SIT787: Mathematics for Artificial Intelligence
Topic 2: Linear Algebra
Part 1

Asef Nazari

School of Information Technology, Deakin University

Linear Algebra

- A good understanding of Linear Algebra
  - is essential for understanding and working with many machine learning algorithms,
  - especially deep learning algorithms
- The entities we deal with are
  - scalars
  - vectors
  - matrices
  - tensors

Scalars, Vectors, Matrices and Tensors

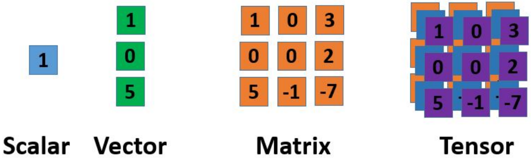
- Scalars: a single number $c \in \mathbb{R}$
- Vectors: an array of numbers
  - order is important, and each element of a vector has an index

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

- Matrices: a 2-D array of numbers
  - each element is identifiedby two indices

$$A_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m\times n}$$

- Tensors: an array with more than two axes
  - A_{ijk}
  - an RGB color image has three axes



Operations

- Transpose
- Addition (subtraction)
- Multiplying by a scalar
- Products
  - vector product
  - matrix product

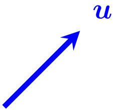
Data tables

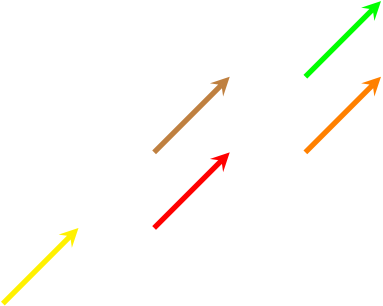
	A	B	C	D	E	F	G	H	I
1			Database clients of Jolly Day						
2	№	Customer	Type	Country	City	Contract Number	Date	Limitation years	Contact Manager
3	1	Intersection	com.network	USA	New York	2314589	12.12.2012	2	Aaron
4	2	Magnet	com.network	USA	New York	2432656	27.08.2014	3	Alex
5	3	Perspective korp.	warehouse	Belarus	Minsk	2456983	31.12.2014	2	Ashley
6	4	Driveway	enterprise	USA	New York	2408570	24.04.2014	5	Aaron
7	5	near	enterprise	USA	Los Angeles	2481553	06.05.2015	2	Ashley
8	6	Nori	warehouse	Japan	Tokyo	2506369	09.09.2015	2	Blake
9	7	Nevsky comp.	com.network	Russia	Moscow	2337735	15.04.2013	1	Caroline
10	8	Perspective korp.	enterprise	Belarus	Minsk	2361112	17.08.2013	2	Daniel
11	9	in touch	warehouse	USA	San Francisco	2384723	20.12.2013	2	Alex
12	10	Nardis	com.network	Japan	Tokyo	2531433	14.01.2016	3	Blake

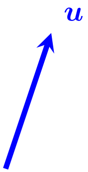
- 2D array
- each case is a vector
- each variable is a vector

Vectors

- a quantity with length and direction

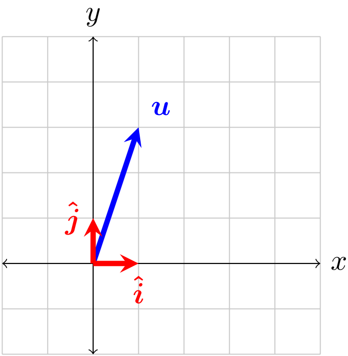






$$\boldsymbol{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \hat{\boldsymbol{i}} + 3\hat{\boldsymbol{j}}$$

$$\hat{\boldsymbol{i}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \hat{\boldsymbol{j}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



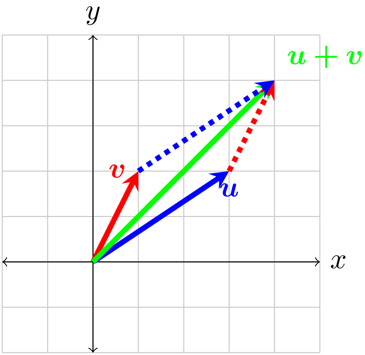
Vector Addition

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \text{ and } \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}$$

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \text{ and } \boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

$$\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix}$$



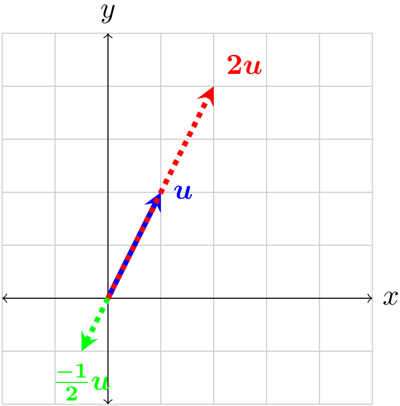
- $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$
- $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n$$

For $c \in \mathbb{R}$

$$c\boldsymbol{u} = \begin{bmatrix} cu_1 \\ \vdots \\ cu_n \end{bmatrix} \in \mathbb{R}^n$$

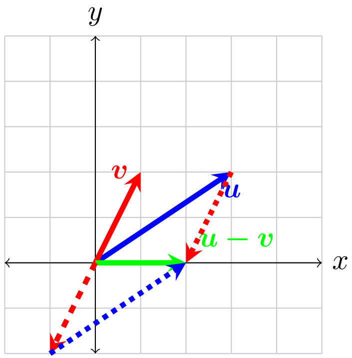
- $c\boldsymbol{u}$ depends on $\boldsymbol{u}$, so they are linearly dependent, and they are not linearly independent!
- $c\boldsymbol{u}$ and $\boldsymbol{u}$ are parallel



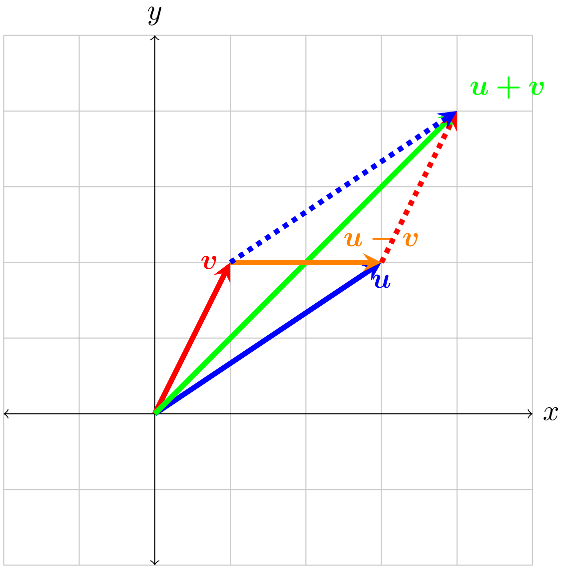
- Vectors $\boldsymbol{u} \in \mathbb{R}^n$ and $\boldsymbol{v} \in \mathbb{R}^n$ are given
- To find $\boldsymbol{u} - \boldsymbol{v}$
- find $-\boldsymbol{v}$ first
- then $\boldsymbol{u} + (-\boldsymbol{v})$

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n \text{ and } \boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$$

$$\boldsymbol{u} - \boldsymbol{v} = \begin{bmatrix} u_1 - v_1 \\ \vdots \\ u_n - v_n \end{bmatrix} \in \mathbb{R}^n$$

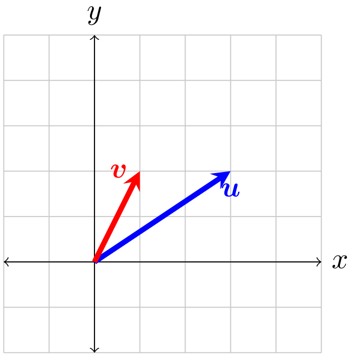


Vector Addition and Subtraction: Paralleogram



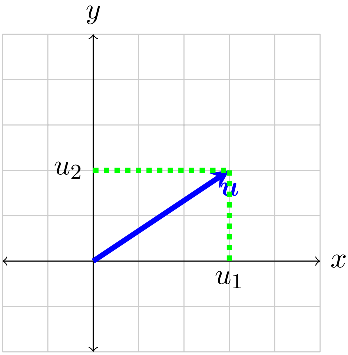
Independent vectors

- There is no way one can express $\boldsymbol{u}$ as a scalar product of $\boldsymbol{v}$, or $\boldsymbol{v}$ as a scalar product of $\boldsymbol{u}$.
- They are independent vectors.
- But $\boldsymbol{u}$ and $c\boldsymbol{u}$ are dependent vectors. They are parallel.



- $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in \mathbb{R}^2$
- $\boldsymbol{u} = u_1 \hat{\boldsymbol{i}} + u_2 \hat{\boldsymbol{j}}$
- Pythagoras' theorem:
  length of $\boldsymbol{u} = \sqrt{u_1^2 + u_2^2}$
- $||\boldsymbol{u}|| = \sqrt{u_1^2 + u_2^2}$
- If $\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n$, then

$$||\boldsymbol{u}|| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$$

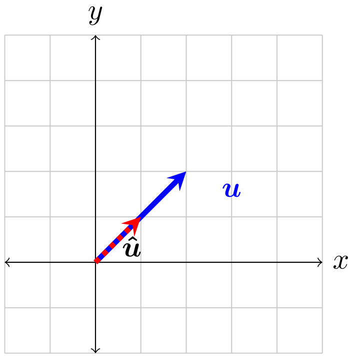


$$||\boldsymbol{u}|| = \sqrt{\sum_{i=1}^{n} u_i^2}$$

The Direction of a vector: Unit vectors

- A vector has a length and direction
- for $\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$,

$$||\boldsymbol{u}|| = \sqrt{\sum_{i=1}^{n} u_i^2}$$



- The unit vector in the doirection of $\boldsymbol{u}$ is $\hat{\boldsymbol{u}}$

$$\hat{\boldsymbol{u}} = \left(\frac{1}{\text{length } \boldsymbol{u}}\right)\boldsymbol{u} = \frac{1}{||\boldsymbol{u}||}\boldsymbol{u}$$

$$\boldsymbol{u} = ||\boldsymbol{u}||\hat{\boldsymbol{u}}$$

Dot product or Inner Product

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \text{ and } \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2$$

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \text{ and } \boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

- $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$
- $\boldsymbol{u} \cdot \boldsymbol{u} = \sum_{i=1}^{n} u_i u_i = \sum_{i=1}^{n} u_i^2 = ||\boldsymbol{u}||^2 \implies ||\boldsymbol{u}|| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}$
- $\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w}$

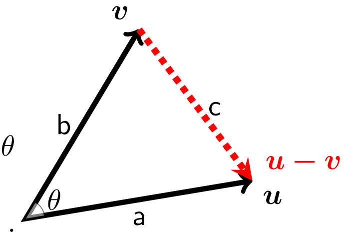
Some Questions

- $\boldsymbol{u} \cdot (\boldsymbol{v} \cdot \boldsymbol{w}) = (\boldsymbol{u} \cdot \boldsymbol{v}) \cdot \boldsymbol{w}$?
- For $c_1, c_2 \in \mathbb{R}$, $(c_1 \boldsymbol{u}) \cdot (c_2 \boldsymbol{v}) = c_1 c_2 \boldsymbol{u} \cdot \boldsymbol{v}$?
- For $c \in \mathbb{R}$, $c(\boldsymbol{u} \cdot \boldsymbol{v}) = (c\boldsymbol{u}) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot (c\boldsymbol{v})$

Dot Product: Another Formula

- Cosine rule in triangles
  $c^2 = a^2 + b^2 - 2ab\cos\theta$

  $||\boldsymbol{u}-\boldsymbol{v}||^2 = ||\boldsymbol{u}||^2+||\boldsymbol{v}||^2-2||\boldsymbol{u}||||\boldsymbol{v}||\cos\theta$



- We know that $\boldsymbol{u}\cdot\boldsymbol{u} = ||\boldsymbol{u}||^2$
  - $||\boldsymbol{u}-\boldsymbol{v}||^2 = (\boldsymbol{u}-\boldsymbol{v})\cdot(\boldsymbol{u}-\boldsymbol{v}) =$
    $(\boldsymbol{u}\cdot\boldsymbol{u}) + (\boldsymbol{v}\cdot\boldsymbol{v}) - 2(\boldsymbol{u}\cdot\boldsymbol{v}) = ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2 - 2(\boldsymbol{u}\cdot\boldsymbol{v})$
  - $||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2 - 2(\boldsymbol{u}\cdot\boldsymbol{v}) = ||\boldsymbol{u}||^2 + ||\boldsymbol{v}||^2 - 2||\boldsymbol{u}||||\boldsymbol{v}||\cos\theta$
    Then

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

Dot (inner) Product

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \text{ and } \boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

- $\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n = \sum_{i=1}^{n} u_i v_i$
- $\boldsymbol{u} \cdot \boldsymbol{v} = ||\boldsymbol{u}||||\boldsymbol{v}|| \cos\theta$
- Cosine between two vectors is a measure of their similarity

$$\cos\theta = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{||\boldsymbol{u}||||\boldsymbol{v}||} = \frac{\sum_{i=1}^{n} u_i v_i}{\left(\sqrt{\sum_{i=1}^{n} u_i^2}\right)\left(\sqrt{\sum_{i=1}^{n} v_i^2}\right)}$$

The benefits of dot product

- Finding vector length $||\boldsymbol{u}|| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}$
  - is used to find unit vectors
- finding angle between vectors

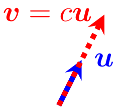
$$\cos\theta = \frac{\boldsymbol{u} \cdot \boldsymbol{u}}{||\boldsymbol{u}||||\boldsymbol{v}||}$$

  - To compare vectors: How similar they are!
  - The less the angle between thwm the more similar they are!
  - If two vectors are perpendicular (orthogonal), the angle between them is $\theta = \frac{\pi}{2}$, and $\cos\frac{\pi}{2} = 0$, then

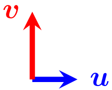
$$\boldsymbol{u} \cdot \boldsymbol{v} = ||\boldsymbol{u}||||\boldsymbol{v}||\underbrace{\cos\theta}_{=0} = 0$$

  - Orthogonal vectors $\boldsymbol{u} \perp \boldsymbol{v} \Leftrightarrow \boldsymbol{u} \cdot \boldsymbol{v} = 0$
  - Orthonormal vectors

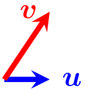
The impact of θ



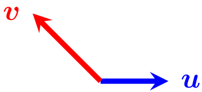
$$\theta = 0$$
$$\boldsymbol{u} \cdot \boldsymbol{v} = ||\boldsymbol{u}||||\boldsymbol{v}||$$
$$c > 0$$



$$\theta = \frac{\pi}{2}$$
$$\boldsymbol{u} \cdot \boldsymbol{v} = 0$$



$$\theta < \frac{\pi}{2}$$
$$\boldsymbol{u} \cdot \boldsymbol{v} > 0$$

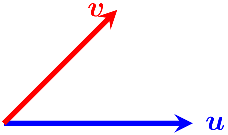


$$\theta > \frac{\pi}{2}$$
$$\boldsymbol{u} \cdot \boldsymbol{v} < 0$$

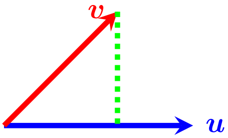
$$\theta = \pi$$
$$\boldsymbol{u} \cdot \boldsymbol{v} = -||\boldsymbol{u}||||\boldsymbol{v}||$$
$$c < 0$$

Projection

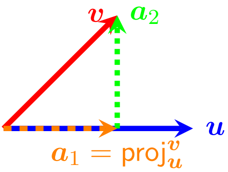
- Find the projection of $\boldsymbol{v}$ over $\boldsymbol{u}$



- draw a vertical line from the end of $\boldsymbol{v}$ to $\boldsymbol{u}$

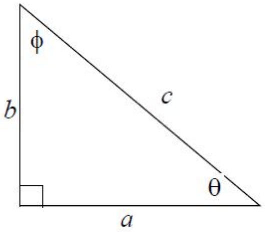


- aim: find $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$



Projection

- In a right angle triangle
  - $\cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}} = \frac{a}{c}$
  - $\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}} = \frac{b}{c}$

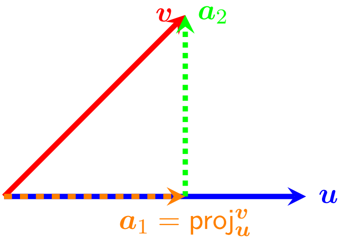


Projection

- $\boldsymbol{a}_1$
  - Length: $\cos\theta = \frac{||\boldsymbol{a}_1||}{||\boldsymbol{v}||} \rightarrow ||\boldsymbol{a}_1|| = ||\boldsymbol{v}|| \cos\theta$
  - Direction: same as $\boldsymbol{u}$, which is $\frac{\boldsymbol{u}}{||\boldsymbol{u}||}$
  - as $\boldsymbol{u} \cdot \boldsymbol{v} = ||\boldsymbol{u}|| ||\boldsymbol{v}|| \cos\theta$,

$$||\boldsymbol{a}_1|| = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{||\boldsymbol{u}||}$$

$$\boldsymbol{a}_1 = \text{(length)(direction)} = \left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{||\boldsymbol{u}||}\right)\left(\frac{\boldsymbol{u}}{||\boldsymbol{u}||}\right) = \left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\boldsymbol{u} \cdot \boldsymbol{u}}\right) \boldsymbol{u}$$



- $\boldsymbol{a}_2 = \boldsymbol{v} - \boldsymbol{a}_1 = \boldsymbol{v} - \left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\boldsymbol{u} \cdot \boldsymbol{u}}\right) \boldsymbol{u}$

Linear Combination of vectors

- consider $\{\boldsymbol{v}\}$. For different values of $c \in \mathbb{R}$, $c\boldsymbol{v}$ produces a line.
- consider $\{\boldsymbol{u}, \boldsymbol{v}\}$. For $c_1 \in \mathbb{R}$ and $c_2 \in \mathbb{R}$, a linear combination of those two vectors is $c_1\boldsymbol{u} + c_2\boldsymbol{v}$
- Note that he linear combination of any number of vectors is a vector.
- In general, for $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$, a linear combination of vectors of this set is

$$c_1\boldsymbol{v}_1 + \ldots + c_k\boldsymbol{v}_k$$

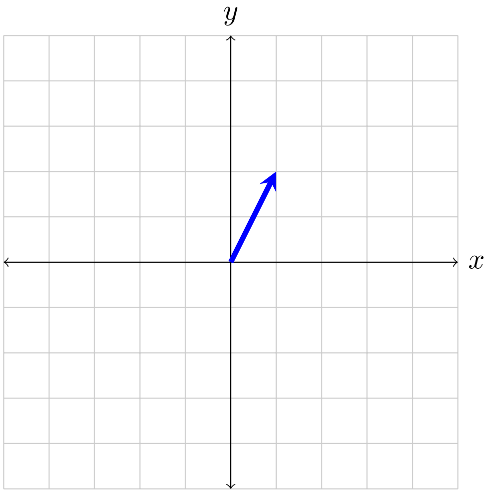
- Consider $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$. The span of S is the set of all linear combinations of its vectors

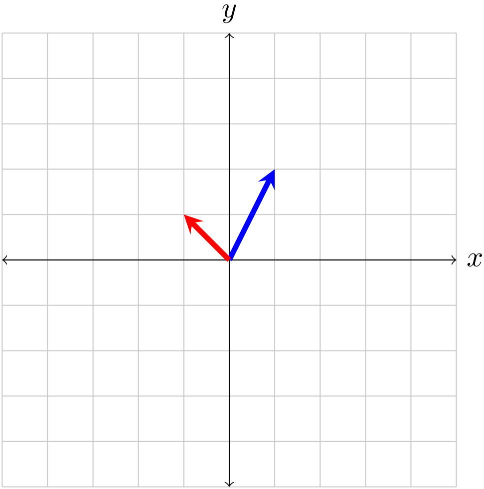
$$\text{span}(S) = \left\{ \sum_{i=1}^{k} c_i\boldsymbol{v}_i | \text{ for all } c_i \in \mathbb{R} \right\}$$

Linear Combination of vectors: Examples

- For $S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ find span(S).
- For $S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$ find span(S).
- For $S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ find span(S).

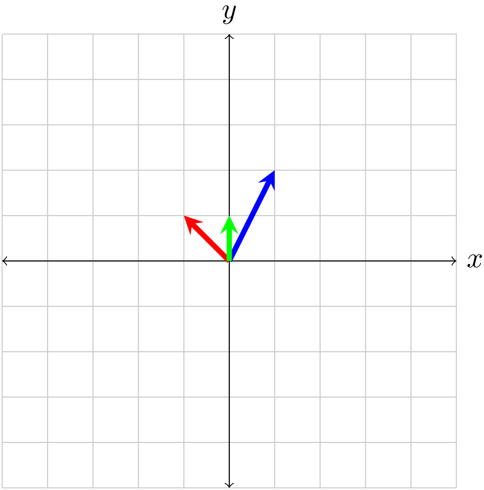
Linear Combination of vectors: Examples





y
x

Linear Combination of vectors: Examples



Vectors

- A n-vectors: a list (tuple) of n numbers.
- the set of all n-vectors is $\mathbb{R}^n$

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$$

- 2-vectors,

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2$$

- 1-vectors,

$$\boldsymbol{v} = \begin{bmatrix} v_1 \end{bmatrix} \in \mathbb{R}$$

Vectors Transpose

- A column vector and a row vector

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$$

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

- $(\boldsymbol{v}^T)^T = \boldsymbol{v}$
- Zero vector

$$\mathbf{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n$$

- Vector of ones

$$\mathbf{1}_n^T = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \in \mathbb{R}^n$$

Vector spaces

- A set of elements (which are called vectors), with two properties
  - The elements can be added to each other

$$\mathbf{x}, \mathbf{y} \in V, \rightarrow \mathbf{x} + \mathbf{y} \in V$$

  - an element can be multiplied by a scaler

$$\mathbf{x} \in V \text{ and } c \in \mathbb{R}, \rightarrow c\mathbf{x} \in V$$

- sometimes it is called linear space
- Every vector space has a null element $\mathbf{0} \in V$,

$$\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$$

- All elements has an additive inverse $-\mathbf{x}$,

$$\mathbf{x} + -\mathbf{x} = \mathbf{0}$$

- Addition:

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

- Scaling

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \text{ and } c \in \mathbb{R} \rightarrow c\boldsymbol{v} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}$$

- Linear combination

$$c\boldsymbol{v} + d\boldsymbol{w} = \begin{bmatrix} cv_1 + dw_1 \\ cv_2 + dw_2 \\ \vdots \\ cv_n + dw_n \end{bmatrix}$$

Vector space examples

- Euclidean space
  - 1-Dimensional space $\mathbb{R}$
  - 2-Dimensional space $\mathbb{R}^2$
  - 3-Dimensional space $\mathbb{R}^3$
  - n-Dimensional space $\mathbb{R}^n$

Properties

- Addition: $\boldsymbol{v}, \boldsymbol{w} \in V \rightarrow \boldsymbol{v}+\boldsymbol{w} \in V$
- Commutativity: $\boldsymbol{v}+\boldsymbol{w}=\boldsymbol{w}+\boldsymbol{v}$
- Zero vector: $\mathbf{0}$
- Identity element: $\boldsymbol{v}+\mathbf{0}=\mathbf{0}+\boldsymbol{v}=\boldsymbol{v}$
- Inverses: $\boldsymbol{v}+(-\boldsymbol{v})=(-\boldsymbol{v})+\boldsymbol{v}=\mathbf{0}$
- Associativity: $\boldsymbol{v}+(\boldsymbol{w}+\boldsymbol{z})=(\boldsymbol{v}+\boldsymbol{w})+\boldsymbol{z}$

Vector space examples

- $V = \{$all real polynomials of degree 3 or less$\}$
- $ax^3 + bx^2 + cx + d, \quad a, b, c, d \in \mathbb{R}$
- addition
- scaling
- zero vector
- inverse

Vector space examples

- $V = \{f(x) | f(x) \text{ is continuous on } \mathbb{R}\}$
- addition
- scaling
- zero vector
- inverse

Subspace of a vector space

- Vector spaces can contain other vector spaces.
- $L \subset V$
  - $\mathbf{0} \in L$
  - $\mathbf{x}, \mathbf{y} \in L, \rightarrow \mathbf{x} + \mathbf{y} \in L$
  - $\mathbf{x} \in L$ and $c \in \mathbb{R}, \rightarrow c\mathbf{x} \in L$
- $V \subset V$
- $\{\mathbf{0}\} \subset V$ trivial subspace
- a line passing through the origin is a subspace of Euclidean space

Linear combinations

- Consider a vector space V
- $\mathbf{x}_1, \ldots, \mathbf{x}_n \in V$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$
- A linear combination

$$\mathbf{y} = \alpha_1 \mathbf{x}_1 + \ldots + \alpha_n \mathbf{x}_n \in V$$

Span

- Consider $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$
- The set of all linear combinations of members of X,

$$\text{span}(X) = \text{span}(\mathbf{x}_1, \ldots, \mathbf{x}_n) =$$

$$\{\alpha_1 \mathbf{x}_1 + \ldots + \alpha_n \mathbf{x}_n | \alpha_1, \ldots, \alpha_n \in \mathbb{R}\}$$

- a linear span of X
- Span of a set of vectors: a set obtained by a linear combination of those vectors

Linear dependence

- $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$
- If $\alpha_1\mathbf{x}_1 + \ldots + \alpha_n\mathbf{x}_n = \mathbf{0}$ implies that all the scalars are zero, we say vectors in X are independent.
- $\alpha_1\mathbf{x}_1 + \ldots + \alpha_n\mathbf{x}_n = \mathbf{0}$ and at least one of the scalars is not zero, we say vectors in X are dependent.

Basis for vector space

- $B = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset V$ is a basis for V if and only if
  - B is linearly independent
  - $V = \text{span}(B)$
- Example $V = \mathbb{R}^n$

$$B = \left\{ \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$$

- a basis is not unique.
- If a vector space is spanned by a finite number of vectors, it is said to be finite-dimensional. Otherwise it is in infinite-dimensional.

Dimension

- The number of vectors in a basis for a finite-dimensional vector space V is called the dimension of V and denoted $\dim(V)$.
- $\dim(V) = |B|$
- $\dim(\mathbb{R}^n) = n$
- $\dim(\mathbb{R}^2) = 2$
- $\dim(\{\mathbf{0}\}) = 0$

Normed spaces

- Norms generalise the notion of length from Euclidean space
- A norm on a real vector space V is a function $||.|| : V \to \mathbb{R}$ that satisfies
  for all $\boldsymbol{x}, \boldsymbol{y} \in V$ and $\alpha \in \mathbb{R}$
  - $||\boldsymbol{x}|| \geq 0$ with equality if and only if $\boldsymbol{x} = \mathbf{0}$
  - $||\alpha \boldsymbol{x}|| = |\alpha| ||\boldsymbol{x}||$
  - $||\boldsymbol{x} + \boldsymbol{y}|| \leq ||\boldsymbol{x}|| + ||\boldsymbol{y}||$ (the triangle inequality)
- A vector space provided with a norm is called a normed vector space, or simply a normed space.
- Any norm on V induces a distance metric on V

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

Frequent norms on $\mathbb{R}^n$

For $\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$

- 1-norm

$$||\boldsymbol{x}||_1 = \sum_{i=1}^{n} |x_i|$$

- 2-norm (Euclidean norm)

$$||\boldsymbol{x}||_2 = \sqrt{\sum_{i=1}^{n} x_i^2} = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}$$

- p-norm

$$||\boldsymbol{x}||_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \qquad p \geq 1$$

- ∞-norm or max norm

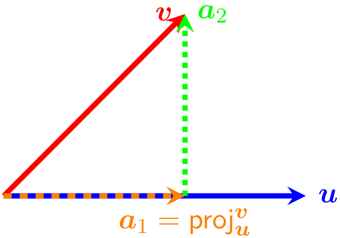
$$||\boldsymbol{x}||_\infty = \max_{i=1,\ldots,n} |x_i|$$

- Let's see $||\boldsymbol{v}||_p = 1$ for $\boldsymbol{v} = \begin{bmatrix} v_1, v_2 \end{bmatrix}^T \in \mathbb{R}^2$
  - $p = 2$, $v_1^2 + v_2^2 = 1$
  - $p = 1$, $|v_1| + |v_2| = 1$
  - $p = \infty$, $\max\{|v_1|, |v_2|\} = 1$
- Note that $|\boldsymbol{x}| = ||\boldsymbol{x}||_2$ and generally lenth of a vector is shown as $||\boldsymbol{x}||$.
- Cauchy-Schwarz inequality

$$\text{for } \boldsymbol{x}, \boldsymbol{y} \in V, \quad |\boldsymbol{x} \cdot \boldsymbol{y}| \leq ||\boldsymbol{x}||||\boldsymbol{y}||$$

Gram-Schmidt Process

- You remember from the projection that
  - For given vectors $\boldsymbol{u}$ and $\boldsymbol{v}$
  - we can find the projection of $\boldsymbol{v}$ over $\boldsymbol{u}$
  - this gives us two vectors
    - $\boldsymbol{a}_1 = \text{proj}_{\boldsymbol{u}}^{\boldsymbol{v}} = \left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\boldsymbol{u}\cdot\boldsymbol{u}}\right)\boldsymbol{u}$
    - $\boldsymbol{a}_2 = \boldsymbol{v} - \boldsymbol{a}_1 = \boldsymbol{v} - \left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\boldsymbol{u}\cdot\boldsymbol{u}}\right)\boldsymbol{u}$
    - $\boldsymbol{a}_1$ is the component of $\boldsymbol{v}$ in the direction of $\boldsymbol{u}$
    - $\boldsymbol{a}_2$ is the component of $\boldsymbol{v}$ in the direction operpendicular to $\boldsymbol{u}$
    - $\boldsymbol{a}_1 \perp \boldsymbol{a}_2$ or $\boldsymbol{a}_1 \cdot \boldsymbol{a}_2 = 0$



Example

- Vectors $\boldsymbol{v} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ are linearly independent
- Two independent vectors in $\mathbb{R}^2$ form a basis.
- but not orthogonal.
- But $\boldsymbol{v}$ and $\boldsymbol{a}_2 = \boldsymbol{v} - \text{proj}_{\boldsymbol{u}}^{\boldsymbol{v}} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ are independent and orthogonal.
- Therefore, $\{\boldsymbol{v}, \boldsymbol{a}_2\}$ form an orthogonal basis for $\mathbb{R}^2$.

Gram-Schmidt Process

- Vectors in $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \subset V$ are called mutually orthogonal when any two different members are orthogonal

$$\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0 \text{ for } i \neq j$$

- **Theorem:** If the vectors in a set are mutually orthogonal and nonzero then that set is linearly independent.
- Using Gram-Schmidt Process we orthognolise this set.
- we make $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ so that they are mutually orthogonal
  1. $\boldsymbol{u}_1 = \boldsymbol{v}_1$
  2. $\boldsymbol{u}_2 = \boldsymbol{v}_2 - \text{proj}_{\boldsymbol{u}_1}^{\boldsymbol{v}_2}$
  3. $\boldsymbol{u}_3 = \boldsymbol{v}_3 - \text{proj}_{\boldsymbol{u}_1}^{\boldsymbol{v}_3} - \text{proj}_{\boldsymbol{u}_2}^{\boldsymbol{v}_3}$
  4. $\ldots$
  5. $\boldsymbol{u}_k = \boldsymbol{v}_k - \text{proj}_{\boldsymbol{u}_1}^{\boldsymbol{v}_k} - \ldots - \text{proj}_{\boldsymbol{u}_{k-1}}^{\boldsymbol{v}_k}$
  6. Finally, normalise each vector $\boldsymbol{u}_k$ by dividing it by its length to get an orthonormal set.