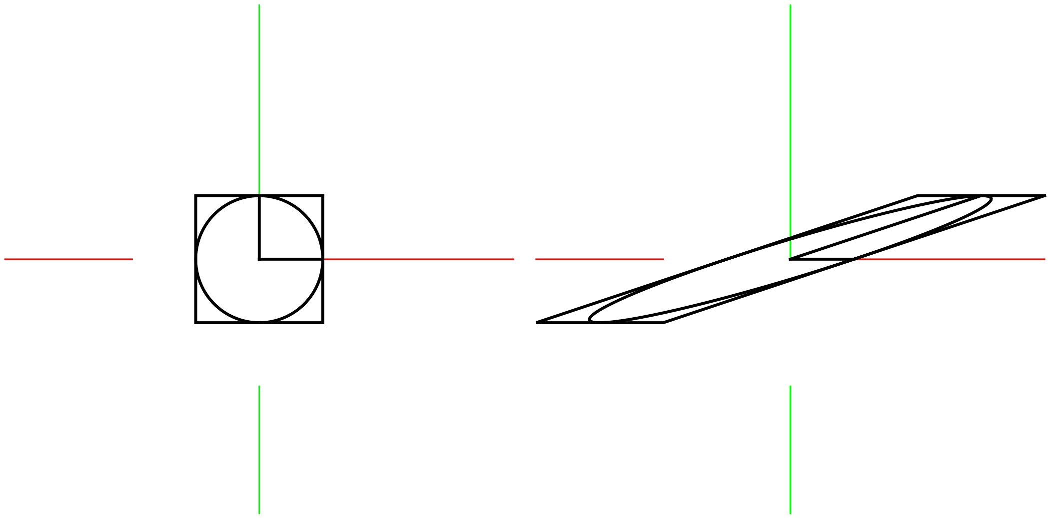

Figure 2: The left side is a shape in the plane. The right side is the same shape, after each of its points $\vec{v}$ has been sheared by the shear matrix with $k = 3$.

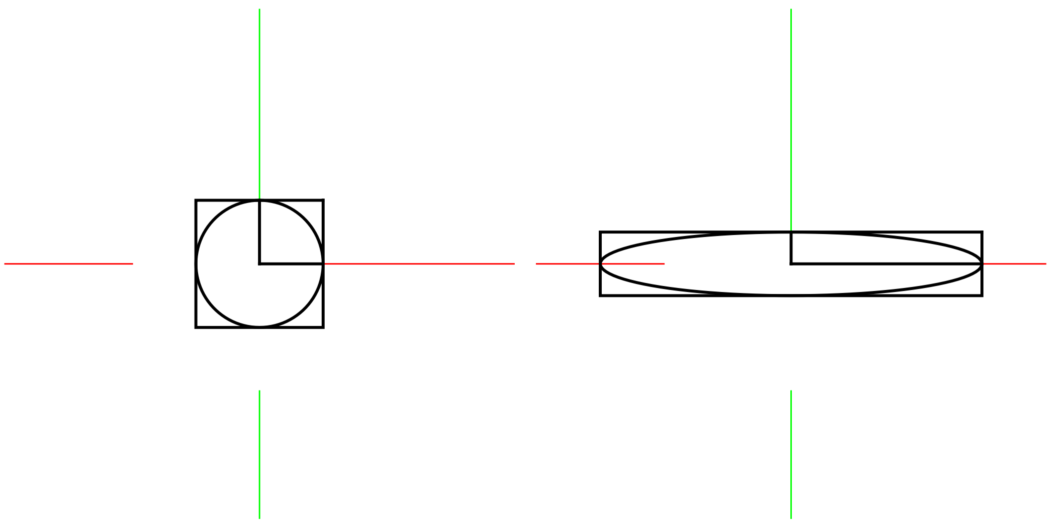

Figure 3: The left side is a shape in the plane. The right side is the same shape, after each of its points $\vec{v}$ has been distorted by the distortion matrix with $k = 3$ and $\ell = 1/2$.

represents a distortion that stretches the plane horizontally by a factor of k and vertically by a factor of ℓ. See Fig. 3. Actually, "stretch" is a good descriptor only if $k, \ell > 1$. For example, if $k = 1/2$, then M compresses the plane in the horizontal direction. If $\ell = -3$, then M's vertical 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 $\theta = 0 = 0^\circ$, shear by $k = 0$, and distortion by $k = \ell = 1$.

3 Coordinates

Let's take a moment to remember that vectors and points are not the same thing as 2×1 matrices. Similarly, linear transformations are not the same thing as 2×2 matrices. Vectors and linear transformations have a Platonic existence, and we can write them as matrices only after choosing a coordinate system.

If we stick to a single coordinate system, then we can largely ignore this distinction. However, in this course we will sometimes use many coordinate systems at once. For example, when we organize a scene using a scene graph, each node in the graph will have its own coordinate system. We'll deal with that when the time comes.

4 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 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. It turns out that matrix multiplication is associative. 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,

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

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

5 Determinant

The *determinant* of a 2×2 matrix M is

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

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 (Fig. 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 (Fig. 2).

The distortion matrix, on the other hand, has determinant $k\ell$ and changes the area of a square by that factor (Fig. 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.

6 Inversion

In the real numbers, every non-zero number x has a multiplicative inverse $x^{-1} = 1/x$, meaning that $xx^{-1} = 1$. 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 $\vec{v}$. (If $\det M = 0$, then M^{-1} does not exist.)

If you've studied linear algebra, then you've probably learned a complicated algorithm for inverting matrices. The standard algorithm is difficult to program and extremely difficult to make numerically stable. To my knowledge it is never used in actual computers. Fortunately, you don't need any such algorithm for 2×2 matrices, because here is an explicit expression for the inverse:

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

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.

On paper, you should check that the expression for M^{-1} satisfies $MM^{-1} = I$. You should also compute the inverses of the rotation, shear, and distortion matrices above. Do the answers make geometric sense? Depending on your math background, that last question might not be easy. Talk to me about it.

7 Solving linear systems

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

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

These two equations can be rewritten as one equation of 2×1 columns,

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

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. If M^{-1} exists, then the unique solution is

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

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 probably won't come up in this course.)