Matrici

Eugenio Animali

15 09 2022

1 Basis Vectors

The basis vectors of a coordinate system are a set of linearly independent vectors (vectors that cannot be described as a scaled sum of the other basis vectors), that when scaled and summed, can reach the full span of the whole coordinate system.

By default, in a 2d coordinate system, these vectors are: $\hat{i} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\hat{j} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

When one states a vector such as $\vec{V} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$, one is scaling and summing the basis vectors $\hat{i}$ and $\hat{j}$ to reach a position on the coordinate system:

$$(4 \times \hat{i}) + (2 \times \hat{j})$$

2 Matrices

Matrices describe a linear transformation that can be applied to any vector to move the vector to a new position. The beauty of linear transformations, is that they do not have to be applied directly to the vector to be transformed, which would be a long and dynamic thing to do; They can be applied directly to the basis vectors, with the same effect! In this way, to describe such a transformation, one only needs to describe the transformation of the basis vectors:

The matrix values are displayed in a grid format. Each column in the matrix describes transformation of one of the basis vectors of the old coordinate system, which creates the new basis vector. In each column, the different values are the instructions for how the basis vector must move in each dimension.

Here is an example of a simple 2 by 2 matrix, transforming the simple vector $\vec{W} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$:

$$\begin{bmatrix} 2 & 3 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Here, vector $\vec{W}$ is the sum of the scaled vectors $\vec{i}$ and $\vec{j}$. To understand what this transformation has done, one can consider that $\vec{W}$ is still equal to $3\vec{i} + 2\vec{j}$; It is the basis vectors

themselves that have changed in the way described by the matrix: now, $\vec{i} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ and $\vec{j} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

so the new coordinates of $\vec{W}$ are

$$3 \begin{bmatrix} 2 \\ -2 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 6 \\ -6 \end{bmatrix} + \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 12 \\ -4 \end{bmatrix}$$

This transformation can be done to as many vectors as you want, meaning that it can be performed on another matrix (which might be the description of the new basis vectors for your coordinate system). The same process is applied to each vector (column) of the matrix.

Note: matrix transformations are associative
$A(BC) = (AB)C$

but not communicative
$AB =! BA$

3 Determinant

When a matrix is used on a coordinate system, it will squash, rotate, and stretch the coordinate space in different ways. When working with areas, you will therefore want to know how areas, volumes, and alike are affected. Luckily, when working with strictly linear transformations, for an established matrix, There will be a single coefficient which will describe the change of area of any shape. This is what the value of the determinant provides.

The determinant of a vector space is the area enclosed by a parallelogram, parallelopiped, or higher dimension equivalent, drawn by the basis vectors. In a simple 2d vector space with the normal $\hat{i}$ and $\hat{j}$ basis vectors, this will be the 1 by 1 square between the origin and (1,1). As the determinant increases in a stretched vector space, so will all area, volume and alike by the same factor.

How to find the determinant for 2d matrices:

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

$$Det(M) = ad - bc$$

For matrices of higher dimensions:

$$N = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

$$Det(N) = a \times Det(\begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \times Det(\begin{bmatrix} d & f \\ g & i \end{bmatrix} + c \times Det(\begin{bmatrix} d & e \\ g & h \end{bmatrix}$$

4 Inverse Matrix

Simply put, the Inverse of a matrix is that matrix which undoes the transformation done by the matrix, restoring the basis vectors to $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

This is useful when used to solve a linear system of equations:

$$\begin{cases} 3x + 4y + 2z = 5 \\ 2x + 3y + 8z = 8 \\ 1x + 0y + 2z = 4 \end{cases}$$

One might notice that this looks like the last steps of a matrix multiplication process, so rather than solving like one used to do, one might use this newfound consideration to their advantage. First lets format it as a matrix:

$$\begin{bmatrix} 3 & 4 & 2 \\ 2 & 3 & 8 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \\ 4 \end{bmatrix}$$

Now it is evident that the way to proceed is to multiply both sides by the inverse matrix to get unknowns on one side and numbers on the other.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1} \begin{bmatrix} 5 \\ 8 \\ 4 \end{bmatrix}$$

How do you find the inverse? Solution for 2x2:

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

$$P^{-1} = \frac{1}{Det(P)} \begin{bmatrix} d & -b \\ -d & a \end{bmatrix}$$

5 Dot Product $\vec{v} \cdot \vec{w}$

The first way to multiply two vectors is the dot product, which returns a scalar value. This value can be found in two ways:
Lets use two vectors $\vec{v}$ and $\vec{w}$.

1. This can be found by multiplying the projection of $\vec{v}$ on $\vec{w}$ by $\vec{w}$.

2. Another way to find the dot product is to sum the multiplication of the rows:

$$\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} \cdot \begin{bmatrix} x_v \\ y_v \\ z_v \end{bmatrix} = x_1 x_2 + y_1 y_2 + z_1 z_2$$

This duality seems pretty unexpected. To explain it, method 1 can be considered as a transformation of $\vec{v}$ onto the plane where lies $\vec{w}$. The matrix for this transformation can be found like so:

Consider that the projections of $\hat{i}$, $\hat{j}$ and $\hat{k}$ on $\vec{w}$, when multiplied by $|\vec{w}|$ will be the same as the projection of $\vec{w}$ on the respective axes (ie x_w, y_w, z_w). These projections of the basis vectors are all that is needed to transform $\vec{v}$ onto the plane where lies $\vec{w}$; Therefore the resulting matrix is:

$$\begin{bmatrix} x_w & y_w & z_w \end{bmatrix}$$

resulting in the transformation

$$\begin{bmatrix} x_w & y_w & z_w \end{bmatrix} \cdot \begin{bmatrix} x_v \\ y_v \\ z_v \end{bmatrix} = x_1 x_2 + y_1 y_2 + z_1 z_2$$

identical to what is drawn in method 2.

Notice that $\vec{w}$ acts as a matrix, transforming the other vector onto its line of existence.

Note: commutative

$$\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$$

but not associative as it returns a scalar, which cannot be dot multiplied with a vector... it makes no sense

6   Cross Product $\vec{v} \times \vec{w}$

When crossing vectors in two dimensions, the computation returns the area enclosed in the parallelogram between the vectors. This can be done by finding the determinant of the matrix that contains those two vectors.

$$Det \left(\begin{bmatrix} x_v & x_w \\ y_v & y_w \end{bmatrix} \right)$$

When crossing 3 vectors in a three dimensional vector space, it returns another vector $\vec{P}$ with magnitude equal to the area enclosed in the parallelogram between the vectors, in the direction perpendicular to the plane on which the vectors lie. This describes a vector for which

$$\vec{P} \cdot \vec{F} = Det \left(\begin{bmatrix} x_F & x_v & x_w \\ y_F & y_v & y_w \\ z_F & z_v & z_w \end{bmatrix} \right)$$

for any vector F.

ie. The dot product with P should give the same answer as creating a parallelopiped with $\vec{v}$ and $\vec{w}$ for any other vector you throw at it.

If you consider the parallelopiped, only the part of F perpendicular to $\vec{v}$ and $\vec{w}$ has any impact at all, so it makes sense that P, being perpendicular, only utilises that aspect of F; and P is proportional to the base area created by $\vec{v}$ and $\vec{w}$ to scale F in the same way as with $\vec{v}$ and $\vec{w}$.

This way we find a dual vector, which can be used as a matrix to find the area value for the parallelopiped. Notice the similarity with the Dot product.

This can be done computationally by finding the determinant of a matrix with $\hat{i}, \hat{j}$ and $\hat{k}$ on

the first row, and the two vectors on the second and third rows, due to a notational trick, which doesn't make sense (the values in a matrix should be scalars for scaling vectors, and having a vector as a value does not make sense).

$$
\begin{bmatrix} x_P \\ y_P \\ z_P \end{bmatrix} \cdot \begin{bmatrix} x_F \\ y_F \\ z_F \end{bmatrix} = Det\left(\begin{bmatrix} x_F & x_v & x_w \\ y_F & y_v & y_w \\ z_F & z_v & z_w \end{bmatrix}\right)
$$

expanded...

$$
x_F(x_P) + y_F(y_P) + z_F(z_P) = x_F(y_v z_w - y_w z_v) + y_F(z_v x_w - x_v z_w) + z_F(x_v y_w - x_w y_v)
$$

One can see that the coordinates of F do not interfere with $P, \vec{v}$ and $\vec{w}$, rather usefully separating the parts of the determinant to become the coordinates of P

$$
x_F(x_P) = x_F(y_v z_w - y_w z_v)
$$
$$
y_F(y_P) = y_F(z_v x_w - x_v z_w)
$$
$$
z_F(z_P) = z_F(x_v y_w - x_w y_v)
$$

$$
x_P = y_v z_w - y_w z_v
$$
$$
y_P = z_v x_w - x_v z_w
$$
$$
z_P = x_v y_w - x_w y_v
$$

And each of these coordinates ends up with a basis vector assigned, as normal, which handily keeps the components separate.

$$
\vec{P} = \hat{i}(y_v z_w - y_w z_v) + \hat{j}(z_v x_w - x_v z_w) + \hat{k}(x_v y_w - x_w y_v)
$$

$$
\vec{P} = \begin{bmatrix} y_v z_w - y_w z_v \\ z_v x_w - x_v z_w \\ x_v y_w - x_w y_v \end{bmatrix}
$$

7   Change of Basis

Sometimes one might need to change their basis vectors, to see the world in a new way. Let's say we have two perspectives A and B of a vector space. To translate between them, we need to be able to describe the basis vectors of one system using the basis vectors of the other system. Matrix C is how A would describe B's basis vectors. If A wants to understand vector $\vec{v}$ in B's system, A must first apply the matrix to make explicit in A's system the vectors that must be scaled and summed.

$\vec{v}$ is a vector described by B

$C\vec{v}$ is the same vector as described by A

To return the information, A must reconvert the vector using the inverse matrix.

$C^{-1}C\vec{v}$ is returned to B

To transform $\vec{v}$ with a matrix in A's system, A must first convert the matrix, transform, and then return the matrix to B

$$C^{-1}MC\vec{v}$$

8 Eigenvectors and Eigenvalues

When a matrix is applied to a coordinate system, it rotates, squishes and stretches space in different directions. Most of the time, there are some lines which remain stationary, meaning that any vector on that line remains on that line after the transformation. All vectors on that line are called eigenvectors. Sometimes, like in a 90 deg rotation, there are no eigenvectors; in other transformations, such as the sheer, there is only one set of eigenvectors. The eigenvalue of a line of eigenvectors is the factor by which the vector is scaled in the transformation. Eigenvectors and eigenvalues are described by this equation:

$$A\vec{v} = \lambda\vec{v}$$

where $\vec{v}$ is an eigenvector and λ is the associated eigenvalue
Meaning that the transformation caused by A is a scalar. To solve this equation, λ must be considered as a matrix transformation. This is done by multiplying it by the identity matrix $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ to get:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\vec{v} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\vec{v}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\vec{v} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\vec{v} = 0$$

$$\begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix}\vec{v} = 0$$

The only way for a non-zero vector to become null, is if the matrix squashes space onto fewer dimensions. In this case, the determinant will be zero. So let's set the determinant to equal zero:

$$(a - \lambda)(d - \lambda) - bc = 0$$

once the possible eigenvalues have been found, one can now find the eigenvectors which satisfy the equation.

$$\begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix}\vec{v} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

ie the vectors perpendicular to any of the vectors in the matrix.

9 A quick way to compute Eigenvalues

First we need to set some notions:

1. The mean value of the diagonals of a matrix is equal to the mean of the eigenvalues.

$$\text{for a matrix } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

$$\frac{a+b}{2} = \frac{\lambda_1 + \lambda_2}{2}$$

2. The determinant of the matrix is equal to the product of the eigenvalues.

$$Det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \lambda_1 \lambda_2$$

as these could be considered the sizes of two valid basis vectors for the transformed coordinate space.

3. From these two, we can reach the last notion:
we know that the two eigenvalues are equally distant from the mean (lets call the mean m, and the distance d), and that the product $p = (m+d)(m-d) = m^2 - d^2$.

$$d^2 = m^2 - p$$
$$d = \pm\sqrt{m^2 - p}$$

$$\lambda_1, \lambda_2 = m \pm \sqrt{m^2 - p}$$

10   A final thought: Abstract Vector Spaces

what is a vector? In reality any set of numbers is a vector, like a function. A function seems to have vector properties (with an infinitely large set of numbers)...
take a polynomial:
$$ax^3 + bx^2 + cx + d$$

These seem to be scaled and summed basis functions, where these basis functions are $b_0 = x^0, b_1 = x^1, b_2 = x^2, \ldots$
So we can find the vector of polynomials like this:

$$\begin{bmatrix} d \\ c \\ b \\ a \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

Then what does it mean to apply a function such as the derivative to this vector? Let's start by acknowledging that the derivative is a linear transformation, because it has the properties of additivity and scaling:

1. $\frac{d}{dx}(\vec{v} + \vec{w}) = \frac{d}{dx}\vec{v} + \frac{d}{dx}\vec{w}$

2. $\frac{d}{dx}(4\vec{v}) = 4\frac{d}{dx}\vec{v}$

This means it can be described as a matrix! The matrix would look like this:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ 0 & 0 & 0 & 3 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ 0 & 0 & 0 & 3 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} d \\ c \\ b \\ a \\ \vdots \end{bmatrix} = \begin{bmatrix} c \\ 2b \\ 3a \\ 0 \\ \vdots \end{bmatrix} = 3ax^2 + 2bx + c$$

To find this matrix, we have to use the derivative of each basis function as a row of the matrix, as that will determine where the basis functions land after the transformation.

Therefore to find the matrix for any transformation of any space of functions, we have to determine the basis functions of the space, which span the full vector space, and find the matrix transformation by applying the transformation to each basis function and placing it in a row of the matrix.