Linear Algebra

Lecture 1
Vectors, Dot Products

Yiping Lu
Based on Dr. Ralph Chikhany's Slide

Logistics

- Course Website: https://2prime.github.io/teaching/2024-linear-algebra
- (anonymous) form: https://forms.gle/Dtw6PRFdnbk8NQWRA

- **Textbook:** Introduction to Linear Algebra - Fifth Edition, Gilbert Strang
- **Reference**: http://web.mit.edu/18.06/www/

- **Grading:**
  - Attendance & Participation 5%
  - Quizzes 15%
  - Problem Sets 10%
  - Exams 70%

Homework

- **6 Problem Sets**
  - Latex and overleaf (not required)
- Late work policy:
  - For your first late assignment within 12 hours after the deadline (as indicated on Gradescope), no point deductions.
  - All subsequent assignments submitted within 12 hours after the deadline will convert to a zero at the end of semester.
  - In all cases, work submitted 12 hours or more after the deadline will not be accepted.

Overview of the Course

Brightspace
Gradescope
Campuswire

What is due next week (and every week)

Access through Gradescope

Problem Set 1 – Friday 2/9 11.59 pm

(Late work policy applies)

Recap Quiz 1 – Sunday 2/4 11.59 am

(No late work accepted)

Note: Recap Quiz 1 is timed for 60 minutes to help you get used to the format.

Future quizzes will be timed for 30-45 minutes

Intro to the Course

What is Linear Algebra?

Linear

- having to do with lines/planes/etc.
- For example, $x + y + 3z = 7$, not $\sin, \log, x^2$, etc.

Algebra

- solving equations involving numbers and symbols
- from al-jebr (Arabic), meaning reunion of broken parts
- 9^{th} century Abu Ja'far Muhammad ibn Muso al-Khwarizmi

study of variables and the rules for manipulating these variables in formulas, rule of calculation

$$\begin{aligned} 2x + y &= 1 \\ x + y &= 1 \end{aligned} \quad \Rightarrow \quad A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, A = \begin{bmatrix} 2, 1 \\ 1, 1 \end{bmatrix}$$

Some Applications

Large classes of engineering problems, no matter how huge, can be reduced to linear algebra:

$$Ax = b \quad \text{or}$$
$$Ax = \lambda x$$

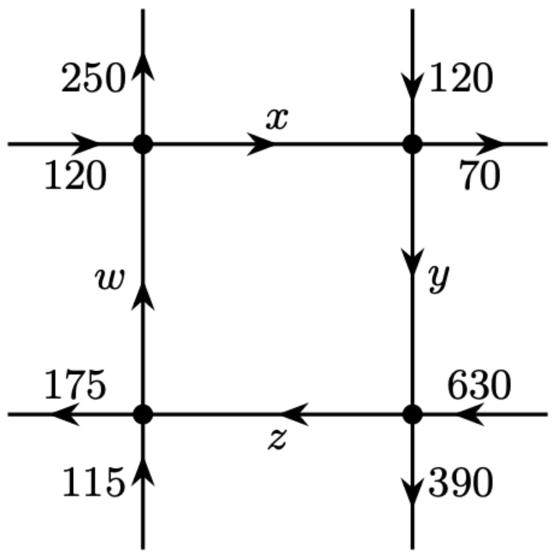
"...and now it's just linear algebra"

Civil Engineering: How much traffic flows through the four labeled segments?

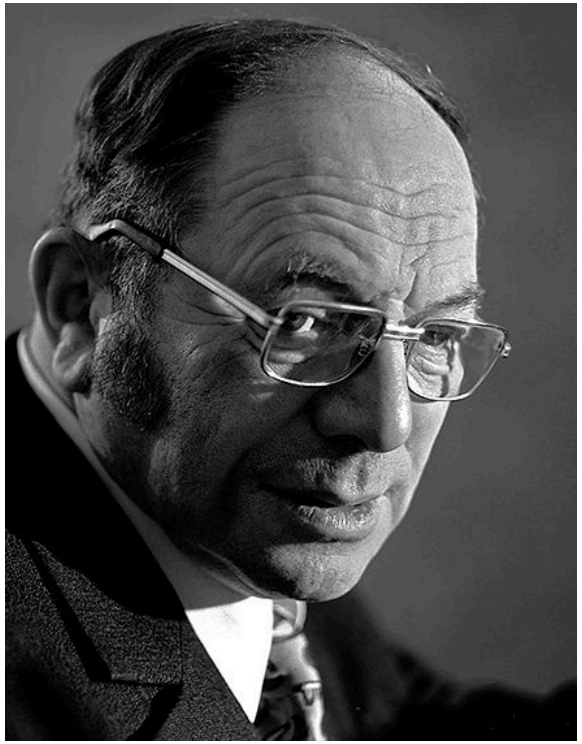
$\rightsquigarrow$ system of linear equations:

$$\begin{aligned} w + 120 &= x + 250 \\ x + 120 &= y + 70 \\ y + 630 &= z + 390 \\ z + 115 &= w + 175 \end{aligned}$$

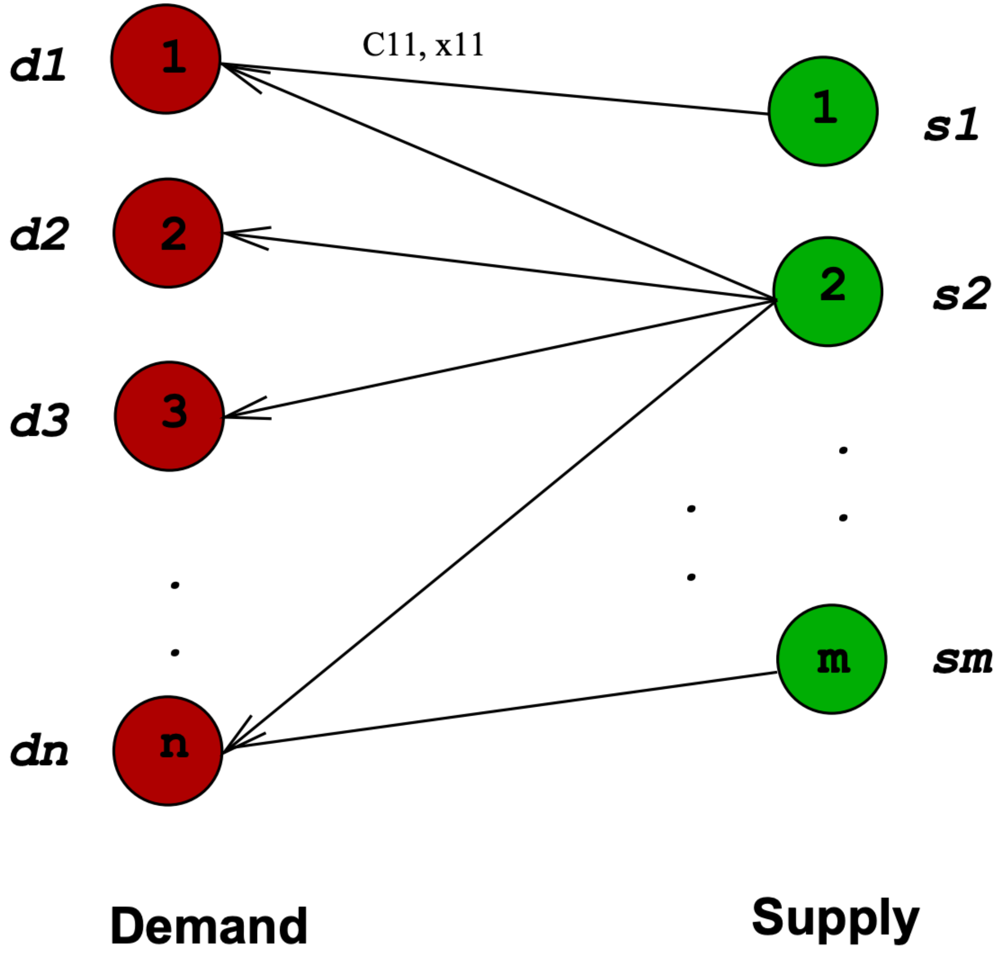
Traffic flow (cars/hr)



Linear Programming



Lenoid Kantorovich
Nobel Prize in Econ (1975)



Some Applications

Chemistry: Balancing reaction equations

$$\underline{\ x\ }\ C_2H_6 + \underline{\ y\ }\ O_2 \rightarrow \underline{\ z\ }\ CO_2 + \underline{\ w\ }\ H_2O$$

$\rightsquigarrow$ system of linear equations, one equation for each element.

$$\begin{aligned} 2x &= z \\ 6x &= 2w \\ 2y &= 2z \end{aligned}$$

Geometry and Astronomy: Find the equation of a circle passing through 3 given points, say $(1,0)$, $(0,1)$, and $(1,1)$. The general form of a circle is $a(x^2+y^2)+bx+cy+d=0$.

$\rightsquigarrow$ system of linear equations:

$$\begin{aligned} a + b \phantom{{}+c} + d &= 0 \\ a \phantom{{}+b} + c + d &= 0 \\ 2a + b + c + d &= 0 \end{aligned}$$

Very similar to: compute the orbit of a planet:

$$ax^2+by^2+cxy+dx+ey+f=0$$

Some Applications

Biology: In a population of rabbits...

- half of the new born rabbits survive their first year
- of those, half survive their second year
- the maximum life span is three years
- rabbits produce 0, 6, 8 rabbits in their first, second, and third years

If I know the population in 2016 (in terms of the number of first, second, and third year rabbits), then what is the population in 2017?

$\rightsquigarrow$ system of linear equations:

$$\begin{aligned} 6y_{2016} + 8z_{2016} &= x_{2017} \\ \tfrac{1}{2}x_{2016} \qquad\qquad\qquad &= y_{2017} \\ \tfrac{1}{2}y_{2016} \qquad\quad &= z_{2017} \end{aligned}$$

Question

Does the rabbit population have an asymptotic behavior? Is this even a linear algebra question? Yes, it is!

Some Applications

Biology: In a population of rabbits...

- half of the new born rabbits survive their first year
- of those, half survive their second year
- the maximum life span is three years
- rabbits produce 0, 6, 8 rabbits in their first, second, and third years

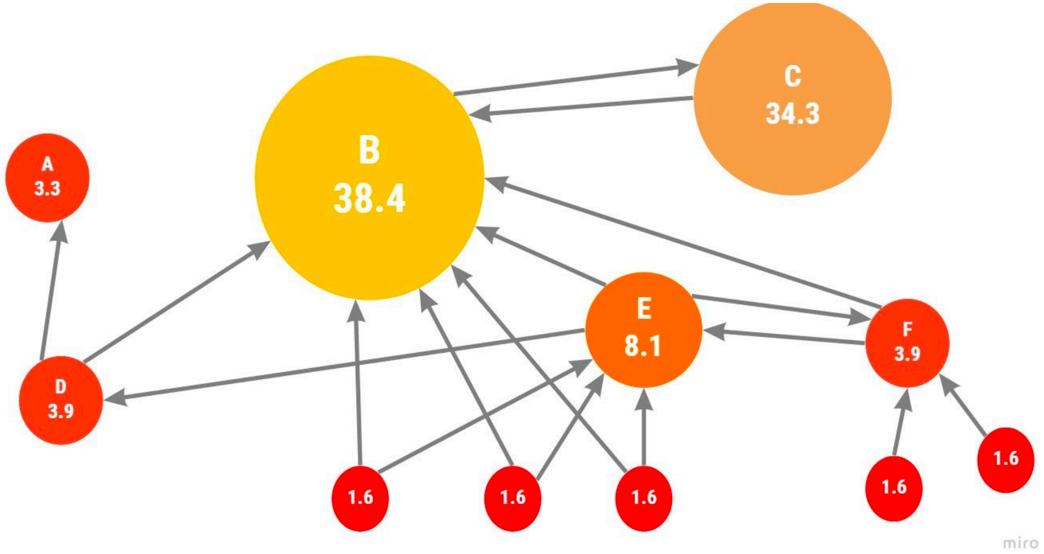
If I know the population in 2016 (in terms of the number of first, second, and third year rabbits), then what is the population in 2017?

⇝ system of linear equations:

$$\begin{aligned} 6y_{2016} + 8z_{2016} &= x_{2017} \\ \tfrac{1}{2}x_{2016} \qquad\qquad &= y_{2017} \\ \tfrac{1}{2}y_{2016} \qquad\quad &= z_{2017} \end{aligned}$$

Question

Does the rabbit population have an asymptotic behavior? Is this even a linear algebra question? Yes, it is!



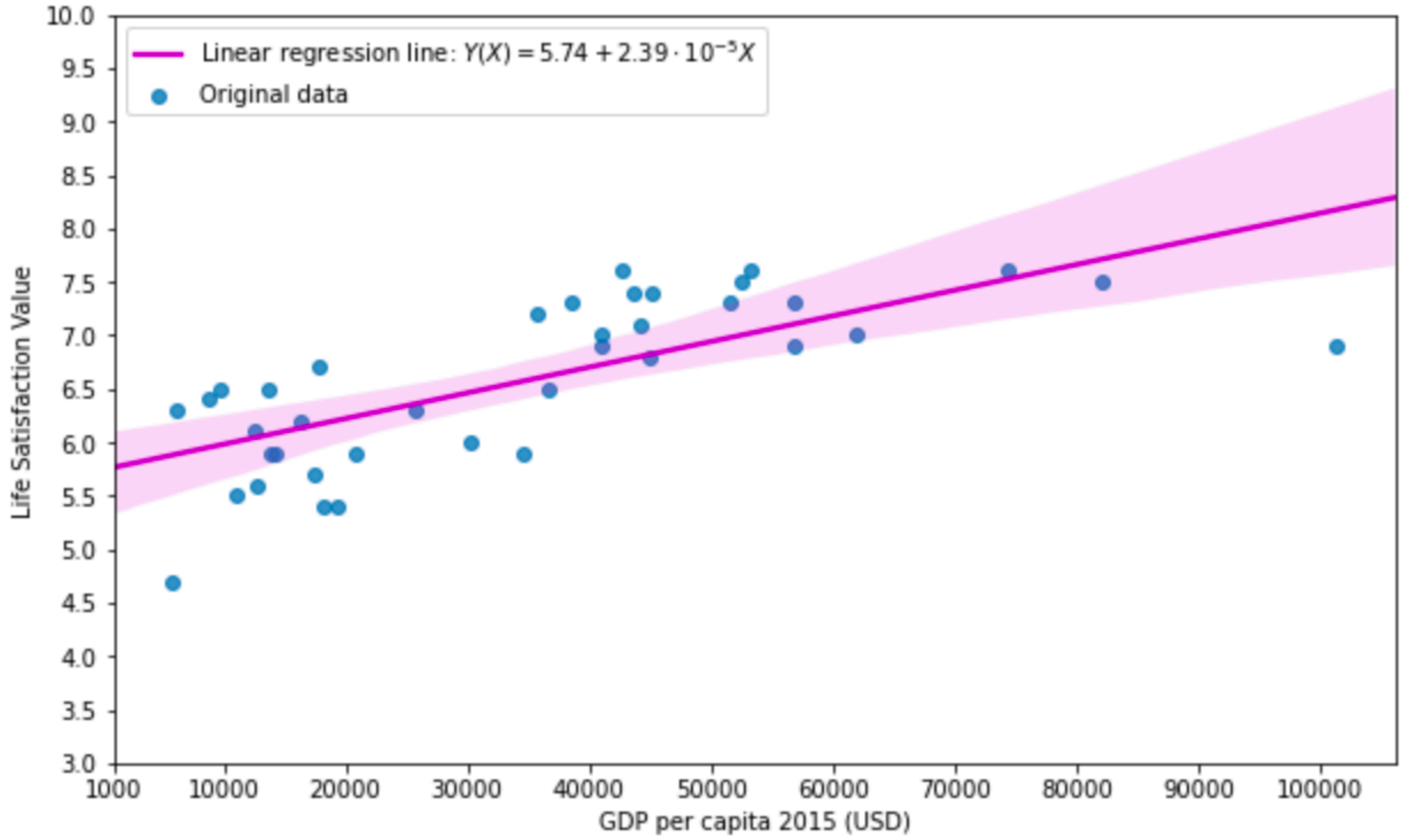
"Pagerank" Algorithm

Google: "The 25 billion dollar eigenvector." Each web page has some importance, which it shares via outgoing links to other pages ⇝ system of linear equations (in gazillions of variables).

Larry Page flies around in a private 747 because he paid attention in his linear algebra class!

Some Application

- Learning from data: https://math.mit.edu/classes/18.065/2019SP/

- 

find the best linear fit!

Overview of the Course

- Solve the matrix equation $Ax = b$
  - Solve systems of linear equations using matrices, row reduction, and inverses.
  - Solve systems of linear equations with varying parameters using parametric forms for solutions, the geometry of linear transformations, the characterizations of invertible matrices, and determinants.
- Solve the matrix equation $Ax = \lambda x$
  - Solve eigenvalue problems through the use of the characteristic polynomial.
  - Understand the dynamics of a linear transformation via the computation of eigenvalues, eigenvectors, and diagonalization.
- Almost solve the equation $Ax = b$
  - Find best-fit solutions to systems of linear equations that have no actual solution using least squares approximations.

Overview of the Course

Your previous math courses probably focused on how to do (sometimes rather involved) computations.

- Compute the derivative of $\sin(\log x)\cos(e^x)$.
- Compute $\int_0^1 (1-\cos(x))\,dx$.

This is important, **but** Wolfram Alpha can do all these problems better than any of us can. Nobody is going to hire you to do something a computer can do better.

> If a computer can do the problem better than you can, then it's just an algorithm: this is not real problem solving.

So what are we going to do?

- About half the material focuses on how to do linear algebra computations—that is still important.
- The other half is on *conceptual* understanding of linear algebra. This is much more subtle: it's about figuring out *what question* to ask the computer, or whether you actually need to do any computations at all.

NYU

Let's get this show started!

Strang Sections 1.1 and 1.2

Course notes adapted from *Introduction to Linear Algebra* by Strang (5th ed), and *Interactive Linear Algebra* by Margalit and Rabinoff.

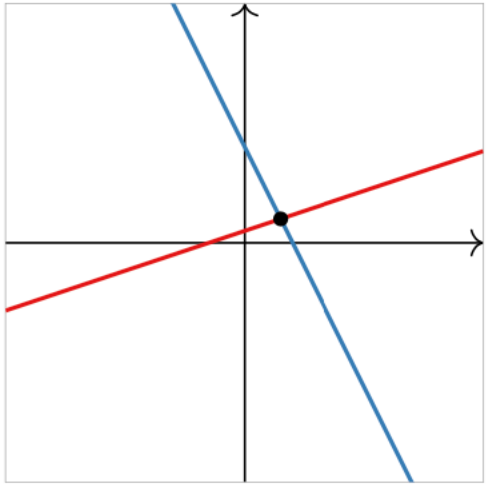
1.1 - Vectors

Course notes adapted from *Introduction to Linear Algebra* by Strang (5th ed), and *Interactive Linear Algebra* by Margalit and Rabinoff.

Motivation

We want to think about the *algebra* in linear algebra (systems of equations and their solution sets) in terms of *geometry* (points, lines, planes, etc).

$$\begin{aligned} x - 3y &= -3 \\ 2x + y &= 8 \end{aligned} \quad \rightsquigarrow$$



This will give us better insight into the properties of systems of equations and their solution sets.

To do this, we need to introduce n-dimensional space $\mathbf{R}^n$, and **vectors** inside it.

Motivation

Recall that $\mathbf{R}$ denotes the collection of all real numbers, i.e. the number line.

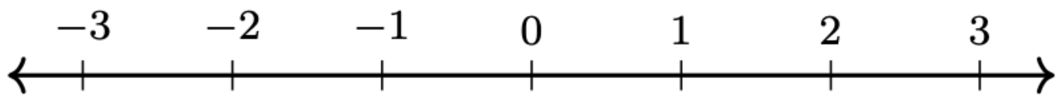
Definition

Let n be a positive whole number. We define

$$\mathbf{R}^n = \text{all ordered } n\text{-tuples of real numbers } (x_1, x_2, x_3, \ldots, x_n).$$

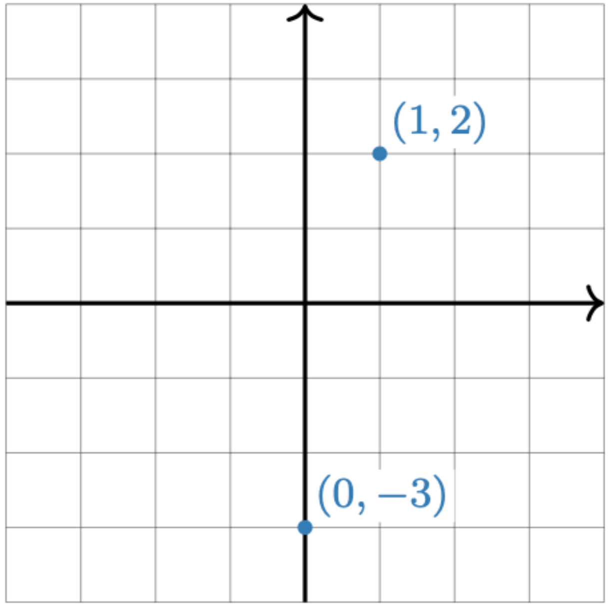
Example

When $n = 1$, we just get $\mathbf{R}$ back: $\mathbf{R}^1 = \mathbf{R}$. Geometrically, this is the *number line*.



Motivation

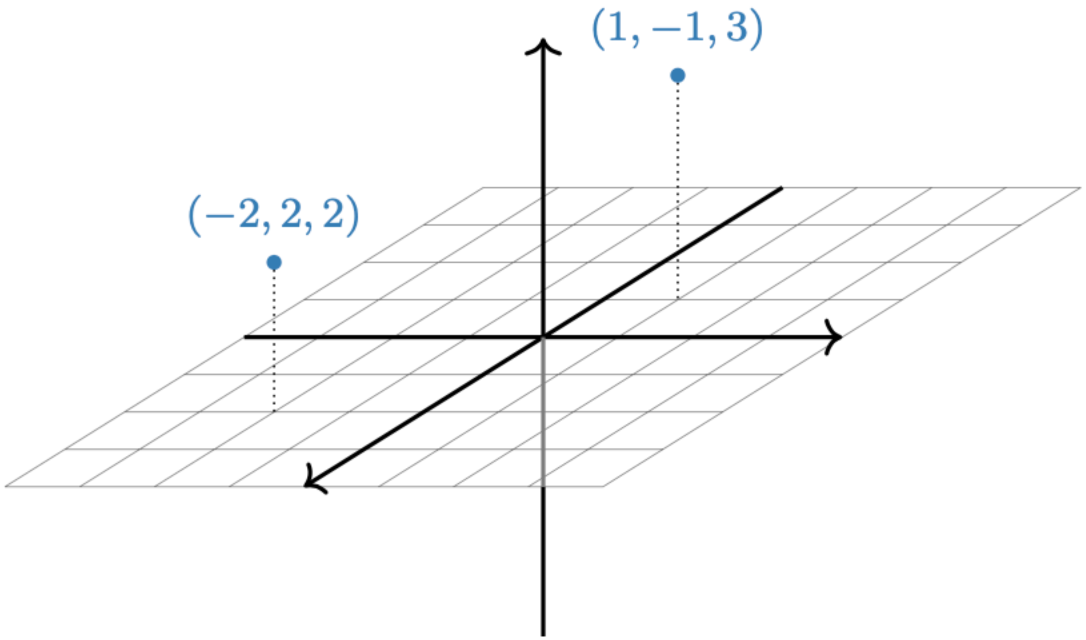
When $n = 2$, we can think of $\mathbf{R}^2$ as the *plane*. This is because every point on the plane can be represented by an ordered pair of real numbers, namely, its x- and y-coordinates.



We can use the elements of $\mathbf{R}^2$ to *label* points on the plane, but $\mathbf{R}^2$ is not defined to be the plane!

Motivation

When $n = 3$, we can think of $\mathbf{R}^3$ as the *space* we (appear to) live in. This is because every point in space can be represented by an ordered triple of real numbers, namely, its x-, y-, and z-coordinates.



Again, we can use the elements of $\mathbf{R}^3$ to *label* points in space, but $\mathbf{R}^3$ is not defined to be space!

Motivation

So what is $\mathbf{R}^4$? or $\mathbf{R}^5$? or $\mathbf{R}^n$?

... go back to the *definition*: ordered n-tuples of real numbers

$$(x_1, x_2, x_3, \ldots, x_n).$$

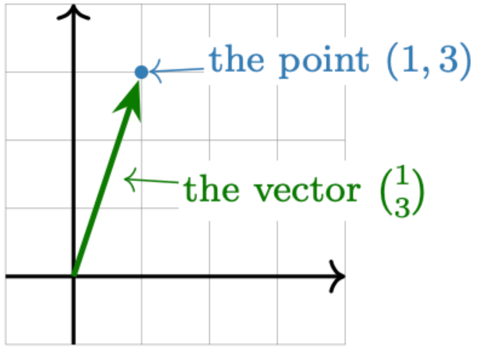
They're still "geometric" spaces, in the sense that our intuition for $\mathbf{R}^2$ and $\mathbf{R}^3$ sometimes extends to $\mathbf{R}^n$, but they're harder to visualize.

We'll make definitions and state theorems that apply to any $\mathbf{R}^n$, but we'll only draw pictures for $\mathbf{R}^2$ and $\mathbf{R}^3$.

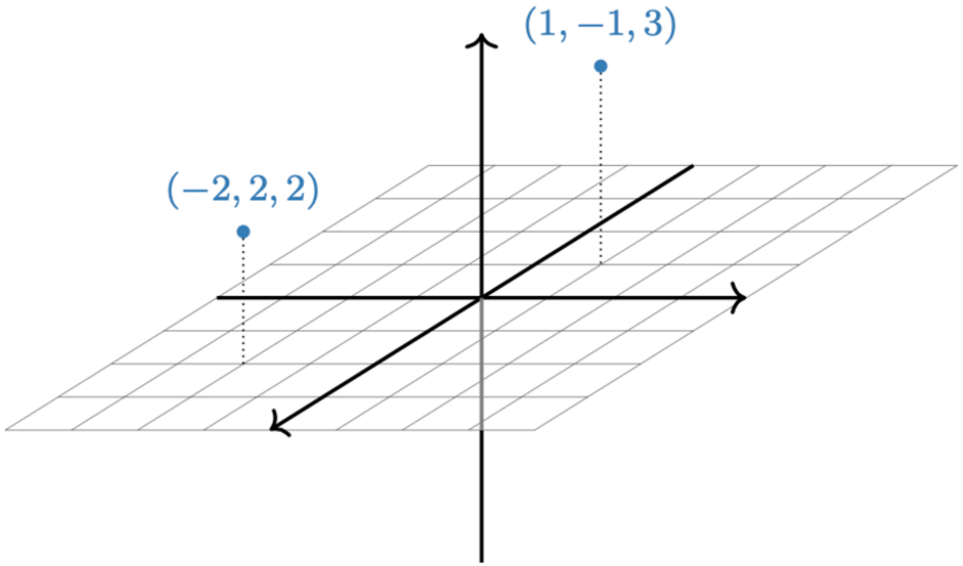
Vectors

In the previous slides, we were thinking of elements of $\mathbf{R}^n$ as **points**: in line, plane, space, etc.

We can also think of them as **vectors**: arrows with a given length and direction.



So the vector points *horizontally* in the amount of its x-coordinate, and *vertically* in the amount of its y-coordinate.



Imagine Manhattan

Vector Algebra

- We can add two vectors together:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a+x \\ b+y \\ c+z \end{pmatrix}.$$

- We can multiply, or **scale**, a vector by a real number c:

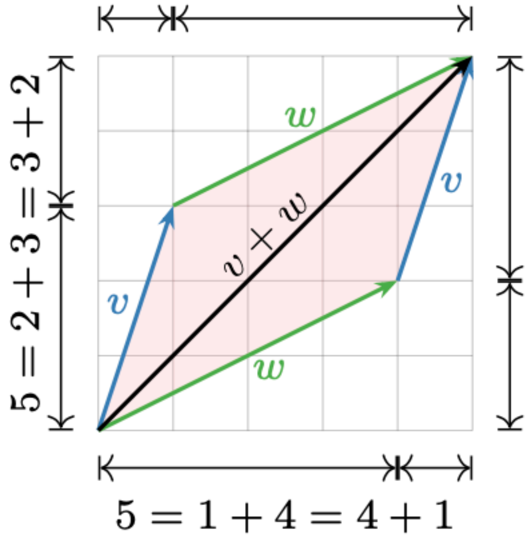
$$c \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} c \cdot x \\ c \cdot y \\ c \cdot z \end{pmatrix}.$$

  We call c a **scalar** to distinguish it from a vector. If v is a vector and c is a scalar, cv is called a **scalar multiple** of v.

(And likewise for vectors of length n.) For instance,

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix} \quad \text{and} \quad -2 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \\ -6 \end{pmatrix}.$$

Vector Addition and Subtraction

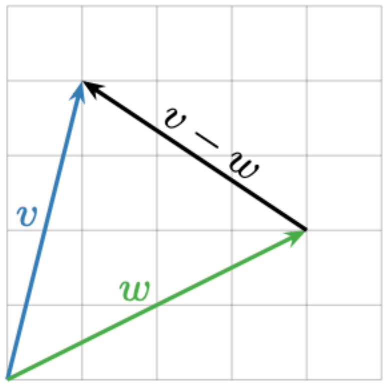


The parallelogram law for vector addition

Geometrically, the sum of two vectors v, w is obtained as follows: place the tail of w at the head of v. Then $v+w$ is the vector whose tail is the tail of v and whose head is the head of w. Doing this both ways creates a parallelogram. For example,

$$\begin{pmatrix}1\\3\end{pmatrix} + \begin{pmatrix}4\\2\end{pmatrix} = \begin{pmatrix}5\\5\end{pmatrix}.$$

Why? The width of $v+w$ is the sum of the widths, and likewise with the heights.



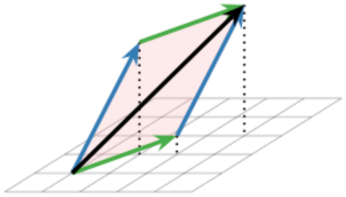
Vector subtraction

Geometrically, the difference of two vectors v, w is obtained as follows: place the tail of v and w at the same point. Then $v - w$ is the vector from the head of v to the head of w. For example,

$$\begin{pmatrix}1\\4\end{pmatrix} - \begin{pmatrix}4\\2\end{pmatrix} = \begin{pmatrix}-3\\2\end{pmatrix}.$$

Why? If you add $v - w$ to w, you get v.

This works in higher dimensions too!

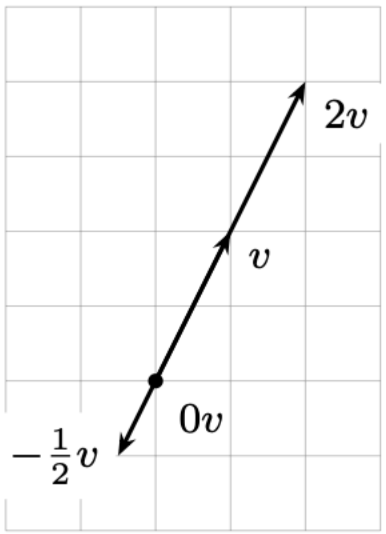


Scalar Multiplication - Geometry

Scalar multiples of a vector

These have the same *direction* but a different *length*.

Some multiples of v.



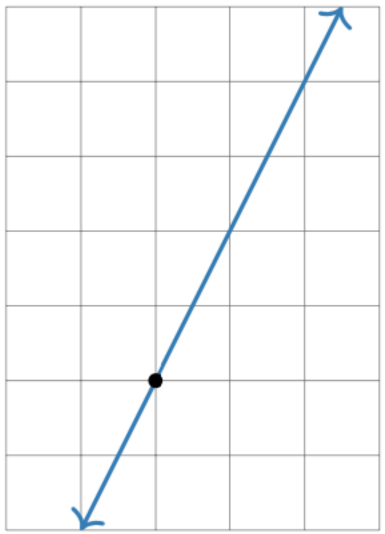
$$v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$2v = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

$$-\frac{1}{2}v = \begin{pmatrix} -\frac{1}{2} \\ -1 \end{pmatrix}$$

$$0v = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

All multiples of v.



So the scalar multiples of v form a *line*.

Linear Combinations

We can add and scalar multiply in the same equation:

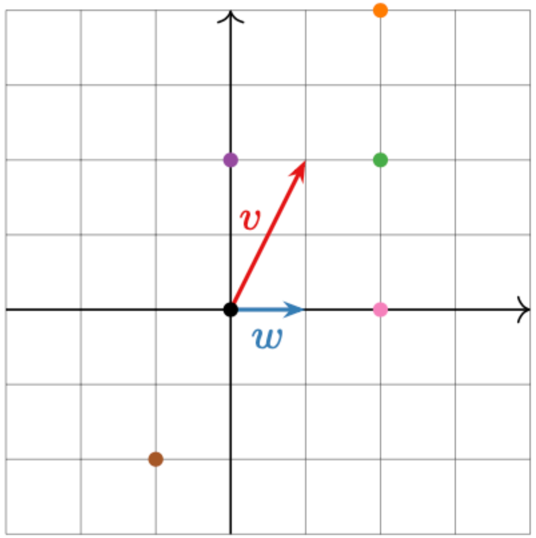
$$w = c_1v_1 + c_2v_2 + \cdots + c_pv_p$$

where $c_1, c_2, \ldots, c_p$ are scalars, $v_1, v_2, \ldots, v_p$ are vectors in $\mathbf{R}^n$, and w is a vector in $\mathbf{R}^n$.

Definition

We call w a **linear combination** of the vectors $v_1, v_2, \ldots, v_p$. The scalars $c_1, c_2, \ldots, c_p$ are called the **weights** or **coefficients**.

Example



Let $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

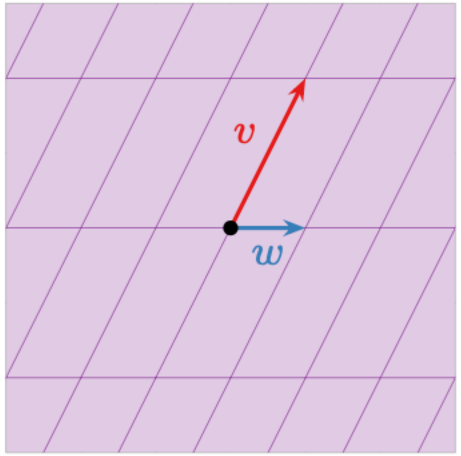
What are some linear combinations of v and w?

- $v + w$
- $v - w$
- $2v + 0w$
- $2w$
- $-v$

Poll

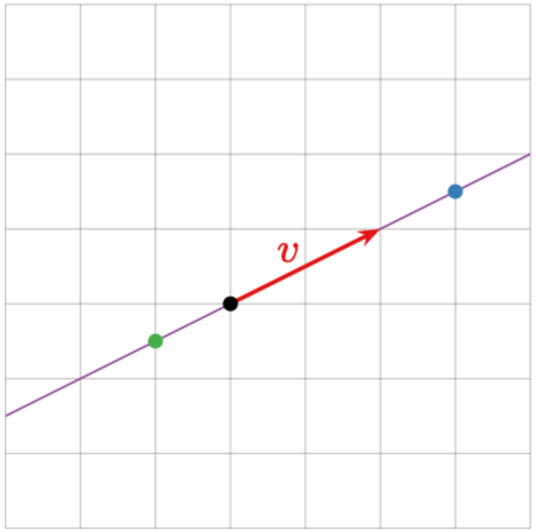
Poll

Is there any vector in $\mathbf{R}^2$ that is *not* a linear combination of v and w?



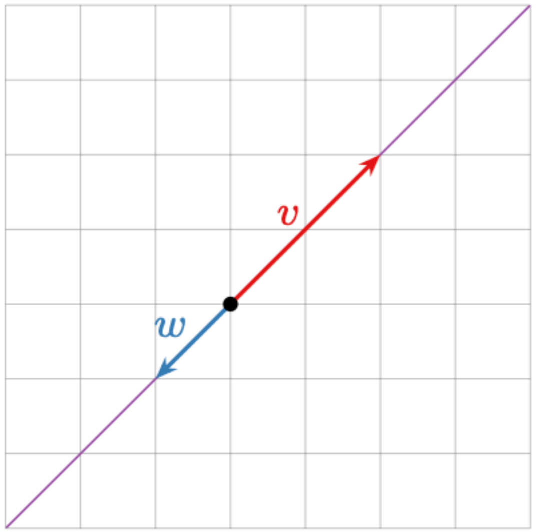
Examples

What are some linear combinations of $v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$?

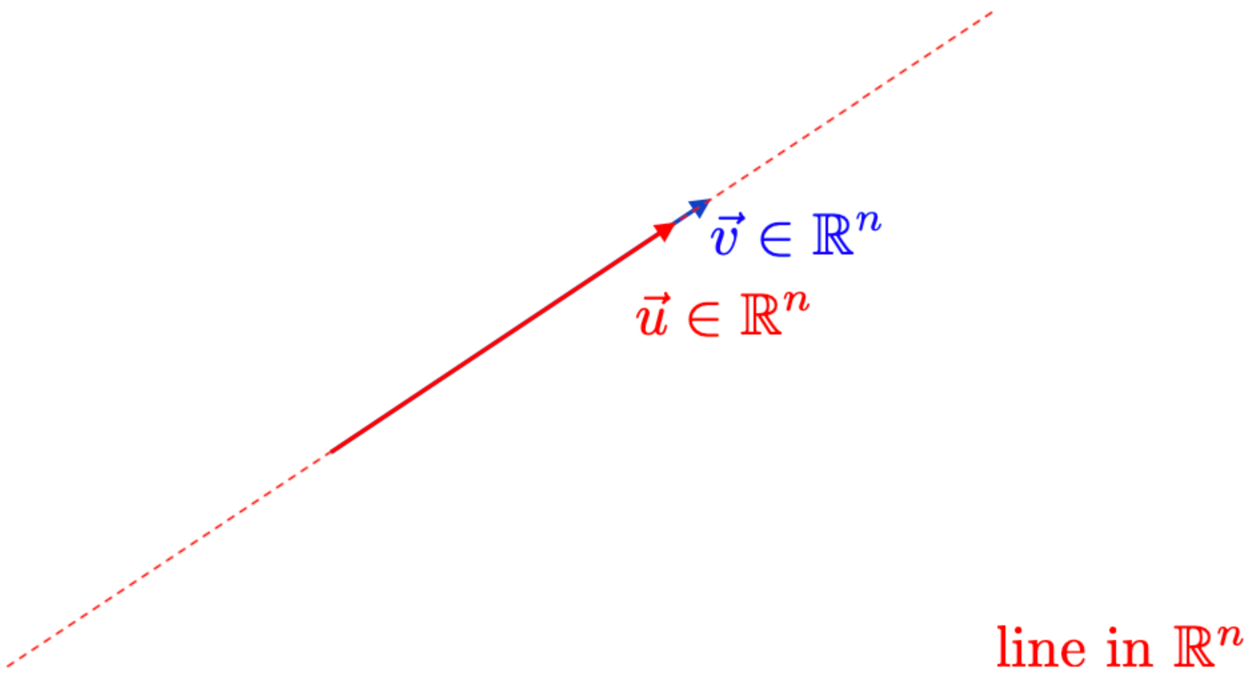


What are all linear combinations of

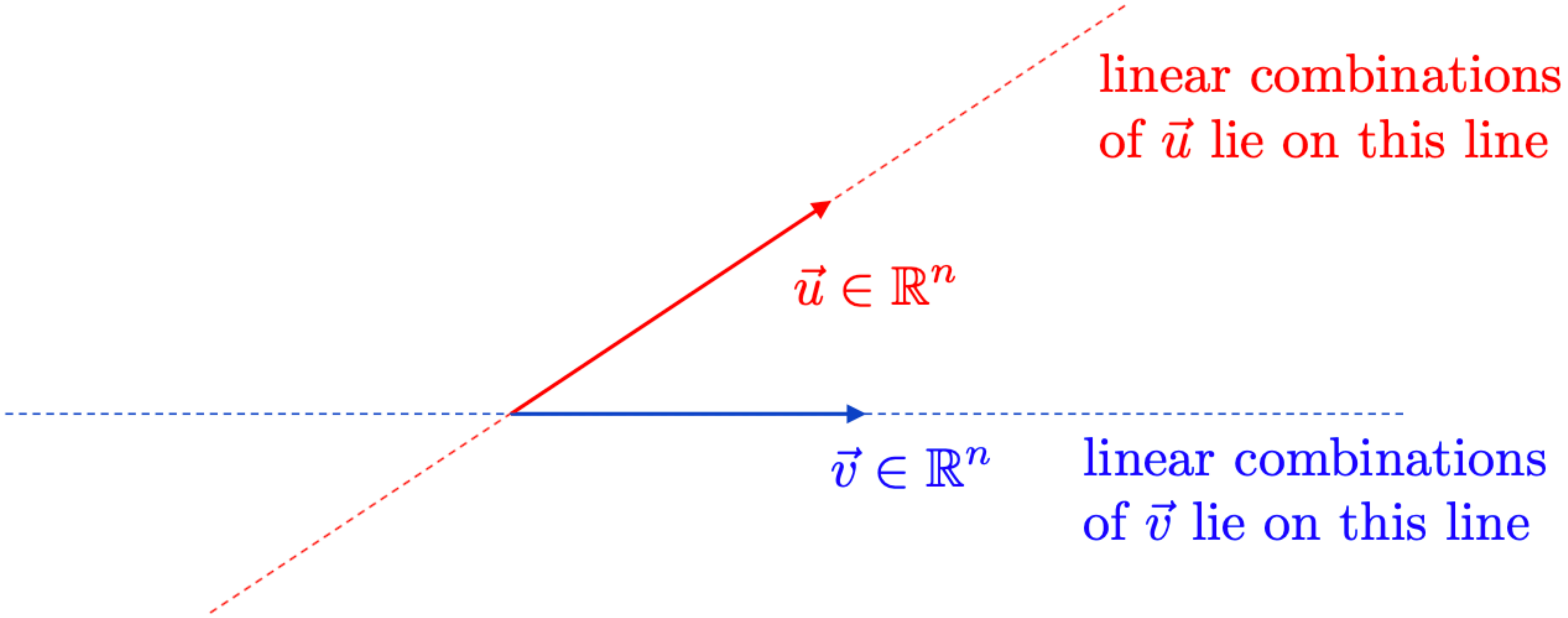
$$v = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} -1 \\ -1 \end{pmatrix}?$$



Geometric Interpretation of Linear Combinations



Geometric Interpretation of Linear Combinations



linear combinations of $\vec{u}$ and $\vec{v}$ lie on a plane in $\mathbb{R}^n$

Vector Equations

Question

Is $\begin{pmatrix} 8 \\ 16 \\ 3 \end{pmatrix}$ a linear combination of $\begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix}$?

1.2 – Lengths and Dot Products

Course notes adapted from *Introduction to Linear Algebra* by Strang (5th ed), and *Interactive Linear Algebra* by Margalit and Rabinoff.

Dot Product

We need a notion of *angle* between two vectors, and in particular, a notion of *orthogonality* (i.e. when two vectors are perpendicular). This is the purpose of the dot product.

Dot Product

We need a notion of *angle* between two vectors, and in particular, a notion of *orthogonality* (i.e. when two vectors are perpendicular). This is the purpose of the dot product.

Definition
The **dot product** of two vectors x, y in $\mathbf{R}^n$ is

$$x \cdot y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \stackrel{\text{def}}{=} x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Thinking of x, y as column vectors, this is the same as $x^T y$.

Example

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 32.$$

Dot Product

Many usual arithmetic rules hold, as long as you remember you can only dot two vectors together, and that the result is a scalar.

- $x \cdot y = y \cdot x$
- $(x + y) \cdot z = x \cdot z + y \cdot z$
- $(cx) \cdot y = c(x \cdot y)$

Dotting a vector with itself is special:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1^2 + x_2^2 + \cdots + x_n^2.$$

Hence:

- $x \cdot x \geq 0$
- $x \cdot x = 0$ if and only if $x = 0$.

Important: $x \cdot y = 0$ does *not* imply $x = 0$ or $y = 0$. For example, $\binom{1}{0} \cdot \binom{0}{1} = 0$.

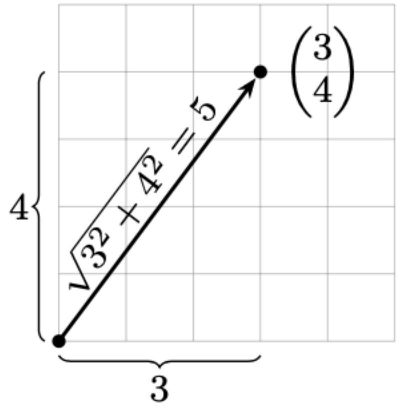
Dot Product and Length

Definition

The **length** or **norm** of a vector x in $\mathbf{R}^n$ is

$$\|x\| = \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Why is this a good definition? The Pythagorean theorem!



$$\left\|\begin{pmatrix}3\\4\end{pmatrix}\right\| = \sqrt{3^2 + 4^2} = 5$$

Fact

If x is a vector and c is a scalar, then $\|cx\| = |c| \cdot \|x\|$.

$$\left\|\begin{pmatrix}6\\8\end{pmatrix}\right\| = \left\|2\begin{pmatrix}3\\4\end{pmatrix}\right\| = 2\left\|\begin{pmatrix}3\\4\end{pmatrix}\right\| = 10$$

Dot Product and Distance

Definition

The **distance** between two points x, y in $\mathbf{R}^n$ is

$$\text{dist}(x, y) = \|y - x\|.$$

This is just the length of the vector from x to y.

Example

Let $x = (1, 2)$ and $y = (4, 4)$. Then

Dot Products

Definition

A **unit vector** is a vector v with length $\|v\| = 1$.

Example

The unit coordinate vectors are unit vectors:

$$\|e_1\| = \left\| \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\| = \sqrt{1^2 + 0^2 + 0^2} = 1$$

Definition

Let x be a nonzero vector in $\mathbf{R}^n$. The **unit vector in the direction of** x is the vector $\dfrac{x}{\|x\|}$.

This is in fact a unit vector:

scalar

$$\left\| \frac{x}{\|x\|} \right\| = \frac{1}{\|x\|}\|x\| = 1.$$

Dot Products

Example

What is the unit vector in the direction of $x = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$?

Orthogonality

Definition

Two vectors x, y are **orthogonal** or **perpendicular** if $x \cdot y = 0$.

Notation: $x \perp y$ means $x \cdot y = 0$.

Some Formulas

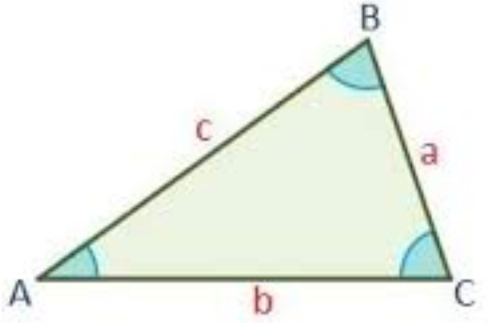
Cosine Formula/Alternate Dot Product Definition:

If u and v are nonzero vectors then

$$\frac{u \cdot v}{\|u\|\|v\|} = \cos\theta$$

The sign of the dot product tells us whether $\theta < \frac{\pi}{2}$ or $\theta > \frac{\pi}{2}$. Alternatively, this can be written as $u \cdot v = \|u\|\|v\| \cos\theta$ for a more general definition of the dot product.

Generalized Pythagorean theorem

- 

$$c^2 = a^2 + b^2 - 2ab \cdot \cos C$$

Some Formulas

Cosine Formula/Alternate Dot Product Definition:

If u and v are nonzero vectors then

$$\frac{u \cdot v}{\|u\|\|v\|} = \cos\theta$$

The sign of the dot product tells us whether $\theta < \frac{\pi}{2}$ or $\theta > \frac{\pi}{2}$. Alternatively, this can be written as $u \cdot v = \|u\|\|v\| \cos\theta$ for a more general definition of the dot product.

Schwarz Inequality

A consequence of the previous formula is that

$$|u \cdot v| \leq \|u\|\|v\|$$

Triangle Inequality

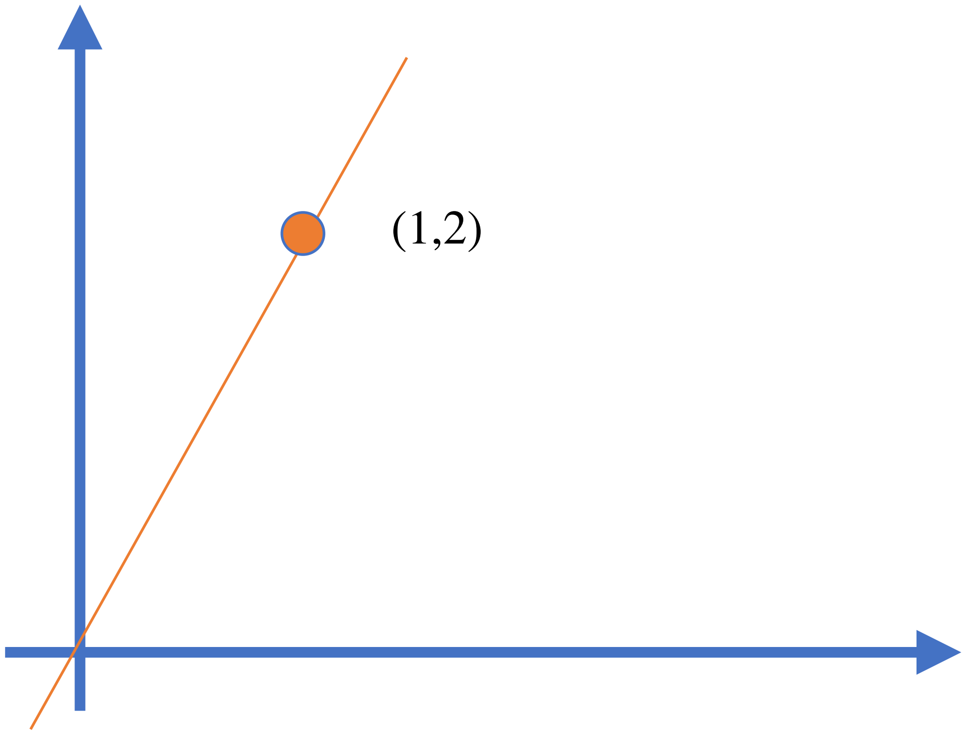
$$\|u + v\| \leq \|u\| + \|v\|$$

Motivation: Best fit of linear equation

Not Required

overdetermined linear system

$$2x = 2$$
$$x = 1$$



$$2x = 1$$
$$x = 1$$