Mathematics for AI

Lecture 2

Vectors, Vector Space, Span, Basis, Coordinates

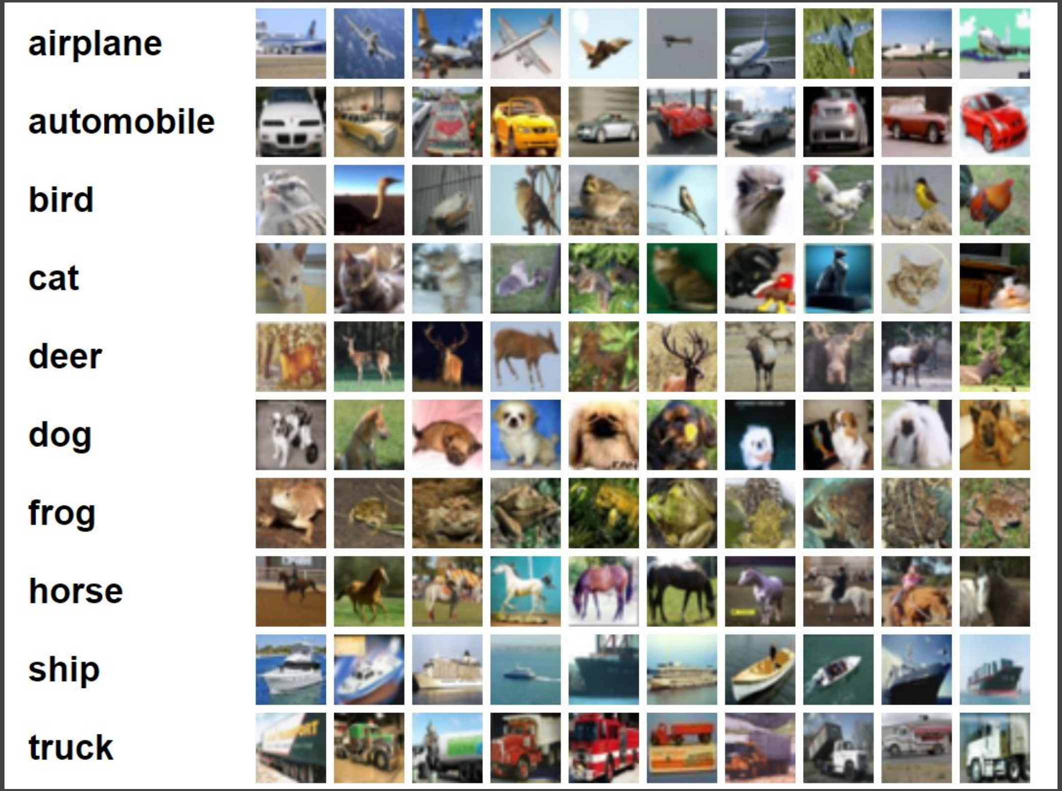
Machine Learning

Learning from data

K. N. Toosi
University of Technology

https://www.analyticsvidhya.com/blog/2018/03/comprehensive-collection-deep-learning-datasets/

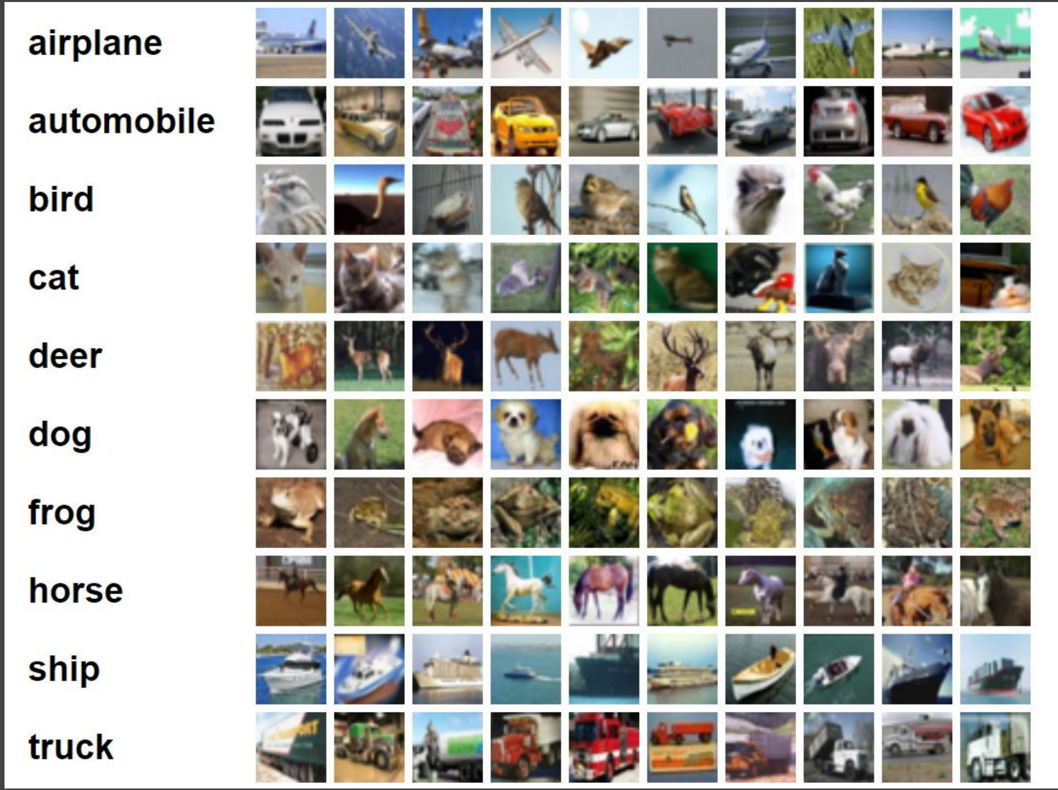
Supervised Learning



http://seansoleyman.com/effect-of-dataset-size-on-image-classification-accuracy/

Supervised Learning

K. N. Toosi
University of Technology



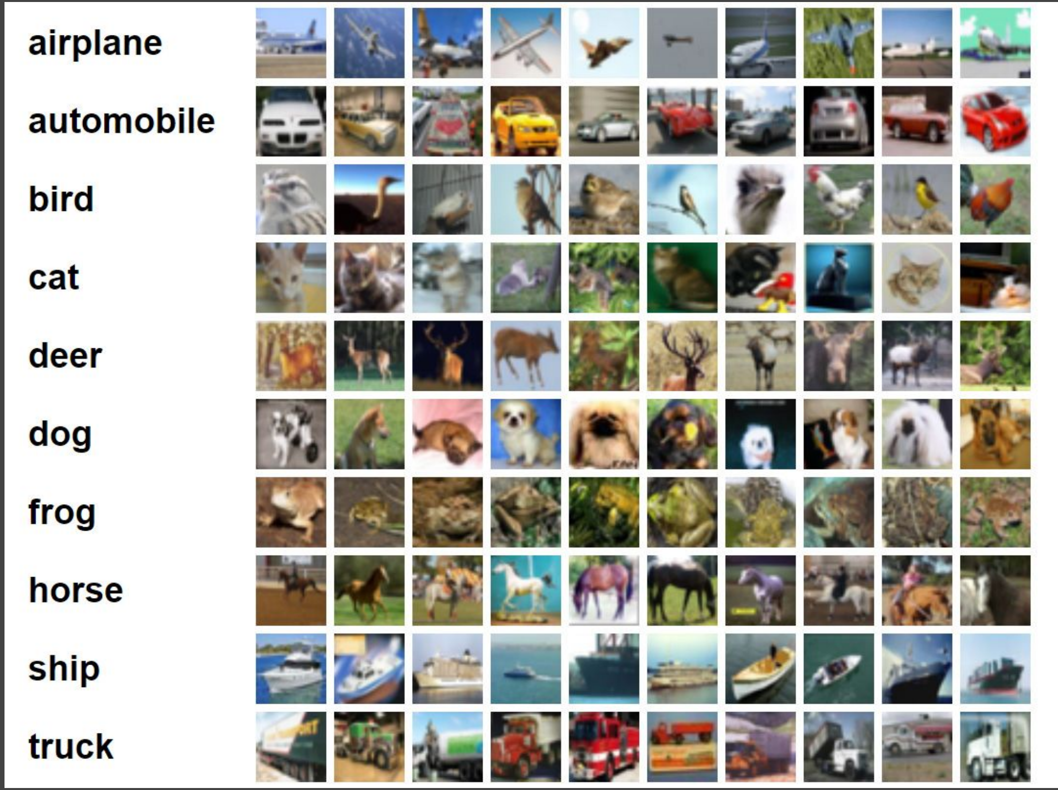
http://seansoleyman.com/effect-of-dataset-size-on-image-classification-accuracy/

Training data:

X_1, y_1
X_2, y_2
X_3, y_3
$\vdots$
X_n, y_n

Supervised Learning

K. N. Toosi
University of Technology

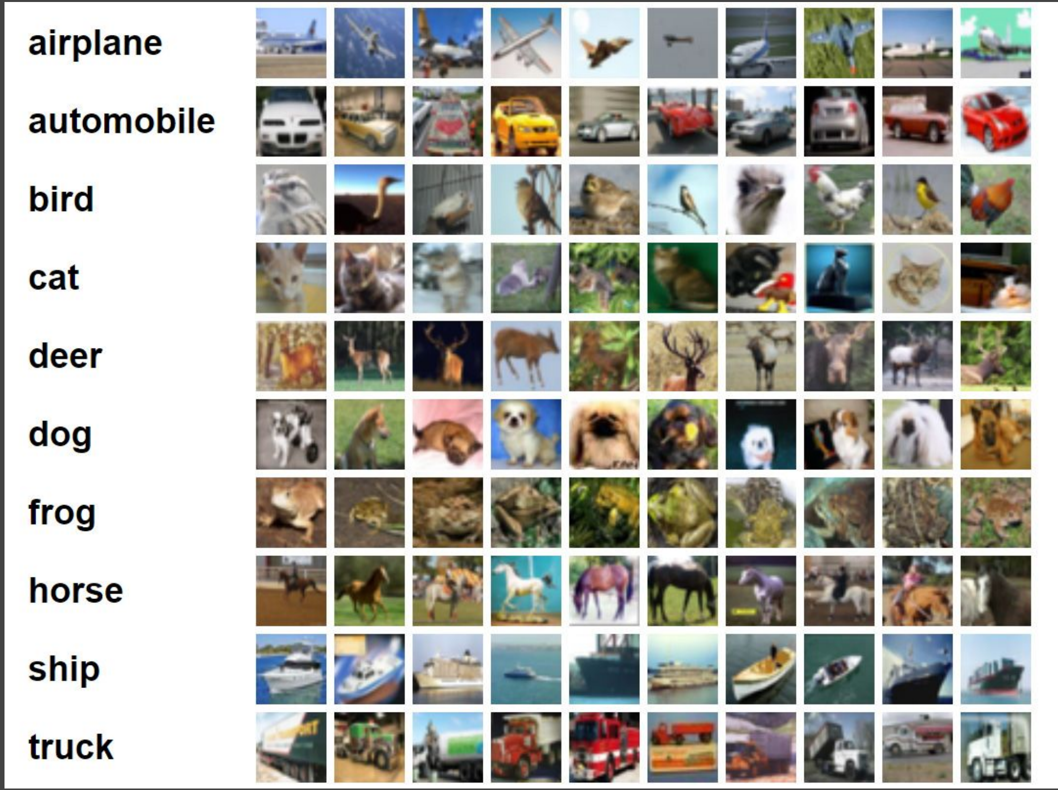


http://seansoleyman.com/effect-of-dataset-size-on-image-classification-accuracy/

Training data:

	Apple
	Apple
	Orange
⋮	
	Orange

Supervised Learning



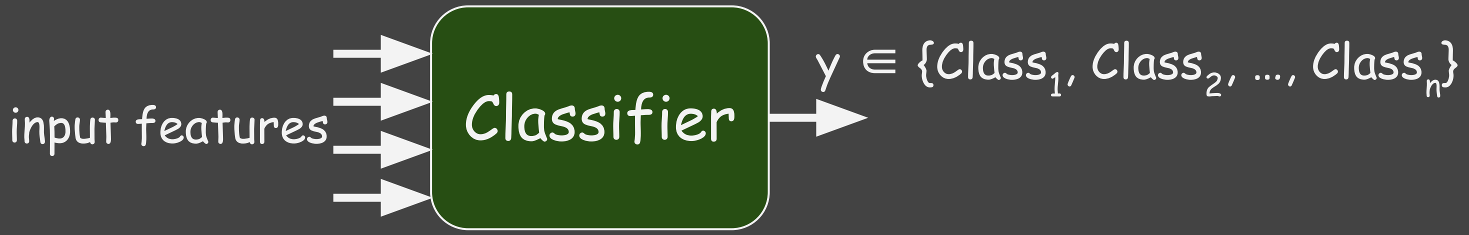
http://seansoleyman.com/effect-of-dataset-size-on-image-classification-accuracy/

Training data:

	0
	0
	1
⋮	
	1

Supervised Learning

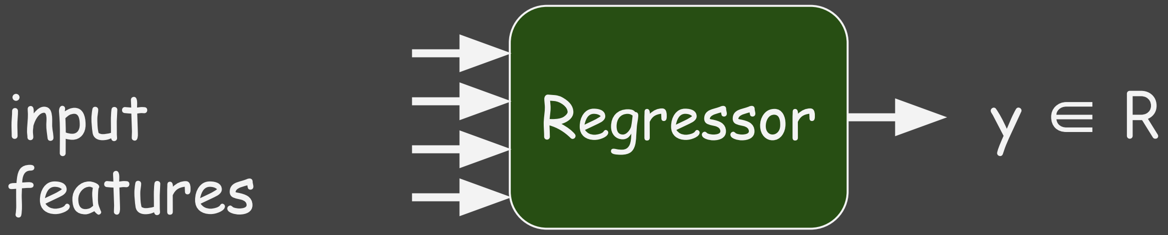
Classification



Classification

Classification

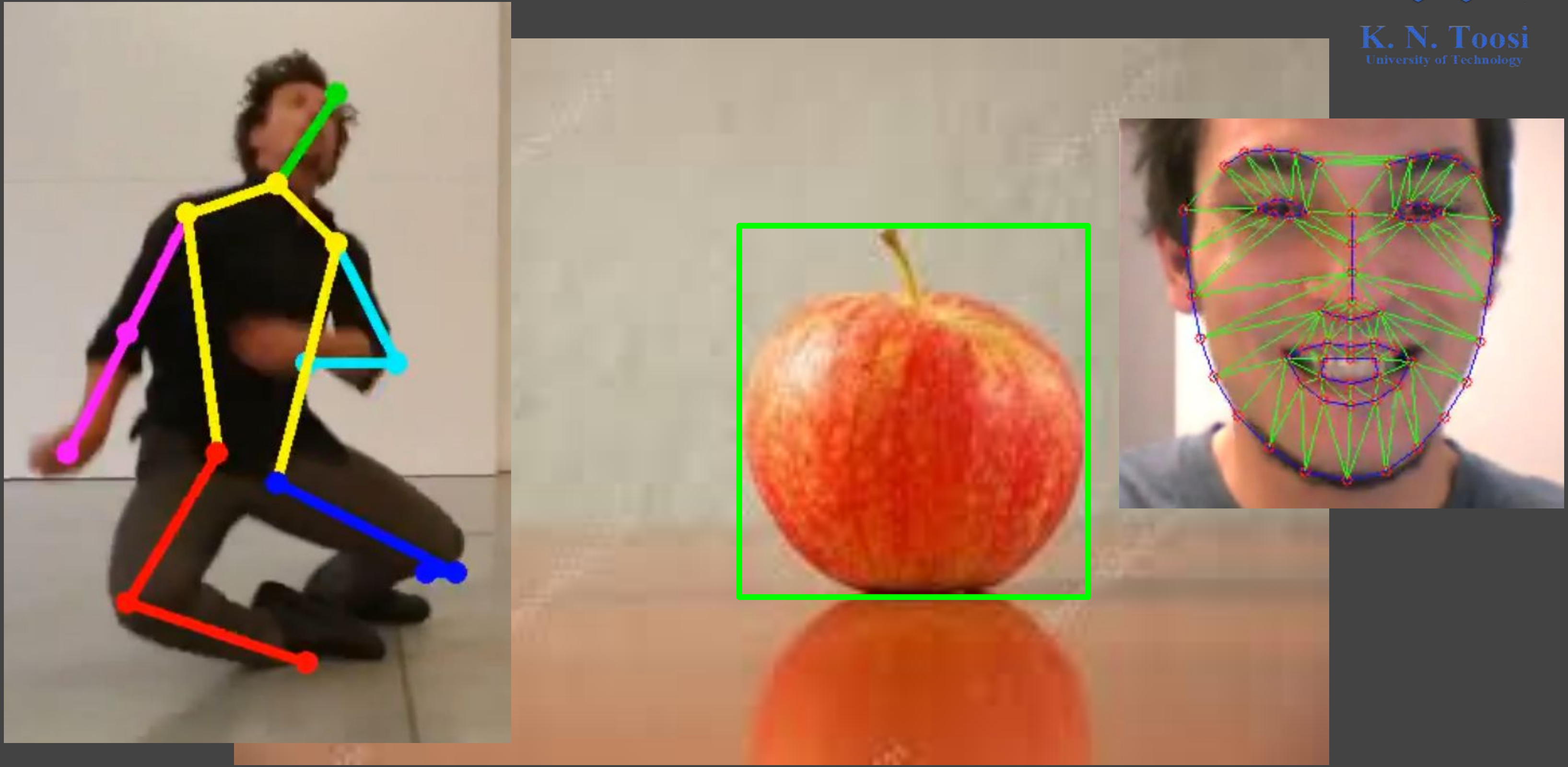
Regression



Regression

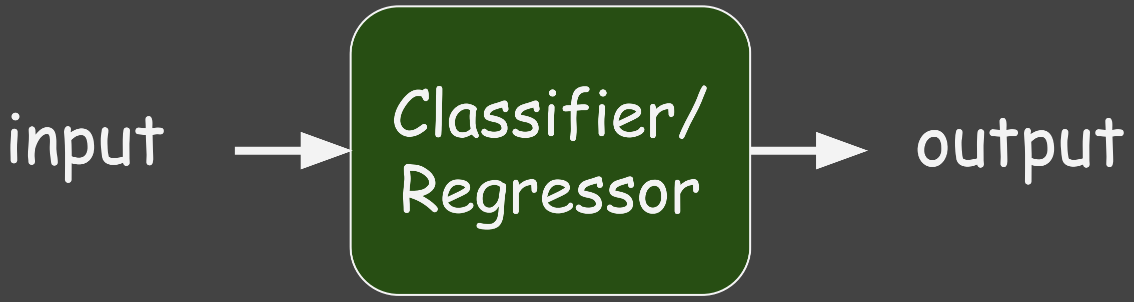
Regression

K. N. Toosi
University of Technology



Learnable Models

K. N. Toosi
University of Technology



Learnable Models: Example

K. N. Toosi
University of Technology

Learnable Models: Example

Learnable Models: Input-output map

$$y = f(x)$$

$$f: R^m \rightarrow R^n$$

Learnable Models: Input-output map

K. N. Toosi University of Technology

$$y = f(x, \theta)$$

$$f: R^m \rightarrow R^n$$

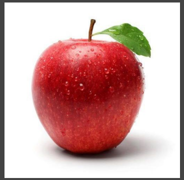
Learnable Models: Input-output map

$$y = f(x, \theta) \qquad \theta \in R^k$$

$$f: R^m \times R^k \rightarrow R^n$$

Learnable Models: Example

K. N. Toosi
University of Technology



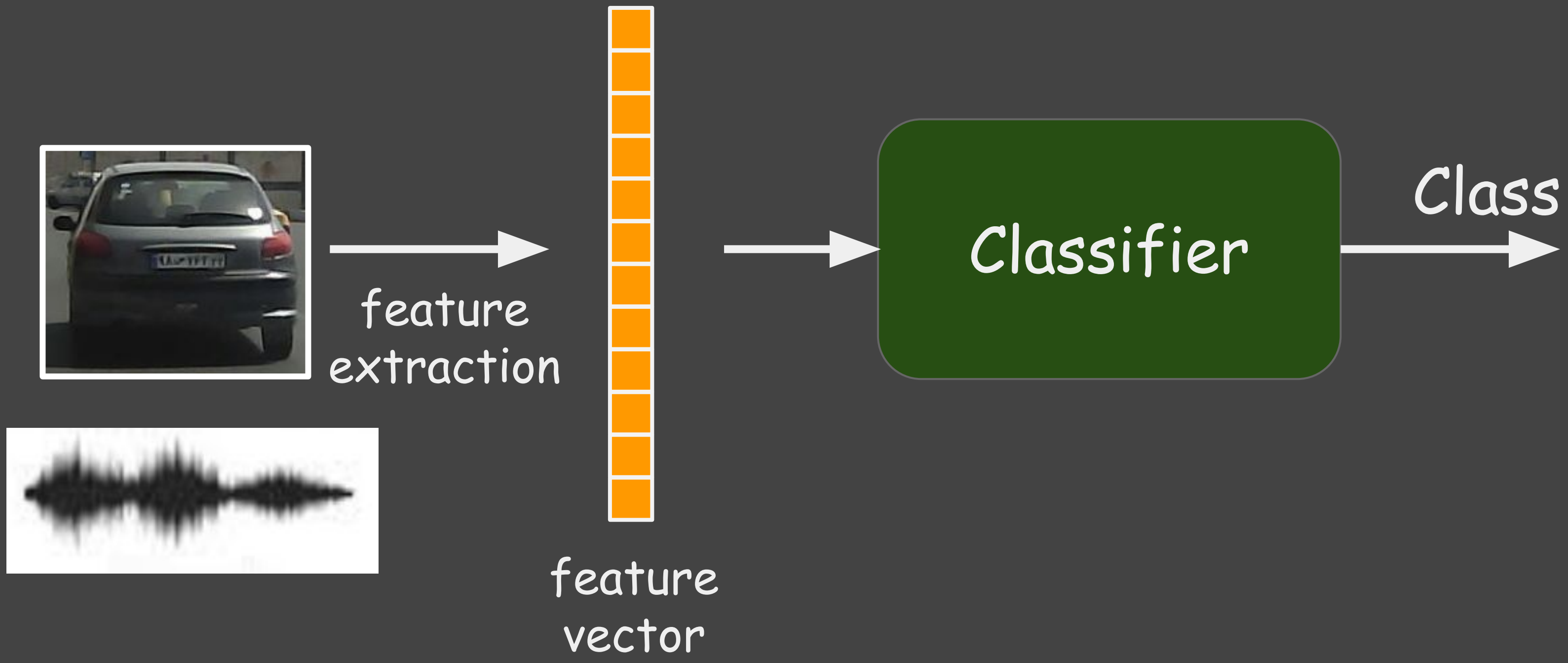
I

x = features(I) → f → y = 0

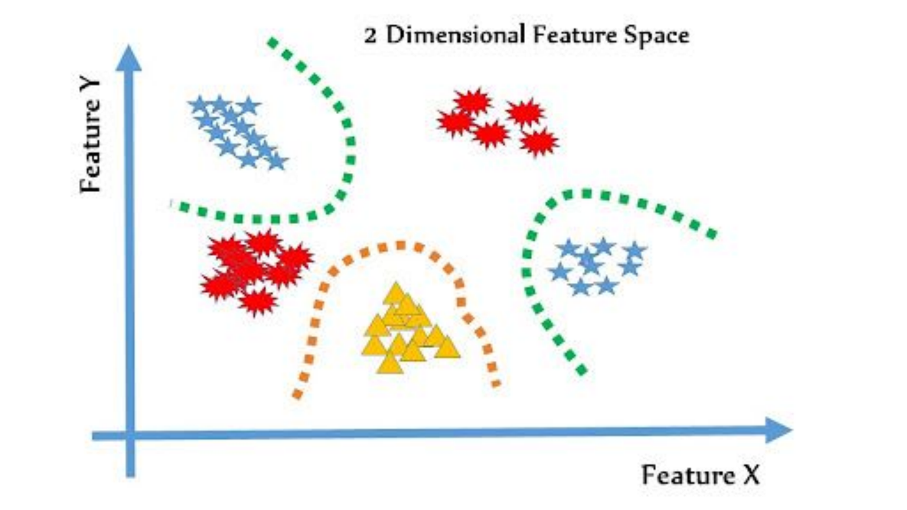
$$y = f(x,\theta)$$

$$f: R^m \times R^k \rightarrow R^n$$

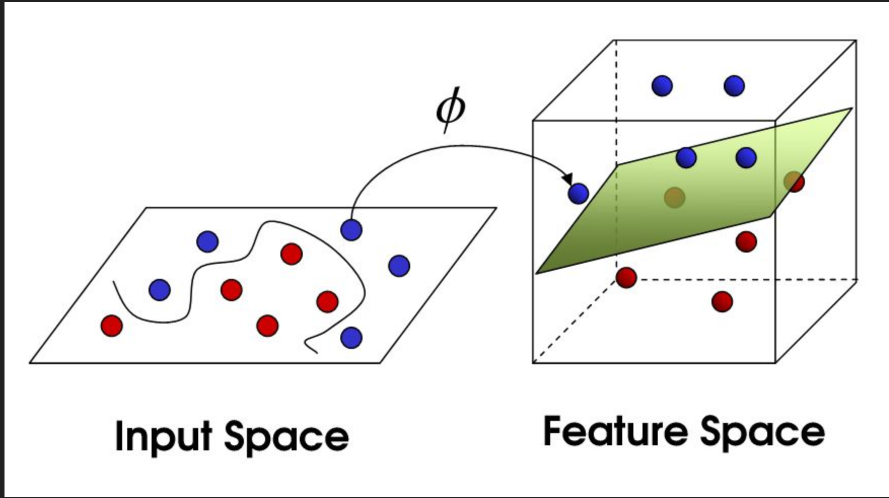
Features



Feature space



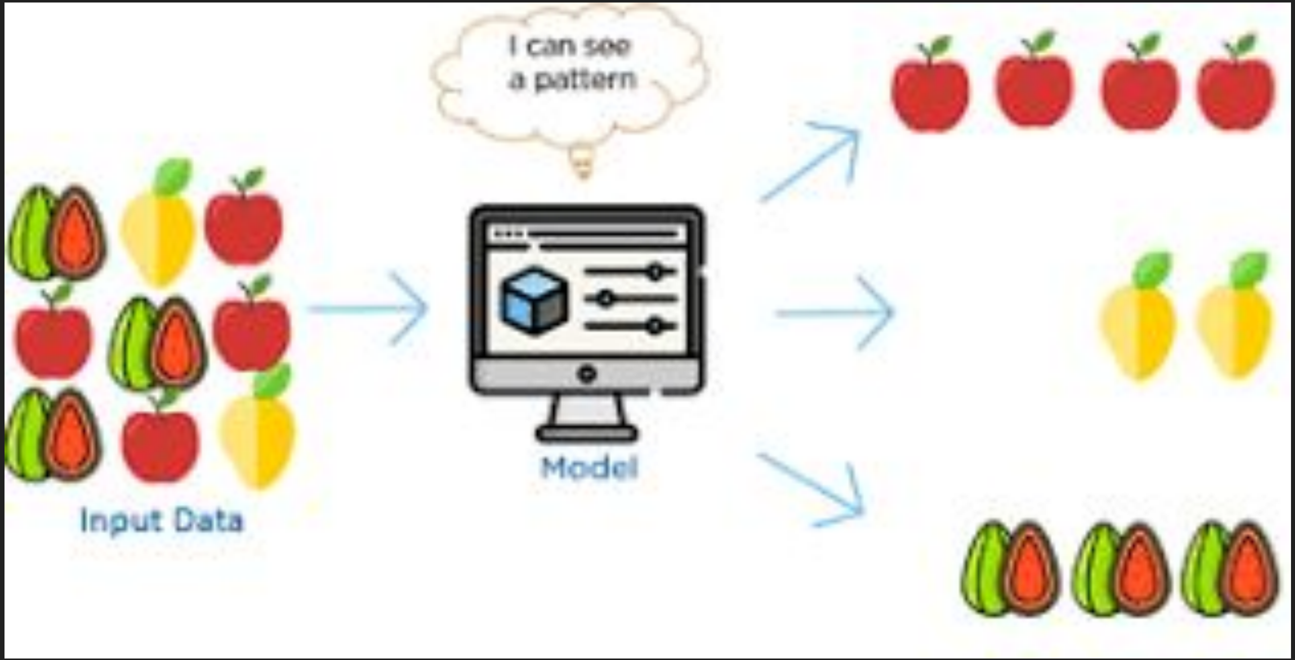
https://www.petersincak.com/news/why-i-do-not-believe-in-error-backpropagation/



https://towardsdatascience.com/the-kernel-trick-c98cdbcaeb3f

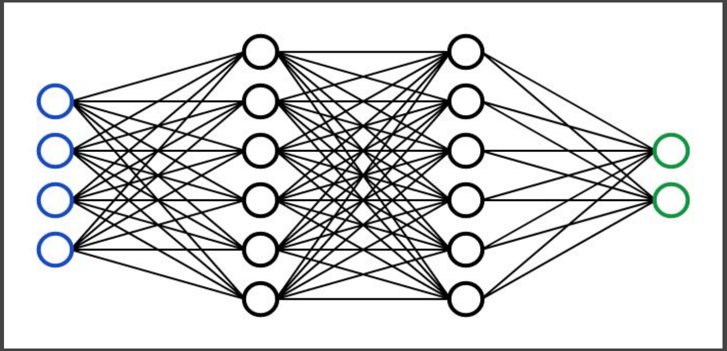
K. N. Toosi
University of Technology

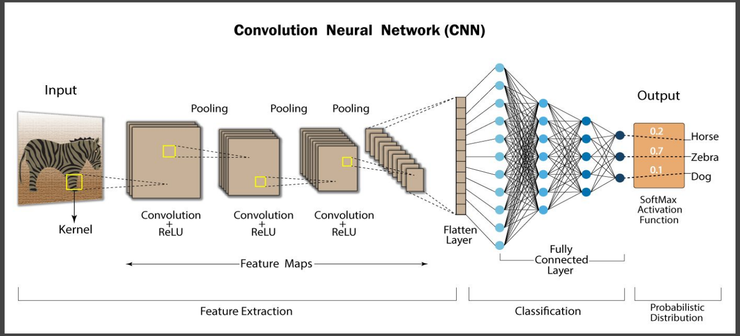
Unsupervised Learning

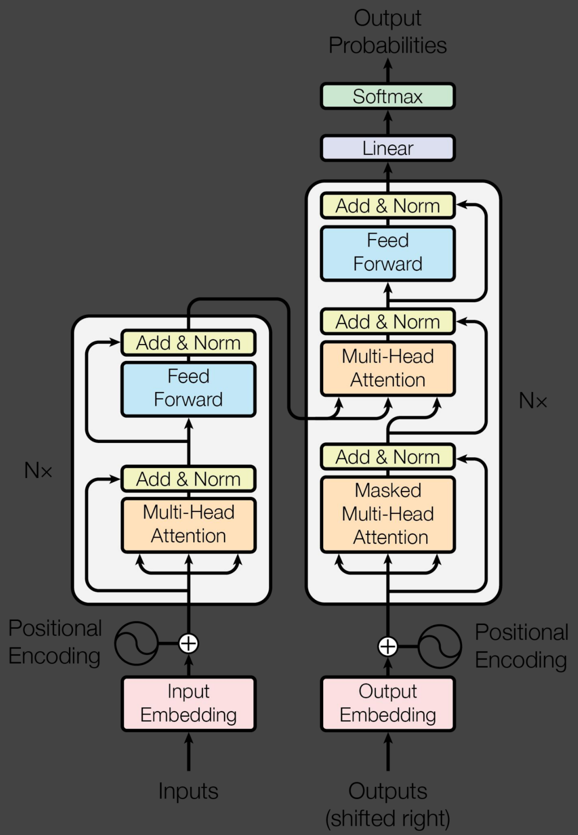


https://towardsdatascience.com/machine-learning-types-and-algorithms-d8b79545a6ec

Neural Networks







Natural Language Processing

K. N. Toosi
University of Technology

ChatGPT

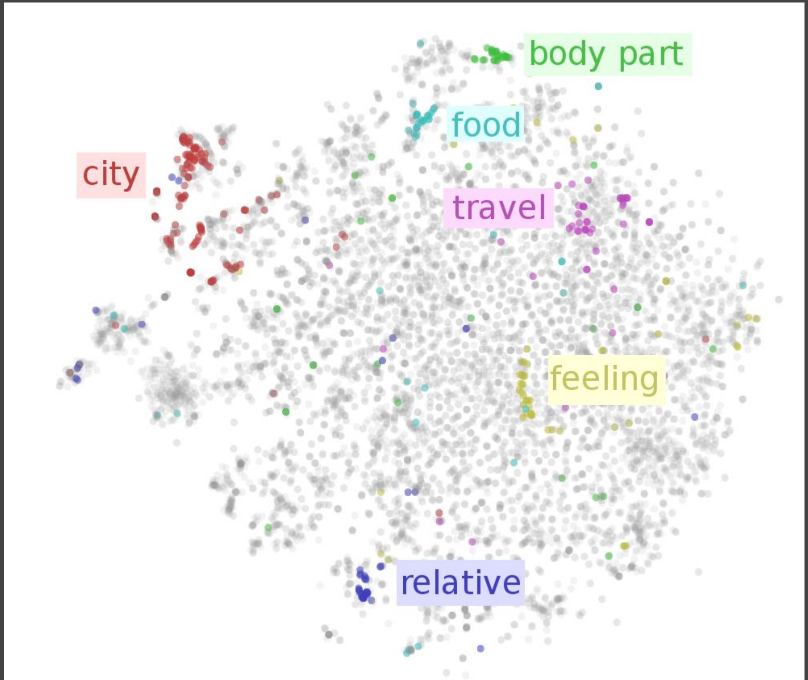
What are the applications of mathematics in AI?

Mathematics forms the foundation of many key concepts in AI, providing the theoretical frameworks, algorithms, and models that drive AI systems. Here are the main areas of mathematics and their applications in AI:

1. Linear Algebra

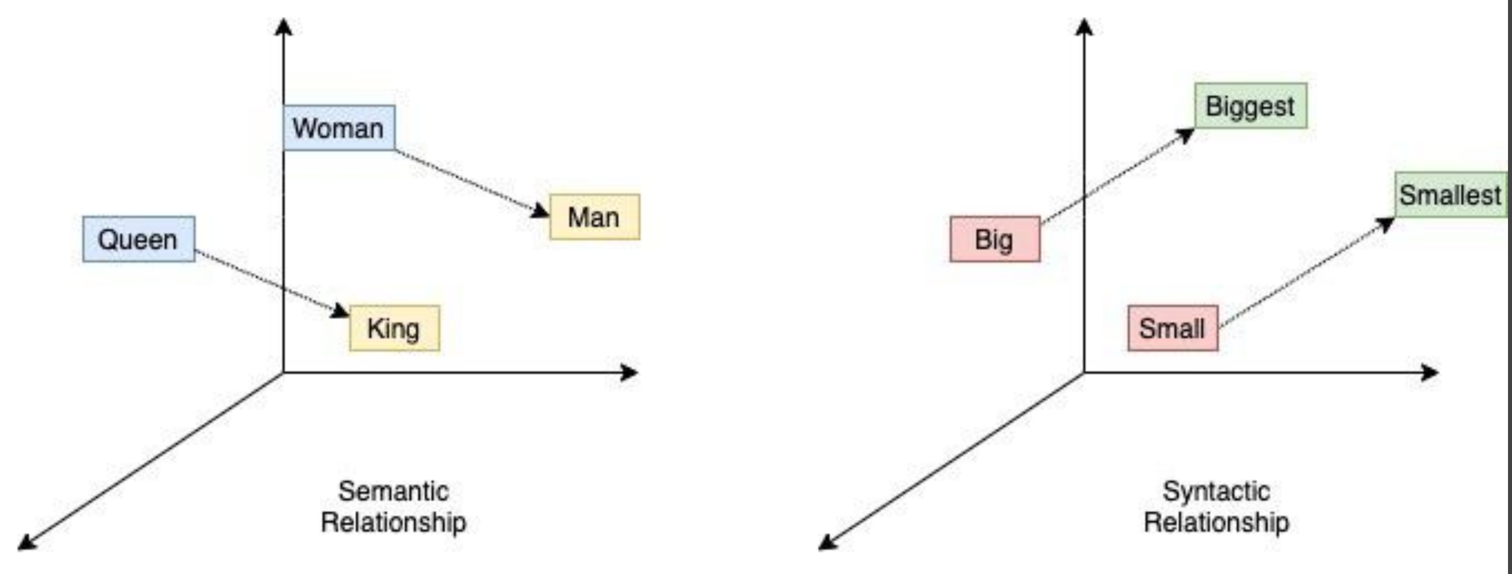
- **Applications**: Linear algebra is essential for representing and manipulating data in AI models. Vectors, matrices, and tensors are widely used to represent inputs, outputs, and transformations in algorithms like neural networks.
- **Specific Use Cases**:

Word Embedding



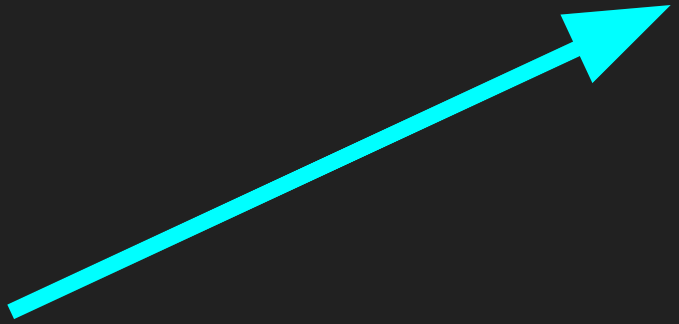
https://ruder.io/word-embeddings-1/

Word2Vec



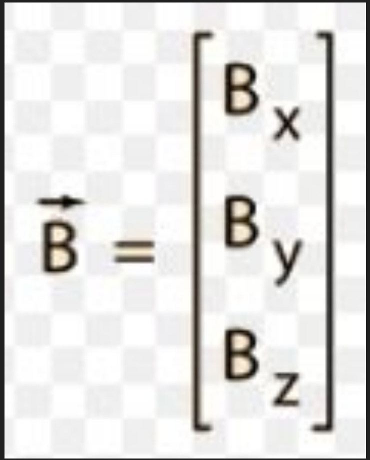
https://towardsdatascience.com/word2vec-research-paper-explained-205cb7eecc30

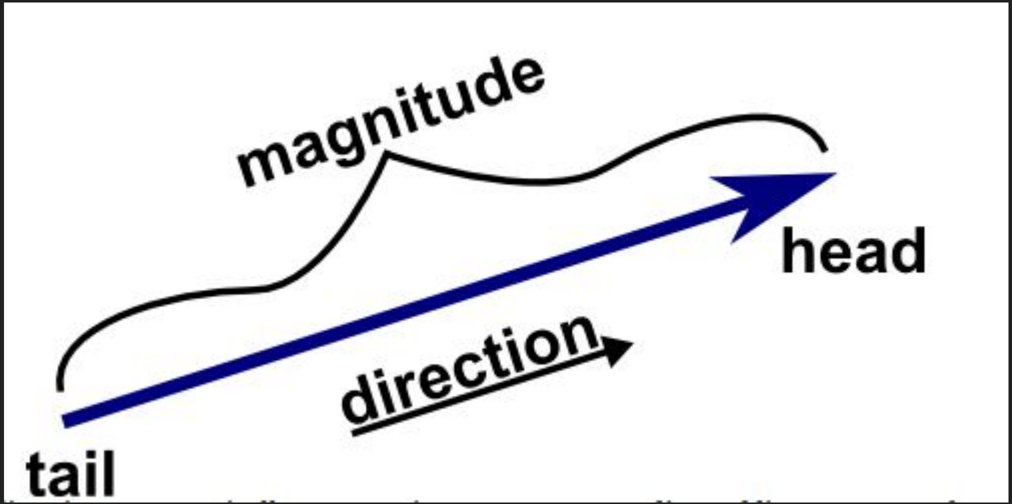
What is a Vector?



What is a Vector?

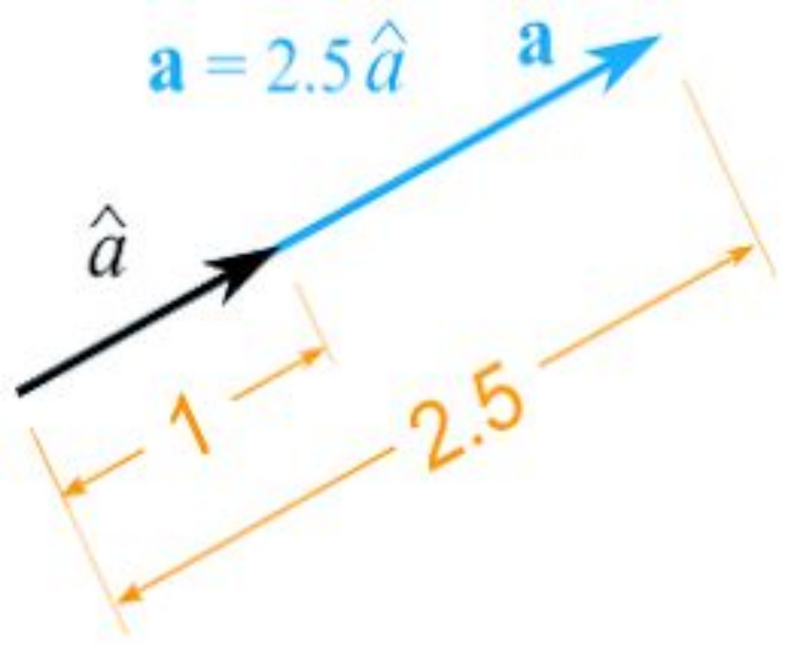
K. N. Toosi
University of Technology





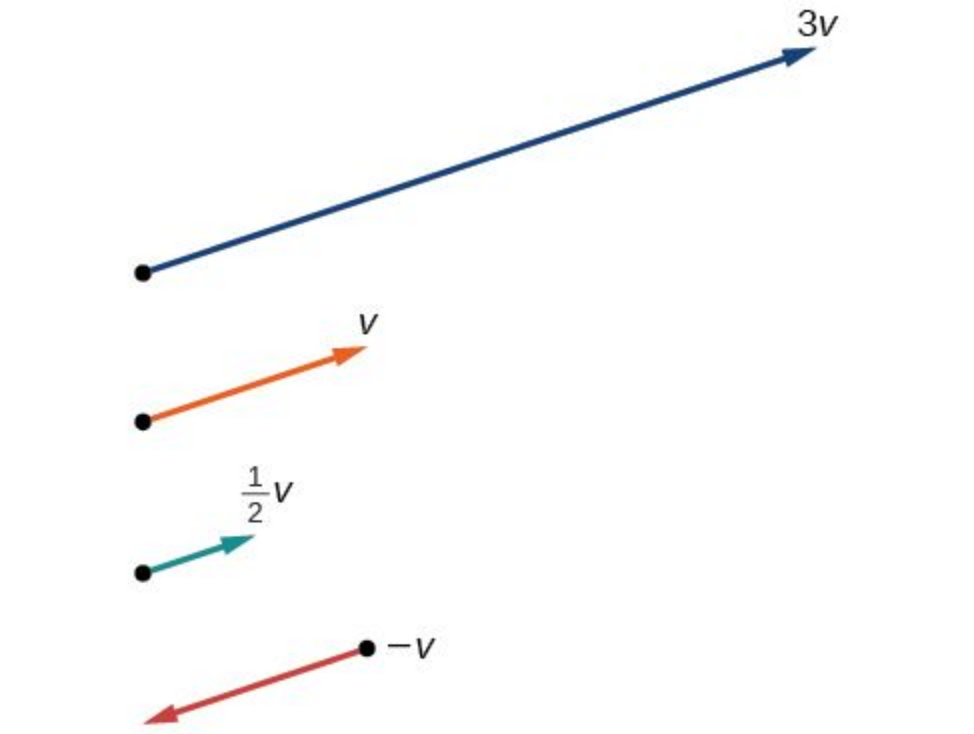
https://mathinsight.org/vector_introduction

Vector Scaling



https://semesters.in/unit-free-forced-fixed-vector/

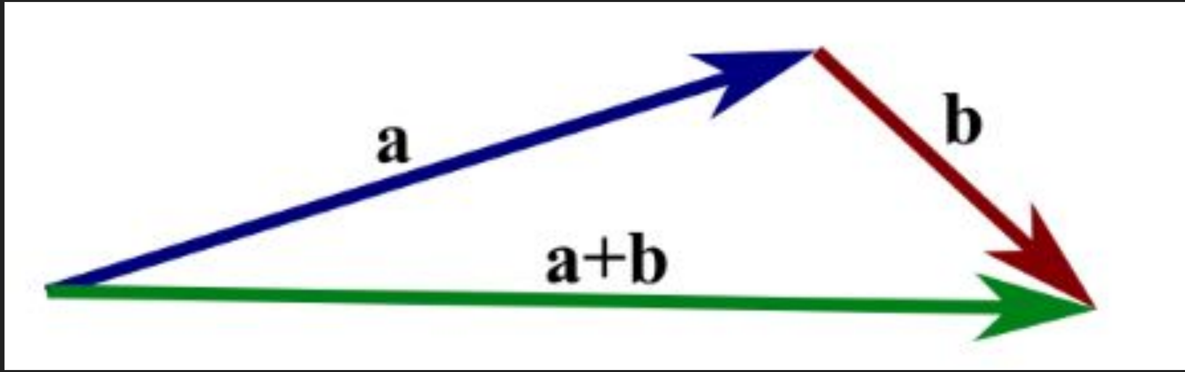
Vector Scaling



https://philschatz.com/precalculus-book/contents/m49412.html

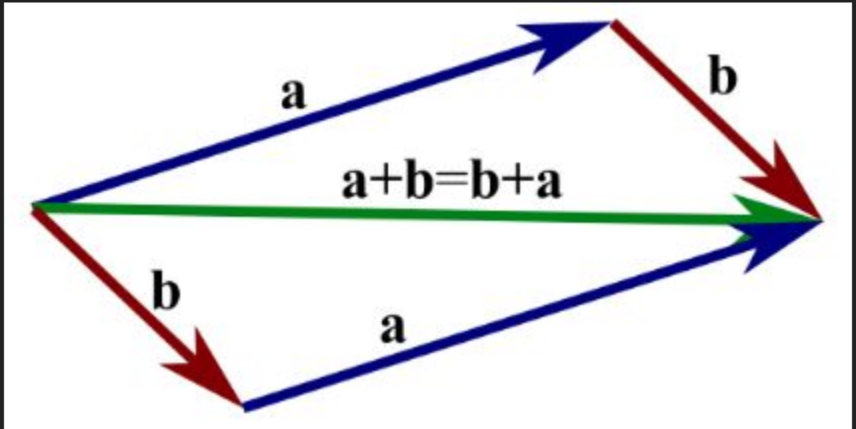
K. N. Toosi
University of Technology

Vector Addition



https://mathinsight.org/vector_introduction

Vector Addition



https://mathinsight.org/vector_introduction

Space

A set with a **structure**

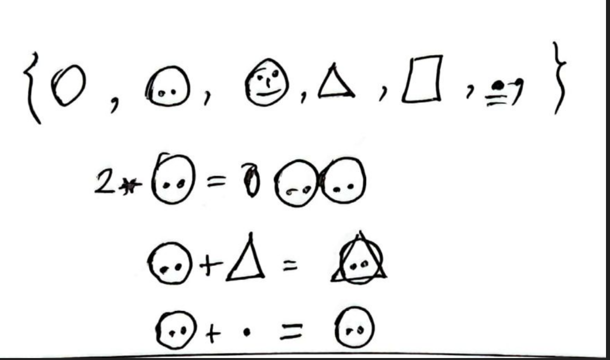
Vector Space

K. N. Toosi
University of Technology

Vector Space

K. N. Toosi
University of Technology

- a set V
- scalars $\in$ R (C, or any field)
- Vector addition + ($u + v$ for $u,v \in V$)
- scalar multiplication ($a\,u$ for $a \in R$, $u \in V$)
  - Commutativity: $u + v = v + u$
  - Associativity: $u + (v + w) = (u + v) + w$
  - Identity element: $\exists\ z \in V: v + z = z + v = v$
  - Inverse: for each $v \in V$ there is v' : $v + v' = z$ (z defined above)
  - $(ab)\,v = a\,(b\,v)$ (a,b are scalars)
  - $1\,v = v$
  - $a\,(u + v) = a\,u + a\,v$ (a is a scalar, u,v are vectors)
  - $(a+b)\,v = a\,v + b\,v$ (a,b are scalars, v is a vector)



Why bother?

K. N. Toosi
University of Technology

Why bother? adding apples and pears?

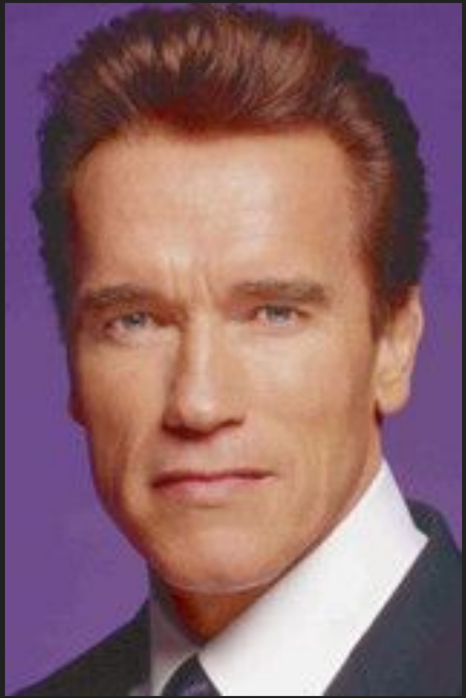
K. N. Toosi
University of Technology

+

= ?

Why bother?

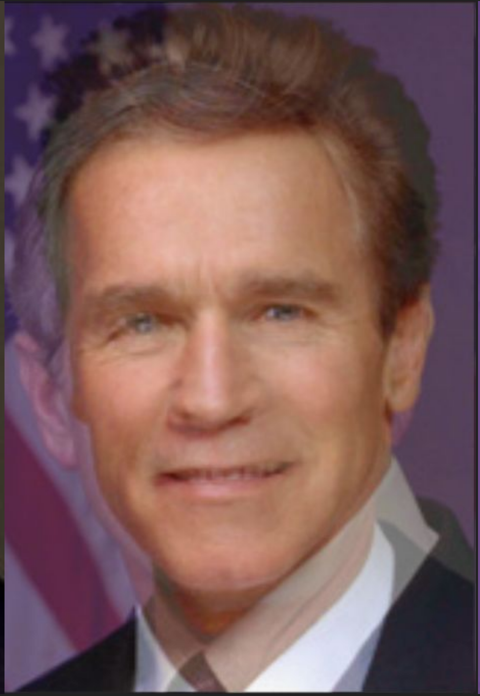
Jerk

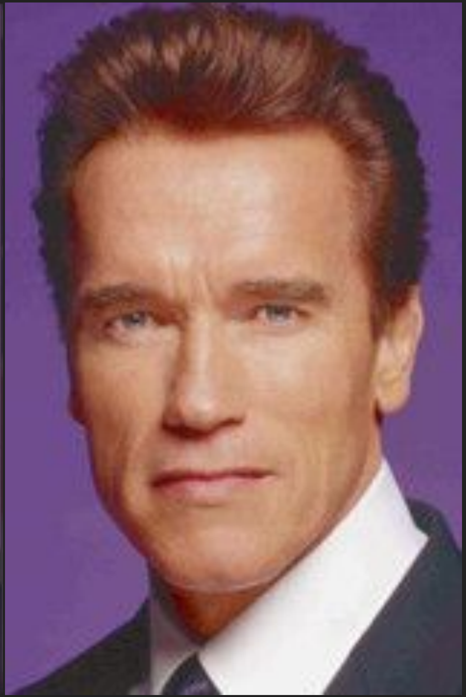


Cyborg

Image Averaging

K. N. Toosi
University of Technology





Jerk

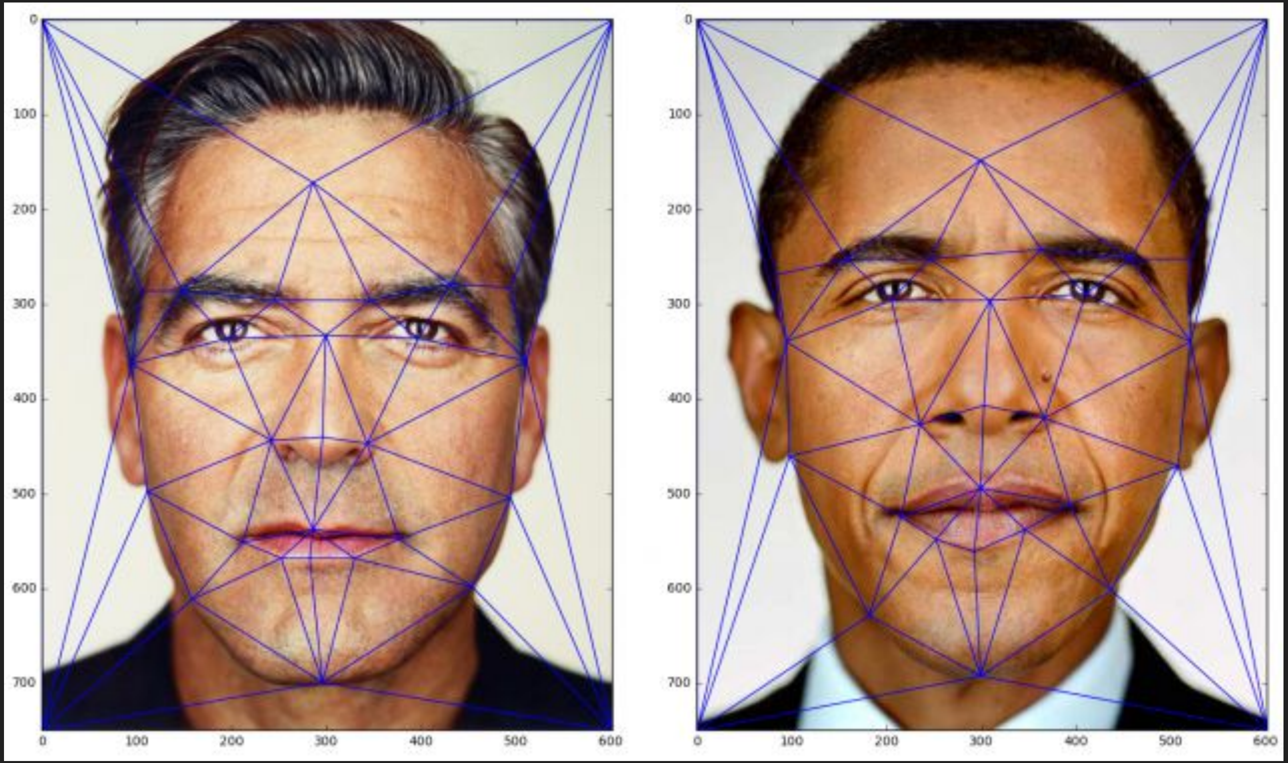
Cyborg

Shape+Appearance Averaging

K. N. Toosi
University of Technology

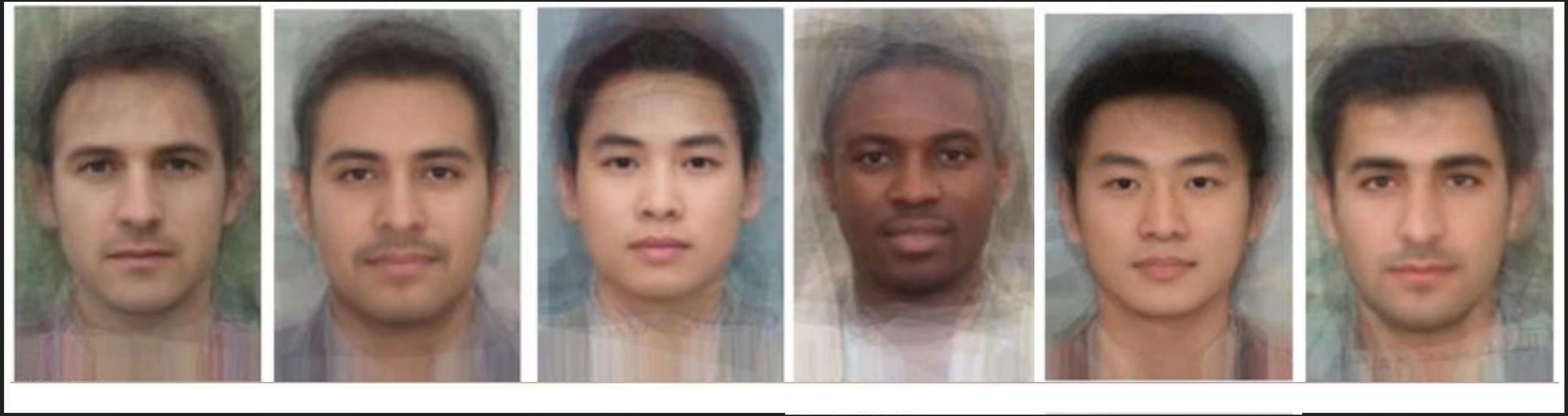
Jerk   Cyjerk   Cyborg

Why bother? Define vector addition and scaling

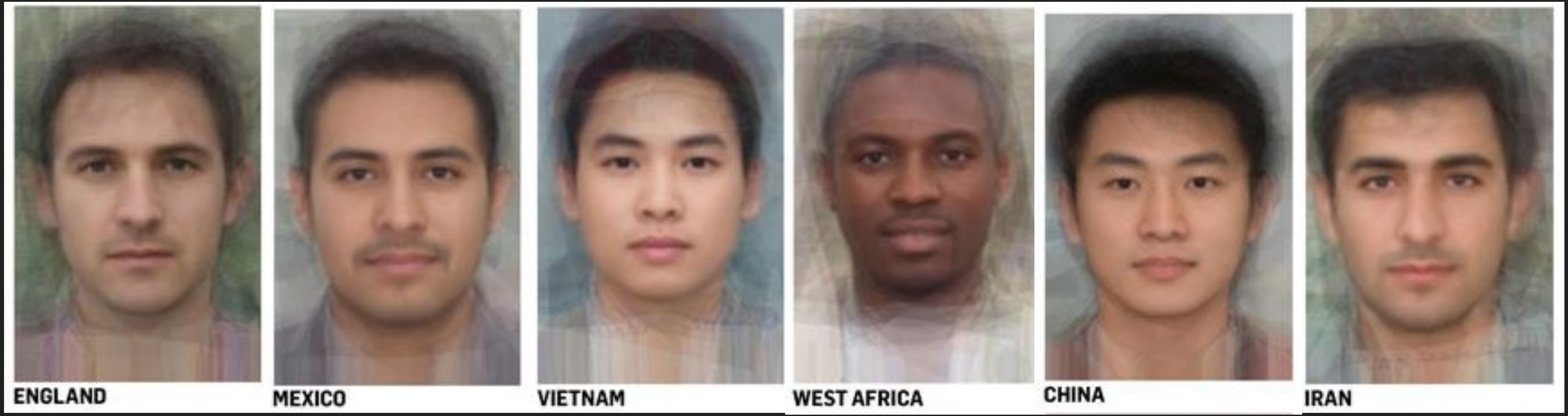


Why bother? Average Faces by country

K. N. Toosi
University of Technology



Why bother? Average Faces by country



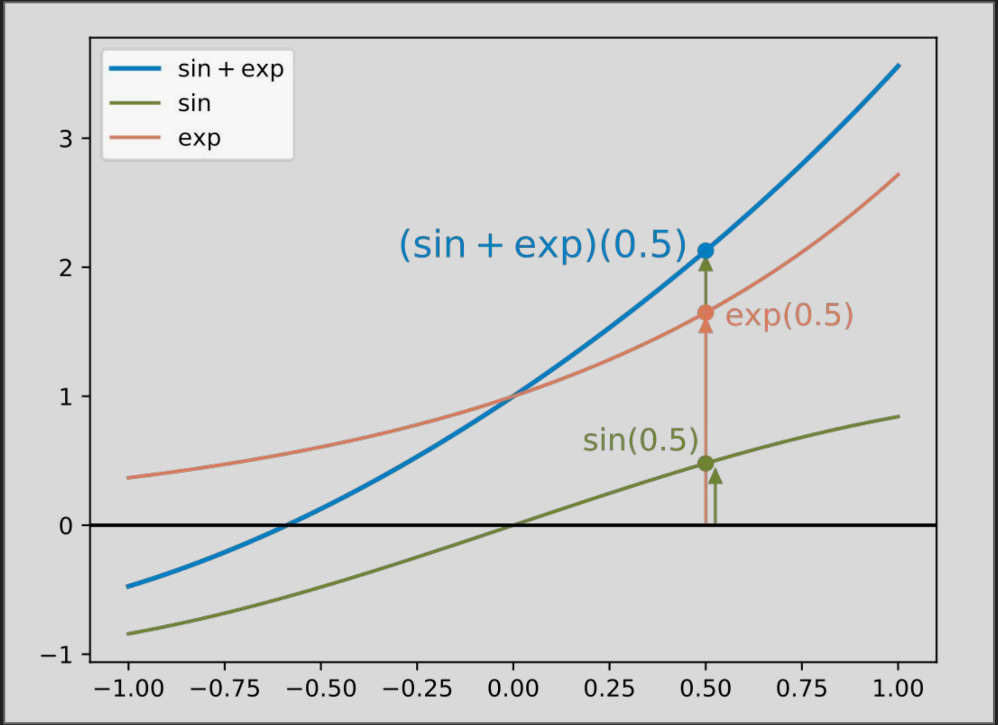
K. N. Toosi
University of Technology

Morphable Shape Models

K. N. Toosi
University of Technology

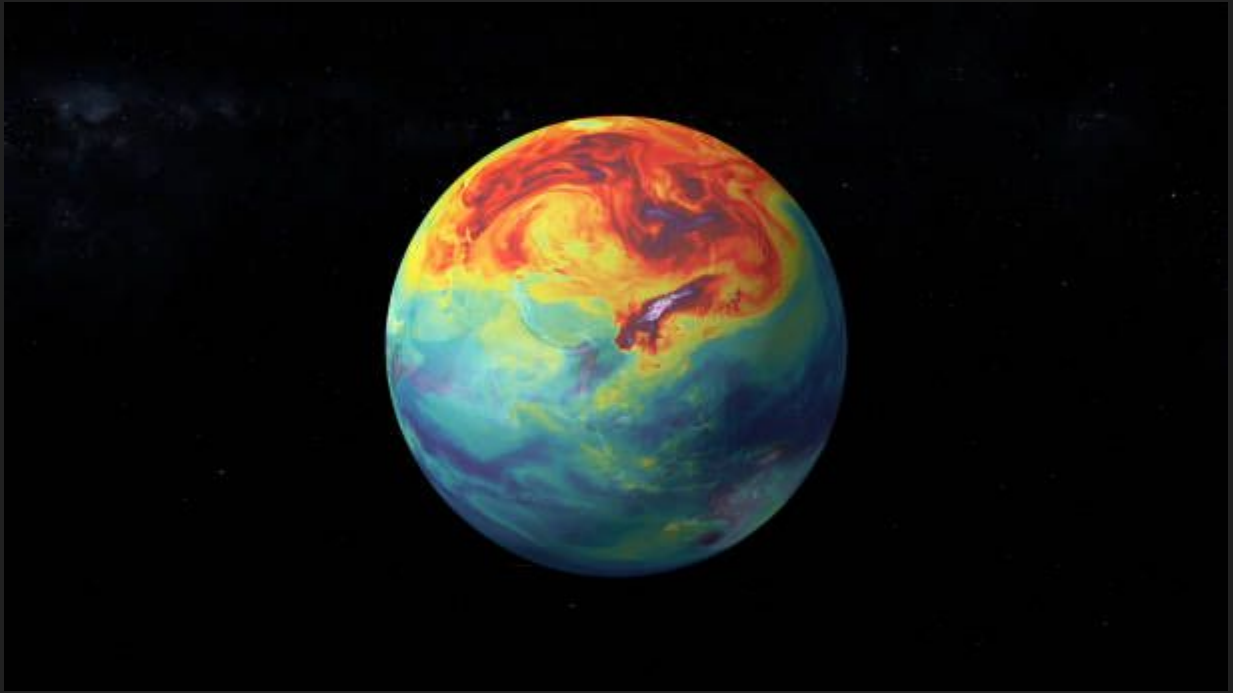
https://www.youtube.com/watch?v=kJPRCLhTEPg&t=36s

Why bother? functions as vectors



https://en.wikipedia.org/wiki/Vector_space

Why bother? functions as vectors



Linear combination

Let $a, b \in R$. The vector $a\,x + b\,y$ is a linear combination of the vectors x and y.

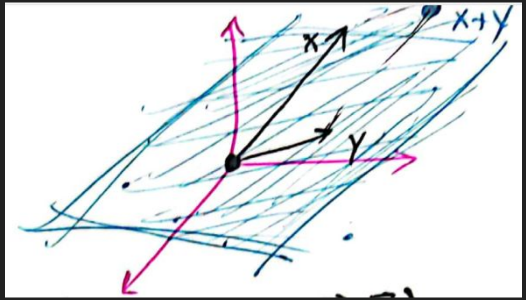
$$2\vec{x} + 3\vec{y}$$

$$a\vec{x} + b\vec{y} \qquad a, b \in \mathbb{R}$$

Let $a_i \in R$. The vector $a_1 x_1 + a_2 x_2 + \dots + a_n x_n$ is a linear combination of the vectors $x_1, x_2, \dots, x_n$.

K. N. Toosi
University of Technology

Span



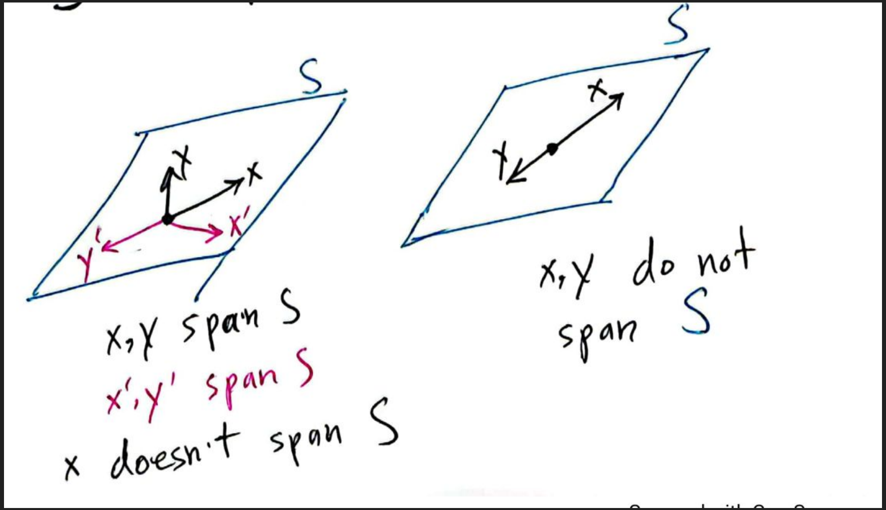
$$\text{span}(x,y) = \{ a\,x + b\,y \mid a,b \in R\}$$

The space of all linear combinations of x and y.

$$\text{span}(x_1, x_2, \ldots, x_n) = \{a_1 x_1 + a_2 x_2 + \ldots + a_n x_n \mid a_i \in R \}$$

Span

We say that $x_1, x_2, \ldots, x_n$ **span** S if $S = \text{span}(x_1, x_2, \ldots, x_n)$.



Linear dependence

K. N. Toosi
University of Technology

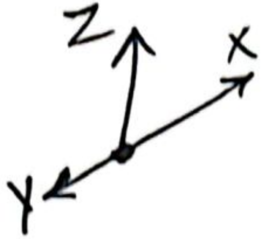
x, y, z are dependent if

- $x \in \text{span}(y,z)$, OR
- $y \in \text{span}(z,x)$, OR
- $z \in \text{span}(x,y)$

that is

- $x = a\,y + b\,z$, for some a, b, OR
- $y = a\,z + b\,x$, for some a,b, OR
- $z = a\,x + b\,y$, for some a,b.

Linear dependence - Example



MA2

$y = -0.5x$

$x \in \text{span}(y, z)$ → $x = -2y + 0z$

$y \in \text{span}(x, z)$

$z \notin \text{span}(x, y)$

x, y, z are linearly dependent

Linear dependence

$x_1, x_2, \ldots, x_n \in V$ are **linearly dependent** if one of them can be written as a linear combination of the others (one of them is in the span of the others).

Linear independence

K. N. Toosi
University of Technology

x,y,z are independent if

- $x \notin \text{span}(y,z)$, AND
- $y \notin \text{span}(z,x)$, AND
- $z \notin \text{span}(x,y)$

Linear independence

$x_1, x_2, \ldots, x_n \in V$ are **linearly independent** if none of them can be written as a linear combination of the others.

Linear independence

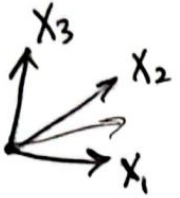
$x_1, x_2, \ldots, x_n \in V$ are **linearly independent** if none of them can be written as a linear combination of the others.

Equivalently:

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = 0 \;\Rightarrow\; a_1 = a_2 = \ldots = a_n = 0$$

Linear independence

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n = 0 \Rightarrow a_1 = a_2 = \dots = a_n = 0$$



$$a_1 \vec{X_1} + a_2 \vec{X_2} + a_3 \vec{X_3} = \vec{0}$$

$$\Rightarrow a_1, a_2, a_3 = 0$$

$$\vec{a_1} \neq 0$$

$$a_1 \vec{X_1} = (-a_2)\vec{X_2} + (-a_3)\vec{X_3}$$

$$\vec{X_1} = \left(-\frac{a_2}{a_1}\right)\vec{X_2} + \left(-\frac{a_3}{a_1}\right)\vec{X_3}$$

Basis

$v_1, v_2, \ldots, v_n \in V$ such that

- $v_1, v_2, \ldots, v_n$ are linearly independent
- $v_1, v_2, \ldots, v_n$ span V

Basis

$v_1, v_2, \dots, v_n \in V$ such that

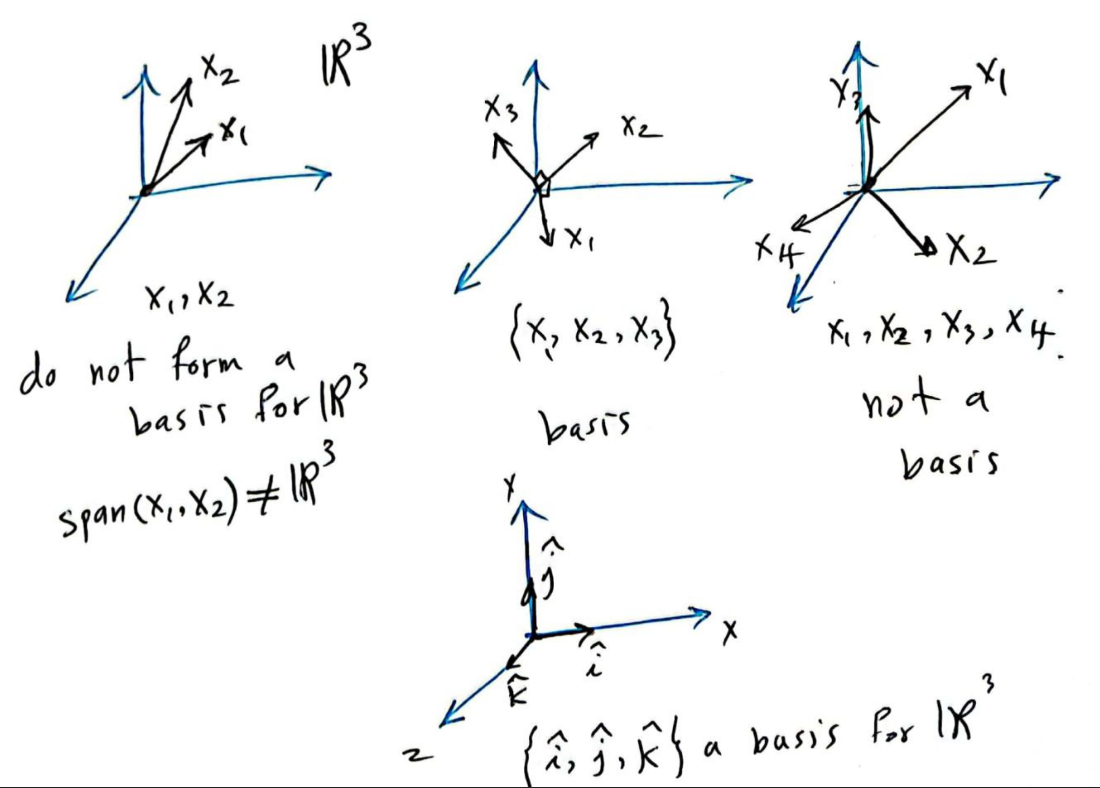
- $v_1, v_2, \dots, v_n$ are linearly independent
- $v_1, v_2, \dots, v_n$ span V

* n is the same for any choice of the basis vectors

* n is called the dimension of V

* There are also **infinite dimensional** vector spaces

Basis - Example



* Basis (general definition)

$\{v_i\}_{i\in I} \subseteq V$ such that

- v_i's are linearly independent
- for any $v \in V$ there is a **finite** set of vectors $v_1, v_2, \ldots, v_d \in \{v_i\}_{i\in I}$ such that $v \in$ **span**$(v_1, v_2, \ldots, v_d)$

* Any vector space has a basis

* cardinality of $\{v_i\}_{i\in I}$ is the same for any choice of the basis vectors

* cardinality of $\{v_i\}_{i\in I}$ is called the dimension of V

Bases and Coordinate Representation

Why is independence needed? => uniqueness

every $x \in V$ can be written **uniquely** as a linear combination of the basis vectors $v_1, v_2, \dots, v_n$.

Proof

MA2 (IV)

V is a finite dimensional vector space.

$\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ form a basis for V.

Consider an arbitrary vector $\vec{x} \in V$.

$$\vec{x} \in V \implies \vec{x} \in \text{span}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n)$$

$$\implies \vec{x} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n$$ for some $a_1, \ldots, a_n$.

$$\vec{x} = b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_n\vec{v}_n$$

Proof

K. N. Toosi
University of Technology

$$\Longrightarrow \vec{x} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n \quad \text{for some } a_1, \ldots, a_n.$$

$$\vec{x} = b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_n\vec{v}_n$$

$$(a_1-b_1)\vec{v}_1 + (a_2-b_2)\vec{v}_2 + \cdots + (a_n-b_n)\vec{v}_n = \vec{0}$$

$v_1 - v_n$ independent

$$\Longrightarrow a_1-b_1 = a_2-b_2 = \cdots = a_n-b_n = 0$$

$$a_1=b_1,\ a_2=b_2,\ \ldots,\ a_n=b_n$$

$$X = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \rightarrow$$

~~com~~ coordinates of x

Bases and Coordinate Representation

$\Rightarrow$ Every $x \in V$ can be written as a unique linear combination of $u_1, \dots, u_n$.

$$x = a_1 u_1 + a_2 u_2 + \dots + a_n u_n$$

$\Rightarrow$ x can be represented as $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}$.

$$x = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}$$ as an array of real numbers.

a_i-s are called coordinates of x

مختصات

مختصات به بردارهای پایه وابسته است

Example: The Euclidean space

K. N. Toosi
University of Technology