AI Summer Camp: Linear Algebra

Elsiddig Elsiddig

King Abdullah University of Science and Technology, Saudi Arabia

July 15, 2023

Vectors

A **vector** is a quantity that has a magnitude and direction.

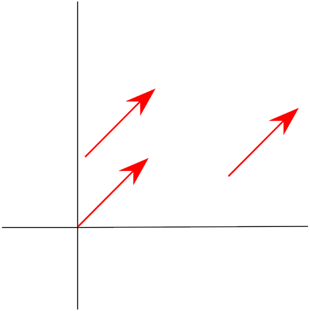
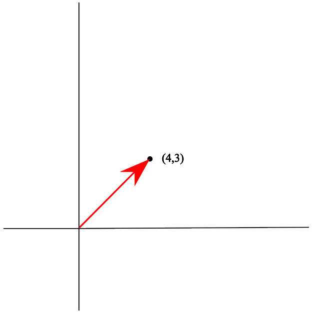


Figure: Representations of the same vector

In the plane we can represent a vector by a pair of real numbers (x, y). The set of such pairs are denoted $\mathbb{R}^2$.



It is staight forward to generalize to $\mathbb{R}^n$ for any integer $n \geq 1$. A **vector x** in $\mathbb{R}^n$ is a sequence of n real numbers $\mathbf{x} = (x_1, ..., x_n)$. n is called the dimension of the space or of the vector. The numbers x_i's are called the components of the vector $\mathbf{x}$.

Although a vector is independent from its starting point is we can sill represent a vector by a starting point and an end point. The vector in this case will just be the vector represented by the subtraction of the components of the points.

Example: The vector that starts at $(3, 1)$ and ends at $(0, 5)$ is represented by

$$(0 - 3, 5 - 1) = (-2, 4)$$

.

Vector Operations

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
By definition we can write $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_n)$.

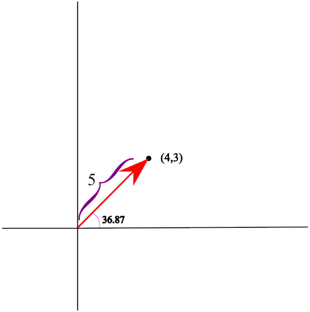
- We say that $\mathbf{x} = \mathbf{y}$ if and only if $x_1 = y_1, x_2 = y_2, ..., x_n = y_n$.
- Addition: $\mathbf{x} + \mathbf{y} = (x_1 + y_1, ..., x_n + y_n)$.
- Scalar Multiplication: $c\mathbf{x} = (cx_1, ..., cx_n)$.
- Length (Norm): $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$.
- Dot (Scalar) Product: $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + ... + x_ny_n$.

Notice that some operations return a vector while some return a real number.

The vector $(4, 3)$ has length

$$\|(4,3)\| = \sqrt{4^2 + 3^2} = 5$$

and its angle $\theta \approx 36.87$ which is the angle that satisfies $\cos(\theta) = 4/5$ and $\sin(\theta) = 3/5$.



Angle between two vectors

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. The angle θ between the vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as the unique angle in the interval $[0, \pi]$ that satisfies

$$\cos(\theta) = \frac{\mathbf{x} \cdot \mathbf{y}}{\|x\| \|y\|}$$

Why do only define the angle on $[0, \pi]$?

Unit Vectors

A vector $\mathbf{x} \in \mathbb{R}^n$ that satisfies $\|x\| = 1$ is called a unit vector. Writing this using the vector components

$$x_1^2 + x_2^2 + ... + x_n^2 = 1$$

This is called the n-dimensional unit sphere equation.

Matrix

An $n \times m$ matrix is a sequence of length$n$ of m-vectors written as rows. A matrix can be viewed as a sequence of length m of n-vectors written as columns.

$$\begin{bmatrix} 2 & 5 & 0.3 \\ 5 & 6 & 10 \\ -100 & \frac{5}{3} & 0 \\ -2 & 1 & 1 \end{bmatrix}$$

Addition of matrices and scalar multiplication are defined analogously to way they were defined for vectors.

We denote the set of $n \times m$ matrices by $\mathbb{R}^{n \times m}$

Matrix Multiplication

$$
\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \times \begin{bmatrix} 10 & 11 \\ 20 & 21 \\ 30 & 31 \end{bmatrix}
$$

$$
= \begin{bmatrix} 1\times10 + 2\times20 + 3\times30 & 1\times11 + 2\times21 + 3\times31 \\ 4\times10 + 5\times20 + 6\times30 & 4\times11 + 5\times21 + 6\times31 \end{bmatrix}
$$

$$
= \begin{bmatrix} 10+40+90 & 11+42+93 \\ 40+100+180 & 44+105+186 \end{bmatrix} = \begin{bmatrix} 140 & 146 \\ 320 & 335 \end{bmatrix}
$$

Identity Matrix

Identity Matrix I_n of dimension n is defined as

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}_{n \times n}$$

The identity is the neutral element of matrix multiplication. It satisfies

$$I_n A = A$$

$$B I_n = B$$

whenever the product is defined for the matrices A and B.

More Matrix Operations

Transpose:

$$\begin{bmatrix} 2 & 5 & 0.3 \\ 5 & 6 & 10 \\ -100 & \frac{5}{3} & 0 \\ -2 & 1 & 1 \end{bmatrix}^T = \begin{bmatrix} 2 & 5 & -100 & -2 \\ 5 & 6 & \frac{5}{3} & 1 \\ 0.3 & 10 & 0 & 1 \end{bmatrix}$$

Inverse: Let $A \in \mathbb{R}^{n\times n}$. A matrix $B \in \mathbb{R}^{n\times n}$ is called the inverse of A if it satisfies

$$AB = I_n$$

and

$$BA = I_n$$

Verifying one of the two equations above implies the other. Hence we only need to check one (any of them).

Some matrices do not have an inverse! When the inverse exists we denote it by A^{-1}

Inverse for $2 \times 2 matrices$

Let

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

The matrix A has an inverse if and only if $ad - bc \neq 0$ and the inverse is given by

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Linear Systems and Matrices

Consider the following linear system

$$\begin{cases} 4x + 3y = -1 \\ 3x - 2y = 12 \end{cases}$$

Let

$$A = \begin{bmatrix} 4 & 3 \\ 3 & -2 \end{bmatrix}, \; X = \begin{bmatrix} x \\ y \end{bmatrix}, \; b = \begin{bmatrix} -1 \\ 12 \end{bmatrix}$$

We can write the linear system using matrix operations as

$$AX = b$$

Solving Linear Systems using Matrices

Suppose A has an inverse and let's multiply both side (from left) of the equation $AX = b$ by A^{-1}. We find

$$\begin{aligned} & A^{-1}AX = A^{-1}b \\ \Leftrightarrow\ & I_n X = A^{-1}b \\ \Leftrightarrow\ & X = A^{-1}b \end{aligned}$$

In our example $A^{-1} = \frac{1}{-17}\begin{bmatrix} -2 & -3 \\ -3 & 4 \end{bmatrix}$. We can calculate X as

$$X = A^{-1}b = \frac{1}{-17}\begin{bmatrix} -2 & -3 \\ -3 & 4 \end{bmatrix}\begin{bmatrix} -1 \\ 12 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

Which the same as saying $x = 2, y = -3$.

Line of Best Fit Again

Suppose we have points $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$. We would like to find the line of the form $y = mx + b$ that best fits the points. We already dealt with this problem using least square methods previously. Here will write an equivalent formulation that is perhaps easier to generalize. The problem of finding the line of best fit mildly asks for

$$\begin{cases} mx_1 + b \approx y_1 \\ mx_2 + b \approx y_2 \\ mx_3 + b \approx y_3 \\ \vdots \\ mx_n + b \approx y_n \end{cases}$$

In the same manner we wrote linear system in matrix format we can write

$$AP \approx \mathbf{y}$$

where

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, P = \begin{bmatrix} m \\ b \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

The meaning of the $\approx$ above can be formulated preisely in term of finding P that minimizes the quantity

$$\|AP - y\|^2$$

This is exactly the least squares problem but written compactly.

Explicit Solution

In the calculus section we provided a systematic method to solve the least squares problem. Here we provide an equivalent but direct to calculate formula. The parameters of the line of best fit are the result of the following

$$\left(A^T A\right)^{-1} A^T \mathbf{y}$$

Check the exercises for an example.