Linear Algebra

Logit

2016

Table of Contents

Intro to Linear Algebra

Linear algebra is a branch of mathematics that deals with with vector spaces and systems of linear equations. It represents problems in terms of vectors and matrices, and is incredibly useful, with applications showing up in:

- mechanical physics
- quantum mechanics
- control theory
- computer vision
- fluid flow in pipe networks
- circuitry
- optimization

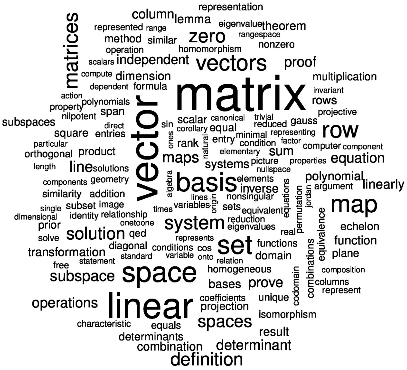


Image Source

Vectors

Scalars

A scalar is a quantity possessing magnitude but no direction. Real-life examples include:

- $25\,^{\circ}\mathrm{C}$ outside temperature
- $40 US dollars
- 60 kg

In other words, ignoring the units used here, a scalar is just a number by itself.

Shown here: a scaler, not a scalar.

Vectors

A vector represents a quantity possessing both magnitude *and* direction. Real-life examples include:

- Driving 45 miles per hour going east.
- Flying upwards on a rocket at 700 meters per second.
- Rotating a vinyl record at 45 revolutions per minute clockwise.

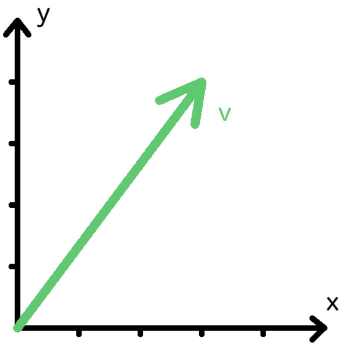
Vectors

A vector is usually represented as an ordered *n*-tuple, and can be visualized as an arrow in *n*-dimensional space with its tail at the origin.

For example, the vector

$$\boldsymbol{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

is a column vector which can be represented as an arrow in 2-dimensional space with its base at the origin (0,0) pointing towards (3,4), as depicted by the green arrow on the right.



Vector Addition and Subtraction

Vector addition and subtraction is done elementwise:

$$\begin{bmatrix} 9 \\ 2 \\ 6 \end{bmatrix} + \begin{bmatrix} 8 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 9+8 \\ 2+5 \\ 6+7 \end{bmatrix} = \begin{bmatrix} 17 \\ 7 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 4 \\ 17 \\ 6 \\ 5 \end{bmatrix} - \begin{bmatrix} 20 \\ 19 \\ 19 \\ 11 \end{bmatrix} = \begin{bmatrix} 4-20 \\ 17-19 \\ 6-19 \\ 15-11 \end{bmatrix} = \begin{bmatrix} -16 \\ -2 \\ -13 \\ 4 \end{bmatrix}$$

Note that we can only add or subtract vectors with the same number of elements.

Vector Addition and Subtraction

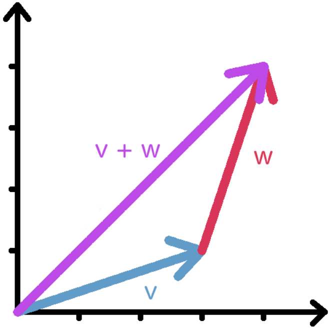
Vector addition can be visualized as putting vectors end-to-end. For instance, if

$$\boldsymbol{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \qquad \boldsymbol{w} = \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

then the sum

$$\boldsymbol{v} + \boldsymbol{w} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

is the same vector that results from putting the tail of w at the head of v and drawing a vector from the tail of v to the head of w.



Scalar Multiplication With a Vector

When we multiply a scalar with a vector, we apply the multiplication to each element of the vector:

$$2 \cdot \begin{bmatrix} 13 \\ 7 \\ 2 \\ 16 \end{bmatrix} = \begin{bmatrix} 2 \cdot 13 \\ 2 \cdot 7 \\ 2 \cdot 2 \\ 2 \cdot 16 \end{bmatrix} = \begin{bmatrix} 26 \\ 14 \\ 4 \\ 32 \end{bmatrix}$$

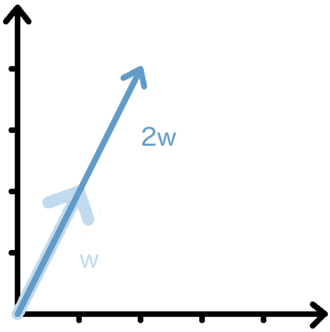
$$10 \cdot \begin{bmatrix} 10 \\ 9 \\ 7 \\ 4 \\ 8 \end{bmatrix} = \begin{bmatrix} 10 \cdot 10 \\ 10 \cdot 9 \\ 10 \cdot 7 \\ 10 \cdot 4 \\ 10 \cdot 8 \end{bmatrix} = \begin{bmatrix} 100 \\ 90 \\ 70 \\ 40 \\ 80 \end{bmatrix}$$

Scalar Multiplication With a Vector

Vector multiplication by a scalar can be visualized as "stretching" the vector. Scalar multiples of a vector keep the same direction as the original vector, but have a different length or magnitude.

$$\boldsymbol{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$2\boldsymbol{w} = \begin{bmatrix} 2 \cdot 1 \\ 2 \cdot 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$



Row and Column Vectors

A vector can be expressed as either a row vector or a column vector. A row vector has the numbers listed horizontally:

$$\boldsymbol{b} = \begin{bmatrix} 16 & 11 & 2 \end{bmatrix}$$

A column vector has the numbers listed vertically:

$$\boldsymbol{c} = \begin{bmatrix} 17 \\ 3 \\ 9 \end{bmatrix}$$

The distinction will become important when we discuss matrices.

Vector Transpose

A row vector can be transformed into a column vector (or vice-versa) by taking the transpose of that vector, usually denoted with a superscript T:

$$\boldsymbol{b} = \begin{bmatrix} 16 & 11 & 2 \end{bmatrix} \qquad \boldsymbol{b}^T = \begin{bmatrix} 16 \\ 11 \\ 2 \end{bmatrix}$$

$$\boldsymbol{c} = \begin{bmatrix} 17 \\ 3 \\ 9 \end{bmatrix} \qquad \boldsymbol{c}^T = \begin{bmatrix} 17 & 3 & 9 \end{bmatrix}$$

Dot Product

If we have two n-dimensional vectors $\boldsymbol{v}$ and $\boldsymbol{w}$:

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \qquad \boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

then the dot product of $\boldsymbol{v}$ and $\boldsymbol{w}$ is defined as

$$\boldsymbol{v} \cdot \boldsymbol{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n$$

For example, if

$$\boldsymbol{v} = \begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \qquad \boldsymbol{w} = \begin{bmatrix} 8 \\ 2 \\ 6 \end{bmatrix}$$

then

$$\boldsymbol{v} \cdot \boldsymbol{w} = 3 \cdot 8 + 7 \cdot 2 + 4 \cdot 6 = 62$$

Vector Norm

The Euclidean vector norm gives us the length or magnitude of a vector. If we have a vector $\boldsymbol{v}$:

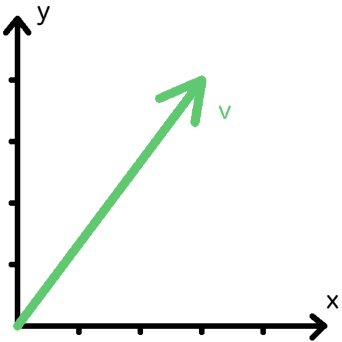
$$\boldsymbol{v} = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}^T$$

then the vector norm $||\boldsymbol{v}||$ is defined as

$$\begin{aligned} ||\boldsymbol{v}|| &= \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \\ &= \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}} \end{aligned}$$

In other words, the dot product of $\boldsymbol{v}$ with itself is the square of the vector norm:

$$\boldsymbol{v} \cdot \boldsymbol{v} = ||\boldsymbol{v}||^2$$



Vector Norm

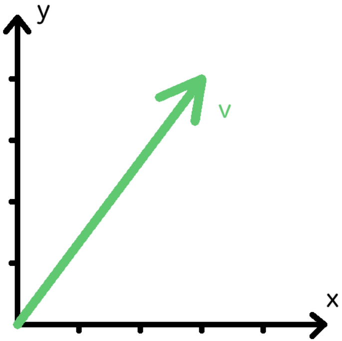
For instance, if we define a vector $\boldsymbol{v}$ like the green vector shown on the right:

$$\boldsymbol{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

the vector norm of $\boldsymbol{v}$ is

$$\begin{aligned} ||\boldsymbol{v}|| &= \sqrt{3^2 + 4^2} \\ &= \sqrt{9 + 16} = 5 \end{aligned}$$

So, the green vector has length 5.



Orthogonal Vectors

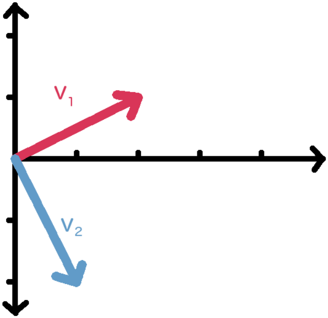
Two vectors are orthogonal to each other if the angle between them is a right angle (90°). For instance, the vectors

$$\mathbf{v_1} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad \mathbf{v_2} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

are orthogonal to each other.

The dot product of two vectors orthogonal to each other is *always zero*:

$$\mathbf{v_1} \cdot \mathbf{v_2} = (2 \cdot 1) + (1 \cdot -2) = 0$$



Normalized Vectors

A unit vector is a vector with norm (length) 1. A vector is normalized when it is scaled to a unit vector. This can be done by dividing the vector by the length of the vector. For instance, a vector $\boldsymbol{w}$ can be normalized to a unit vector $\boldsymbol{u}$ like so:

$$\boldsymbol{u} = \frac{\boldsymbol{w}}{||\boldsymbol{w}||}$$

For instance, if the vector $\boldsymbol{w}$ is defined as

$$\boldsymbol{w} = \begin{bmatrix} 7 & 5 & 5 & 1 \end{bmatrix}^T$$

the norm of $\boldsymbol{w}$ is

$$||\boldsymbol{w}|| = \sqrt{7^2 + 5^2 + 5^2 + 1^2} = \sqrt{100} = 10$$

The normalized vector $\boldsymbol{u}$ is

$$\boldsymbol{u} = \frac{\boldsymbol{w}}{||\boldsymbol{w}||} = \frac{1}{10} \begin{bmatrix} 7 & 5 & 5 & 1 \end{bmatrix}^T = \begin{bmatrix} 0.7 & 0.5 & 0.5 & 0.1 \end{bmatrix}^T$$

Orthonormal Vectors

A set of orthonormal vectors is a set of orthogonal vectors that have been normalized so that all vectors have unit length.

For instance, if we have orthogonal vectors $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$:

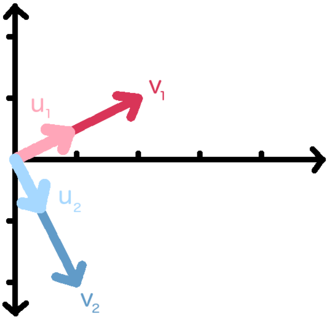
$$\boldsymbol{v_1} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad \boldsymbol{v_1} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

the norms of each of these vectors are

$$||\boldsymbol{v_1}|| = ||\boldsymbol{v_2}|| = \sqrt{5}$$

so the corresponding orthonormal vectors $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$ are

$$\boldsymbol{u_1} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad \boldsymbol{u_2} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$



Linear Combination of Vectors

If we have a set of m vectors:

$$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$$

then a linear combination of those vectors is any sum of the form

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_m\boldsymbol{v}_m$$

where $\{c_1, c_2, \ldots, c_m\}$ are all scalars.

Linear Combination of Vectors

For example, if we have the following row vectors:

$$\begin{aligned} \boldsymbol{v}_1 &= \begin{bmatrix} 7 & 6 & 3 \end{bmatrix} \\ \boldsymbol{v}_2 &= \begin{bmatrix} 8 & 1 & 4 \end{bmatrix} \\ \boldsymbol{v}_3 &= \begin{bmatrix} 2 & 4 & 9 \end{bmatrix} \end{aligned}$$

then

$$3\boldsymbol{v}_1 + \boldsymbol{v}_2 + 2\boldsymbol{v}_3 = \begin{bmatrix} 33 & 27 & 31 \end{bmatrix}$$

is a linear combination of $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$.

Linear Dependence / Independence

A set of vectors

$$\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$$

is said to be <u>linearly dependent</u> if there exist coefficients

$$\{c_1, c_2, \ldots, c_m\} \text{ not all zero}$$

such that

$$c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2 + \cdots + c_m\boldsymbol{w}_m = \mathbf{0} \leftarrow \text{the zero vector}$$

Otherwise, we say that those vectors are <u>linearly independent</u>.

Linear Dependence / Independence

For instance, if we have the following row vectors:

$$\boldsymbol{w}_1 = \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}$$
$$\boldsymbol{w}_2 = \begin{bmatrix} -1 & 0 & 3 \end{bmatrix}$$
$$\boldsymbol{w}_3 = \begin{bmatrix} 4 & 3 & 6 \end{bmatrix}$$

then since

$$3\boldsymbol{w}_1 + 2\boldsymbol{w}_2 - \boldsymbol{w}_3 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

we can say the set $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3\}$ is linearly *dependent*.

Span and Basis

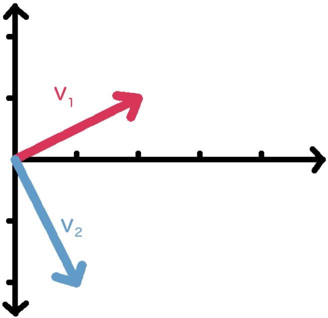
The span of a set of vectors is the space of vectors that can be represented as a linear combination of vectors from that set.

For instance, if we have

$$\mathbf{v}_1 = \begin{bmatrix} 2 & 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 1 & -2 \end{bmatrix}$$

then any vector $\mathbf{w}$ in 2-D space can be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$:

$$\mathbf{w} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2$$



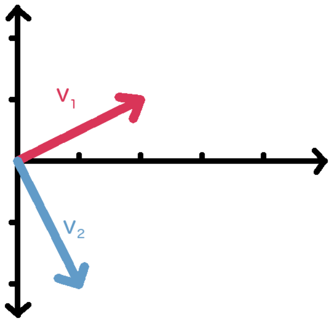
Span and Basis

A linearly independent set of vectors that spans a space is called a basis.

For instance,

$$\boldsymbol{v}_1 = \begin{bmatrix} 2 & 1 \end{bmatrix} \qquad \boldsymbol{v}_2 = \begin{bmatrix} 1 & -2 \end{bmatrix}$$

is a linearly independent set that spans the plane, so it is a basis for the plane.



Span and Basis

The most commonly used basis is known as the canonical basis. For instance, in 3-D space, the vectors

$$\begin{aligned} \boldsymbol{e}_1 &= \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \\ \boldsymbol{e}_2 &= \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \\ \boldsymbol{e}_3 &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

are the most obvious choice for referencing all other vectors in the space. The canonical basis is an example of an orthonormal basis, as the set contains unit length vectors perpendicular to each other.

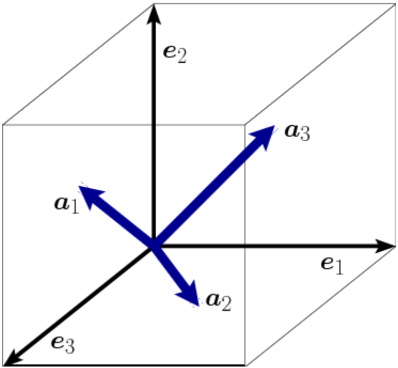


Image Source

Vector Space and Dimension

The region that these vectors inhabit is called a vector space. The dimension of said vector space is defined by the number of basis vectors necessary to represent any other vector in that space as a linear combination of basis vectors.

For instance, a 3 dimensional space requires 3 basis vectors to define any other vector in that space.

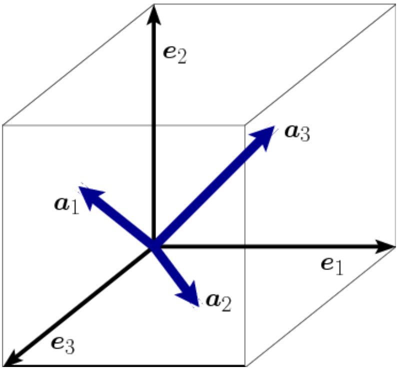


Image Source

Matrices

Matrices

What is a matrix?

A matrix is a rectangular array of numbers, symbols, or expressions arranged in rows and columns. For instance, consider the matrix M:

$$M = \begin{bmatrix} 12 & 7 & 18 \\ 2 & 13 & 11 \\ 19 & 10 & 10 \\ 4 & 1 & 17 \end{bmatrix}$$

M is a 4×3 matrix because it has 4 rows and 3 columns. We can view each matrix as a collection of row vectors or as a collection of column vectors.

Matrix Addition

Matrix addition is done elementwise, similar to vector addition:

$$\begin{bmatrix} 6 & 7 & 11 & 1 \\ 18 & 12 & 7 & 11 \\ 14 & 5 & 6 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 5 & 15 \\ 19 & 7 & 1 & 14 \\ 11 & 3 & 17 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 6+1 & 7+2 & 11+5 & 1+15 \\ 18+19 & 12+7 & 7+1 & 11+14 \\ 14+11 & 5+3 & 6+17 & 4+2 \end{bmatrix}$$

$$= \begin{bmatrix} 7 & 9 & 16 & 16 \\ 37 & 19 & 8 & 25 \\ 25 & 8 & 23 & 6 \end{bmatrix}$$

Note that we can only add (or subtract) matrices with the same dimensions.

Scalar Multiplication With a Matrix

Scalar multiplication with a matrix is also done elementwise:

$$A = \begin{bmatrix} 9 & 6 \\ 1 & 4 \end{bmatrix}$$

$$\begin{aligned} 3A &= 3 \cdot \begin{bmatrix} 9 & 6 \\ 1 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 3 \cdot 9 & 3 \cdot 6 \\ 3 \cdot 1 & 3 \cdot 4 \end{bmatrix} \\ &= \begin{bmatrix} 27 & 18 \\ 3 & 12 \end{bmatrix} \end{aligned}$$

Matrix Multiplication

Suppose we have matrices A and B:

$$A = \begin{bmatrix} 6 & 7 & 9 \\ 1 & 5 & 10 \end{bmatrix} \qquad B = \begin{bmatrix} 10 & 9 & 8 & 4 \\ 5 & 2 & 7 & 1 \\ 7 & 9 & 8 & 6 \end{bmatrix}$$

What is the matrix product $A \cdot B$?

If we set $A \cdot B = C$, we can find the elements of C by considering dot products of rows of A with columns of B.

Matrix Multiplication

$$A \cdot B = C$$

$$\begin{bmatrix} 6 & 7 & 9 \\ 1 & 5 & 10 \end{bmatrix} \cdot \begin{bmatrix} 10 & 9 & 8 & 4 \\ 5 & 2 & 7 & 1 \\ 7 & 9 & 8 & 6 \end{bmatrix} = \begin{bmatrix} 158 & & & \\ & & & \end{bmatrix}$$

For instance, c_{11} is the dot product of the *first* row of A with the *first* column of B.

$$c_{11} = 6 \cdot 10 + 7 \cdot 5 + 9 \cdot 7 = 158$$

Matrix Multiplication

$$A \cdot B = C$$

$$\begin{bmatrix} 6 & 7 & 9 \\ 1 & 5 & 10 \end{bmatrix} \cdot \begin{bmatrix} 10 & 9 & 8 & 4 \\ 5 & 2 & 7 & 1 \\ 7 & 9 & 8 & 6 \end{bmatrix} = \begin{bmatrix} 158 & 149 & & \\ & & & \end{bmatrix}$$

Next, c_{12} is the dot product of the *first* row of A with the *second* column of B.

$$c_{12} = 6 \cdot 9 + 7 \cdot 2 + 9 \cdot 9 = 149$$

Matrix Multiplication

$$A \cdot B = C$$

$$\begin{bmatrix} 6 & 7 & 9 \\ 1 & 5 & 10 \end{bmatrix} \cdot \begin{bmatrix} 10 & 9 & 8 & 4 \\ 5 & 2 & 7 & 1 \\ 7 & 9 & 8 & 6 \end{bmatrix} = \begin{bmatrix} 158 & 149 & 169 & 85 \\ 105 & 109 & 123 & 69 \end{bmatrix}$$

We can continue like this until we have used up every combination of rows in A and columns in B to give us all elements in C.

Matrix Multiplication

We can also view matrix multiplication as a collection of inner products in the following way:

$$A \cdot B = \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix} \begin{bmatrix} | & | & & | \\ b_1 & b_2 & \cdots & b_p \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} a_1^T b_1 & a_1^T b_2 & \cdots & a_1^T b_p \\ a_2^T b_1 & a_2^T b_2 & \cdots & a_2^T b_p \\ \vdots & \vdots & \ddots & \vdots \\ a_m^T b_1 & a_m^T b_2 & \cdots & a_m^T b_p \end{bmatrix}$$

Matrix Multiplication

Notes on matrix multiplication:

- The matrix product $A \cdot B$ is only allowed if the number of columns of A equals the number of rows of B.
- If A is $n \times m$ (n rows, m columns) and B is $m \times p$, then $A \cdot B$ is $n \times p$.
- Matrix multiplication is <u>not</u> commutative. This means $A \cdot B$ is usually not equal to $B \cdot A$ (assuming $B \cdot A$ even exists).
- Since vectors are just special cases of matrices, the dot product between two column vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ can be interpreted as the matrix product $\boldsymbol{v}^T \boldsymbol{w}$.

Matrix Transpose

$$M = \begin{bmatrix} 12 & 7 & 18 \\ 2 & 13 & 11 \\ 19 & 10 & 10 \\ 4 & 1 & 17 \end{bmatrix} \qquad M^T = \begin{bmatrix} 12 & 2 & 19 & 4 \\ 7 & 13 & 10 & 1 \\ 18 & 11 & 10 & 17 \end{bmatrix}$$

Much like vector transposes, a <u>matrix transpose</u> turns the rows into columns and vice-versa. Every element in row i and column j of M is the same as the element in row j and column i of M^T.

A matrix B is <u>symmetric</u> if $B = B^T$ (symmetric along the diagonal).

Identity Matrix

The identity matrix is an $n \times n$ square matrix with ones along the diagonal and zeros everywhere else. For instance, the 4×4 identity matrix looks like this:

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The rows and columns of the identity matrix are the canonical basis vectors for n-dimensional space. The identity matrix is always symmetric.

Identity Matrix

The matrix I is called the identity matrix because it preserves matrices under multiplication. For instance, if we have

$$A = \begin{bmatrix} -5 & 5 & -4 \\ -10 & 3 & 6 \end{bmatrix}$$

then $IA = AI = A$:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -5 & 5 & -4 \\ -10 & 3 & 6 \end{bmatrix} = \begin{bmatrix} -5 & 5 & -4 \\ -10 & 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -5 & 5 & -4 \\ -10 & 3 & 6 \end{bmatrix}$$

Matrix Inverse

Given a square matrix A, the matrix inverse A^{-1} has the property

$$A \cdot A^{-1} = A^{-1} \cdot A = I$$

For instance, if

$$A = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}, \qquad \text{then } A^{-1} = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -2 & 1 \end{bmatrix}$$

and we can see $AA^{-1} = A^{-1}A = I$:

$$\begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Unitary Matrix

A unitary matrix U has the property that its transpose is the same as its inverse, i.e. $U^T = U^{-1}$. For example,

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

is a unitary matrix because $U^T = U^{-1}$.

$$\begin{aligned} UU^T = UU^{-1} &= \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Row and Column Space

- The row space of a matrix is the space spanned by the row vectors in that matrix.
- Column space is defined similarly for the columns of the matrix.
- The rank of a matrix is the dimension of either the row space or the column space.
- A matrix has full rank if it has the largest possible rank for a matrix of the same dimensions.
- A square matrix has an inverse (is invertible) if and only if it is full rank. Otherwise, the matrix is non-invertible, a.k.a. singular.

Kernel / Nullspace

The kernel or nullspace of a matrix A (denoted $\ker(A)$) is the set of vectors $\boldsymbol{x}$ such that

$$A\boldsymbol{x} = \mathbf{0}$$

where $\mathbf{0}$ is the zero vector.

Example: What is the nullspace of A?

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \\ 6 & 3 & 9 \end{bmatrix}$$

Kernel / Nullspace

The kernel or nullspace of a matrix A (denoted $\ker(A)$) is the set of vectors $\boldsymbol{x}$ such that

$$A\boldsymbol{x} = \mathbf{0}$$

where $\mathbf{0}$ is the zero vector.

Example: What is the nullspace of A?

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \\ 6 & 3 & 9 \end{bmatrix}$$

The nullspace of A is the set of scalar multiples of

$$\boldsymbol{x} = \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}$$

because

$$A\boldsymbol{x} = \begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \\ 6 & 3 & 9 \end{bmatrix} \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Rank-Nullity Theorem

The nullity of a matrix is the dimension of its nullspace.

The rank-nullity theorem states that, if A is an $m \times n$ matrix, then

$$\mathrm{rk}(A) + \mathrm{nul}(A) = n$$

In other words, the rank of A plus the nullity of A should equal the number of columns in the matrix.

For instance, when

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & 5 & 6 \\ 6 & 3 & 9 \end{bmatrix}$$

the rank of A is 2, because there are 2 linearly independent rows, while the nullity of A is 1, since its kernel has dimension 1 (from the last slide). This sums to 3, the number of columns of A.

Trace

The trace of a square matrix A is the sum of the elements along the diagonal (upper-left to lower-right) of A. This is denoted $\text{tr}(A)$. For instance, with the matrix

$$A = \begin{bmatrix} 8 & 19 & 3 & 7 & 10 \\ 15 & 14 & 15 & 19 & 18 \\ 2 & 2 & 19 & 15 & 3 \\ 13 & 19 & 5 & 16 & 16 \\ 9 & 15 & 9 & 3 & 20 \end{bmatrix}$$

the trace of A is

$$\text{tr}(A) = 8 + 14 + 19 + 16 + 20 = 77$$

Determinants

Determinants are useful quantities for analyzing matrices and systems of linear equations. The determinant can only be found for square matrices, and is defined for a 2×2 matrix as:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

The determinants of larger square matrices can be found through *cofactor expansion*. For example:

$$\begin{vmatrix} 4 & 5 & 0 \\ 6 & 1 & 8 \\ 1 & 4 & 6 \end{vmatrix} = 4 \cdot \begin{vmatrix} 1 & 8 \\ 4 & 6 \end{vmatrix} - 5 \cdot \begin{vmatrix} 6 & 8 \\ 1 & 6 \end{vmatrix} + 0 \cdot \begin{vmatrix} 6 & 1 \\ 1 & 4 \end{vmatrix}$$

$$= 4(6 - 32) - 5(36 - 8) = -104 - 140 = -244$$

Determinants

A key property of non-invertible matrices is that they have determinant *zero*. For example:

$$\begin{aligned}\begin{vmatrix}1 & 2 & 3\\ 2 & 4 & 6\\ 5 & 4 & 9\end{vmatrix} &= 1\cdot\begin{vmatrix}4 & 6\\ 4 & 9\end{vmatrix} - 2\cdot\begin{vmatrix}2 & 6\\ 5 & 9\end{vmatrix} + 3\cdot\begin{vmatrix}2 & 4\\ 5 & 4\end{vmatrix}\\ &= (36-24) - 2(18-30) + 3(8-20)\\ &= 12 + 24 - 36 = 0\end{aligned}$$

Determinants

If we consider an $n \times n$ matrix as a collection of column vectors, then the determinant of that matrix is analogous to:

- $n = 2$: The area of the parallelogram created by the column vectors.
- $n = 3$: The volume of the paraellelpiped created by the column vectors.
- $n > 3$: The n-dimensional volume of the parallelotope created by the column vectors.

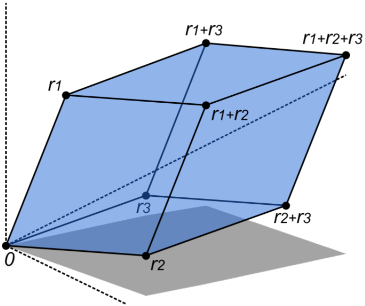


Image Source

Invertible vs. Non-Invertible Matrices

$$A = \begin{bmatrix} 4 & 5 & 0 \\ 6 & 1 & 8 \\ 1 & 4 & 6 \end{bmatrix}$$

- A has 3 linearly independent rows.
- A has rank 3, since the row space has 3 dimensions.
- A is full rank and has nonzero determinant, and therefore invertible.

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 5 & 4 & 9 \end{bmatrix}$$

- B has only 2 linearly independent rows (the second row is a multiple of the first row).
- B has rank 2, since the row space has 2 dimensions.
- B is not full rank and has zero determinant, and therefore non-invertible.

System of Linear Equations

How can we make use of all this matrix math?

Consider the following problem. A movie theater sells tickets at these prices:

$$\begin{aligned} \text{Adults:} &\quad \$10 \\ \text{Children:} &\quad \$7 \end{aligned}$$

If 1000 people showed up during the day, and the theater sold \$8,800 worth of tickets, how many adults and how many children went to that theater today?

System of Linear Equations

How can we make use of all this matrix math?

Consider the following problem. A movie theater sells tickets at these prices:

$$\begin{aligned} \text{Adults:} \quad & \$10 \\ \text{Children:} \quad & \$7 \end{aligned}$$

If 1000 people showed up during the day, and the theater sold \$8,800 worth of tickets, how many adults and how many children went to that theater today?

We can model this problem using a *system of linear equations*. If a is the number of adults and c is the number of children, then

$$\begin{aligned} a + c &= 1000 \\ 10a + 7c &= 8800 \end{aligned}$$

System of Linear Equations

$$
\begin{aligned}
a + c &= 1000 \\
10a + 7c &= 8800
\end{aligned}
$$

We can solve this system using back-substitution:

$$
\begin{aligned}
c &= 1000 - a \\
8800 &= 10a + 7(1000 - a) \\
8800 &= 10a - 7a + 7000 \\
1800 &= 3a \\
600 &= a \\
c &= 1000 - a = 1000 - 600 \\
c &= 400
\end{aligned}
$$

From this, we know 400 children and 600 adults attended that theater today.

System of Linear Equations as Matrix Equation

We can take a linear system of equations and re-interpret it as a matrix equation:

$$A\boldsymbol{x} = \boldsymbol{b}$$

where A is our matrix of coefficients, $\boldsymbol{x}$ is a vector with our variables, and $\boldsymbol{b}$ contains the remaining constants.

If we let x_1 be the number of adults and x_2 be the number of children, we can rewrite the linear system

$$\begin{aligned} x_1 + x_2 &= 1000 \\ 10x_1 + 7x_2 &= 8800 \end{aligned}$$

as $A\boldsymbol{x} = \boldsymbol{b}$, where

$$A = \begin{bmatrix} 1 & 1 \\ 10 & 7 \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1000 \\ 8800 \end{bmatrix}$$

System of Linear Equations as Matrix Equation

We can solve for the vector $\boldsymbol{x}$ in the equation $A\boldsymbol{x} = \boldsymbol{b}$ by multiplying both sides by A^{-1} on the left:

$$A^{-1}A\boldsymbol{x} = \boldsymbol{x} = A^{-1}\boldsymbol{b}$$

If we apply this to our linear system:

$$A = \begin{bmatrix} 1 & 1 \\ 10 & 7 \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} -2.333 & 0.333 \\ 3.333 & -0.333 \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1000 \\ 8800 \end{bmatrix}$$

$$A\boldsymbol{x} = \boldsymbol{b} \quad \rightarrow \quad \begin{bmatrix} 1 & 1 \\ 10 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1000 \\ 8800 \end{bmatrix}$$

$$\boldsymbol{x} = A^{-1}\boldsymbol{b} \quad \rightarrow \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2.333 & 0.333 \\ 3.333 & -0.333 \end{bmatrix} \begin{bmatrix} 1000 \\ 8800 \end{bmatrix} = \begin{bmatrix} 600 \\ 400 \end{bmatrix}$$

We get the same answer: 600 adults and 400 children attended that theater today.

Gaussian Elimination

Systems of equations and matrix inverses can also be found through a process called Gaussian elimination. With Gaussian elimination, we can turn a system of equations into an augmented matrix, with the coefficients on one side and the constants on the other. Gaussian elimination works with the following rules:

- You are allowed to multiply any row by a scalar.
- You are allowed to take any row, multiplied by any scalar, and add it to another row.
- Ideally, you try to reduce the left side of the augmented matrix until it is the identity matrix, or failing that, in *reduced row echelon form* (RREF).

Gaussian Elimination

Suppose we want to solve this system through Gaussian elimination:

$$
\begin{aligned}
x_1 + 5x_2 &= 7 \\
-2x_1 - 7x_2 &= -5
\end{aligned}
$$

Gaussian Elimination

Suppose we want to solve this system through Gaussian elimination:

$$\begin{aligned} x_1 + 5x_2 &= 7 \\ -2x_1 - 7x_2 &= -5 \end{aligned}$$

We can start by turning this into an augmented matrix:

$$\left[\begin{array}{cc|c} 1 & 5 & 7 \\ -2 & -7 & -5 \end{array}\right]$$

Multiply the first row by 2 and add that to the second row:

$$\left[\begin{array}{cc|c} 1 & 5 & 7 \\ 0 & 3 & 9 \end{array}\right]$$

Multiply the second row by $\frac{1}{3}$:

$$\left[\begin{array}{cc|c} 1 & 5 & 7 \\ 0 & 1 & 3 \end{array}\right]$$

Multiply the second row by -5 and add it to the first row:

$$\left[\begin{array}{cc|c} 1 & 0 & -8 \\ 0 & 1 & 3 \end{array}\right]$$

This gives us

$$\begin{aligned} x_1 &= -8 \\ x_2 &= 3 \end{aligned}$$

Gaussian Elimination

What if we want to find the inverse of matrix A through Gaussian elimination?

$$A = \begin{bmatrix} 7 & 3 \\ 5 & 2 \end{bmatrix}$$

Gaussian Elimination

What if we want to find the inverse of matrix A through Gaussian elimination?

$$A = \begin{bmatrix} 7 & 3 \\ 5 & 2 \end{bmatrix}$$

Create an augmented matrix with the identity matrix on the right:

$$\left[\begin{array}{cc|cc} 7 & 3 & 1 & 0 \\ 5 & 2 & 0 & 1 \end{array}\right]$$

Multiply the first row by $\frac{1}{7}$:

$$\left[\begin{array}{cc|cc} 1 & \frac{3}{7} & \frac{1}{7} & 0 \\ 5 & 2 & 0 & 1 \end{array}\right]$$

Multiply the first row by -5 and add that to the second row:

$$\left[\begin{array}{cc|cc} 1 & \frac{3}{7} & \frac{1}{7} & 0 \\ 0 & -\frac{1}{7} & -\frac{5}{7} & 1 \end{array}\right]$$

Multiply the second row by -7:

$$\left[\begin{array}{cc|cc} 1 & \frac{3}{7} & \frac{1}{7} & 0 \\ 0 & 1 & 5 & -7 \end{array}\right]$$

Multiply the second row by $-\frac{3}{7}$ and add that to the first row:

$$\left[\begin{array}{cc|cc} 1 & 0 & -2 & 3 \\ 0 & 1 & 5 & -7 \end{array}\right]$$

$$A^{-1} = \begin{bmatrix} -2 & 3 \\ 5 & -7 \end{bmatrix}$$

Linear Transformations

Matrices can be used to define linear transformations as applied to vectors.

For instance, if we have a vector $\boldsymbol{v}$ in the plane (but defined in 3D coordinates):

$$\boldsymbol{v} = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}$$

and we multiply $\boldsymbol{v}$ on the left by the "reflect about x-axis" matrix, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \\ 0 \end{bmatrix}$$

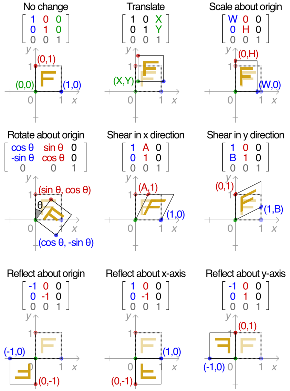
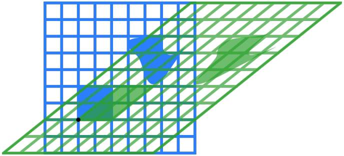


Image Source

Linear Transformations

Top: horizontal shear.

$$\begin{bmatrix} 1 & 1.25 \\ 0 & 1 \end{bmatrix}$$



Bottom: reflection through the vertical axis.

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

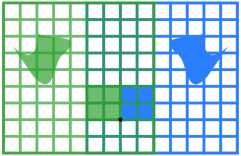
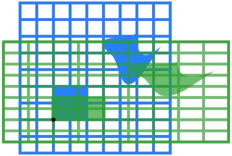


Image Source

Linear Transformations

Top: squeeze mapping.

$$\begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{2}{3} \end{bmatrix}$$



Bottom: rotation of 30°.

$$\begin{bmatrix} \cos(30^\circ) & -\sin(30^\circ) \\ \sin(30^\circ) & \cos(30^\circ) \end{bmatrix}$$

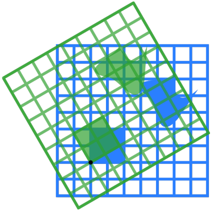


Image Source

Eigenvectors and Eigenvalues

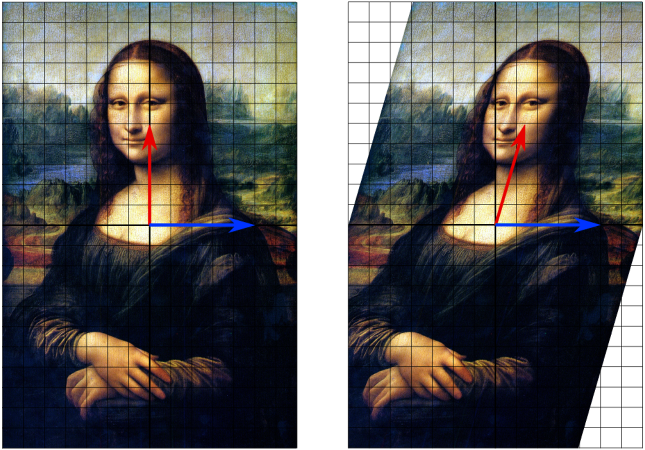


Image Source

Eigenvectors and Eigenvalues

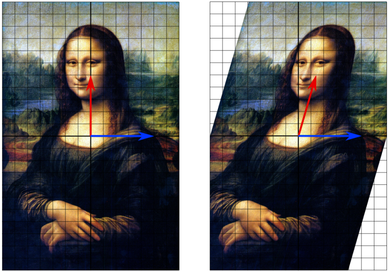


Image Source

A shear transformation as depicted above will change the direction of some vectors (like the red vector), but keep other vectors pointed in their original direction (like the blue vector). A vector whose direction is unchanged by the linear transformation of a matrix is called an eigenvector of that matrix, and the amount by which it is scaled is called its eigenvalue.

Eigenvectors and Eigenvalues

A matrix A has eigenvector $\boldsymbol{v}$ with eigenvalue λ if

$$A\boldsymbol{v} = \lambda \boldsymbol{v}.$$

This is often equivalently written as

$$(A - \lambda I)\boldsymbol{v} = \mathbf{0} \leftarrow \text{(the zero vector)}.$$

This equation has a solution only if the determinant of the matrix $(A - \lambda I)$ is zero, so we can find the eigenvalues and eigenvectors of a matrix by solving the following for λ:

$$\det(A - \lambda I) = |A - \lambda I| = 0$$

Eigenvectors and Eigenvalues

For example, what are the eigenvectors and eigenvalues for the following matrix?

$$A = \begin{bmatrix} 2 & 7 \\ -1 & -6 \end{bmatrix}$$

Eigenvectors and Eigenvalues

For example, what are the eigenvectors and eigenvalues for the following matrix?

$$A = \begin{bmatrix} 2 & 7 \\ -1 & -6 \end{bmatrix}$$

To find the eigenvalues, we first need to solve $\det(A - \lambda I) = 0$ for λ.

$$\begin{aligned} \det(A - \lambda I) &= \det\left(\begin{bmatrix} 2 & 7 \\ -1 & -6 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \\ &= \begin{vmatrix} 2-\lambda & 7 \\ -1 & -6-\lambda \end{vmatrix} \\ &= (2-\lambda)(-6-\lambda) - (-7) \\ &= \lambda^2 + 4\lambda - 5 \\ &= (\lambda+5)(\lambda-1) \end{aligned}$$

This gives us the characteristic polynomial for the matrix, $(\lambda + 5)(\lambda - 1)$. When we set this equal to zero, we get the eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -5$.

Eigenvectors and Eigenvalues

To find eigenvectors, plug the eigenvalues back in and find vectors $\boldsymbol{v}$ that solve $(A - \lambda I)\boldsymbol{v} = \mathbf{0}$.

For $\lambda_1 = 1$:

$$\begin{aligned} \boldsymbol{v}_1 &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ (A - 1I)\boldsymbol{v}_1 &= \begin{bmatrix} 1 & 7 \\ -1 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ x_1 + 7x_2 &= 0 \\ x_1 &= -7x_2 \end{aligned}$$

There are an infinite number of solutions, but one possible solution is

$$\boldsymbol{v}_1 = \begin{bmatrix} -7 \\ 1 \end{bmatrix}$$

Eigenvectors and Eigenvalues

To find eigenvectors, plug the eigenvalues back in and find vectors $\boldsymbol{v}$ that solve $(A - \lambda I)\boldsymbol{v} = \mathbf{0}$.

For $\lambda_2 = -5$:

$$\begin{aligned}
\boldsymbol{v}_2 &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\
(A + 5I)\boldsymbol{v}_2 &= \begin{bmatrix} 7 & 7 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
7y_1 + 7y_2 &= 0 \\
y_1 &= -y_2
\end{aligned}$$

One possible solution is

$$\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Eigenvectors and Eigenvalues

- The trace of a square matrix is equal to the sum of its eigenvalues.
- The determinant of a square matrix is equal to the product of its eigenvalues.
- A non-invertible (singular) square matrix will have 0 as an eigenvalue. Otherwise, an invertible (non-singular) matrix will not have 0 as an eigenvalue.
- Eigenvectors with distinct eigenvalues are linearly independent.
- If A is nonsingular and has λ as an eigenvalue, then A^{-1} has λ^{-1} as an eigenvalue.
- The eigenvalues of A are the same as the eigenvalues of A^T.

Positive or Negative (Semi-) Definite Matrices

For the following bullet points, A is an $n \times n$ symmetric matrix, and $\boldsymbol{x}$ is an n-dimensional vector.

- The matrix A is positive definite if $\boldsymbol{x}^T A\boldsymbol{x} > 0$ for all $\boldsymbol{x}$. Similarly, A is negative definite if $\boldsymbol{x}^T A\boldsymbol{x} < 0$ for all $\boldsymbol{x}$.
- A is positive semidefinite if $\boldsymbol{x}^T A\boldsymbol{x} \geq 0$ for all $\boldsymbol{x}$. Similarly, A is negative semidefinite if $\boldsymbol{x}^T A\boldsymbol{x} \leq 0$ for all $\boldsymbol{x}$.
- If A is square and symmetric, but doesn't fit into the above 4 categories, then A is an indefinite matrix.
- If A is positive definite, then all eigenvalues of A are positive.

Similar Matrices

Matrices A and B are considered similar if there exists a square invertible matrix P such that

$$A = P^{-1}BP$$

Similar matrices represent the same linear transformation, but using different basis vectors. Similar matrices have the same

- rank.
- determinant.
- eigenvalues.

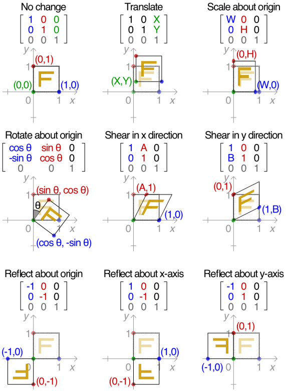


Image Source

Principal Axis Theorem

The Principal Axis Theorem states that, given a real symmetric $n \times n$ matrix A, there exists an orthonormal basis for n-dimensional space made up of eigenvectors of A.

This fact will come in handy when we explore *principal component analysis* (PCA), as the principal components come from the eigenvectors and eigenvalues of the covariance matrix (which is symmetric).

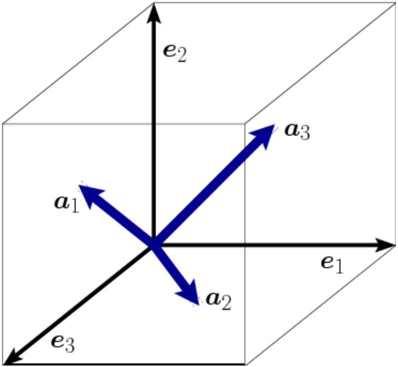


Image Source