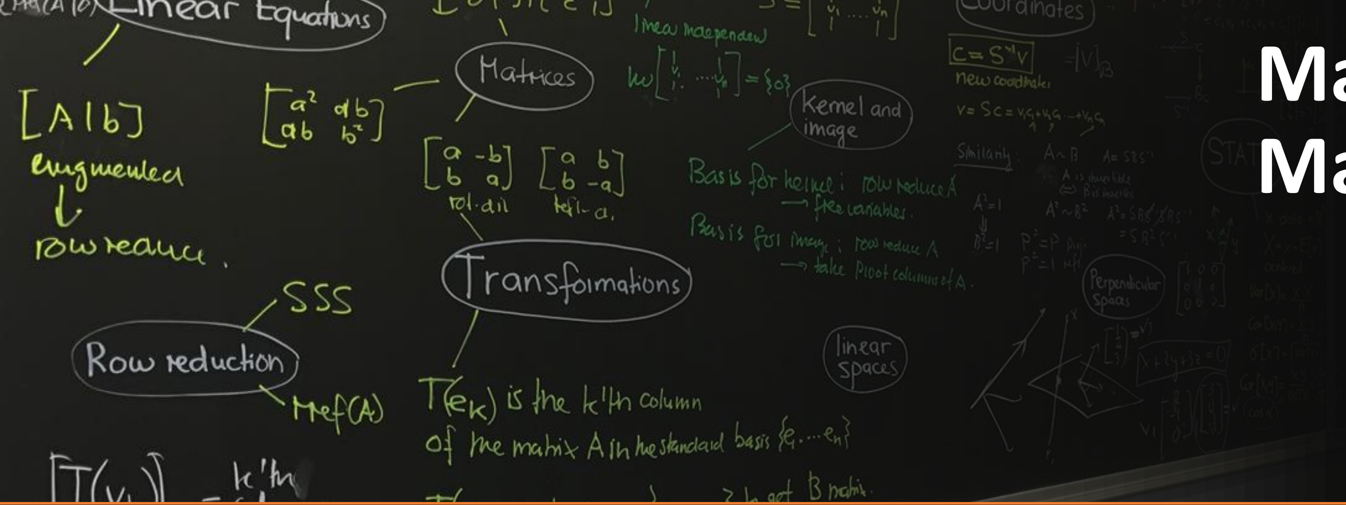


Mathematical Foundations for Machine Learning and Data Science

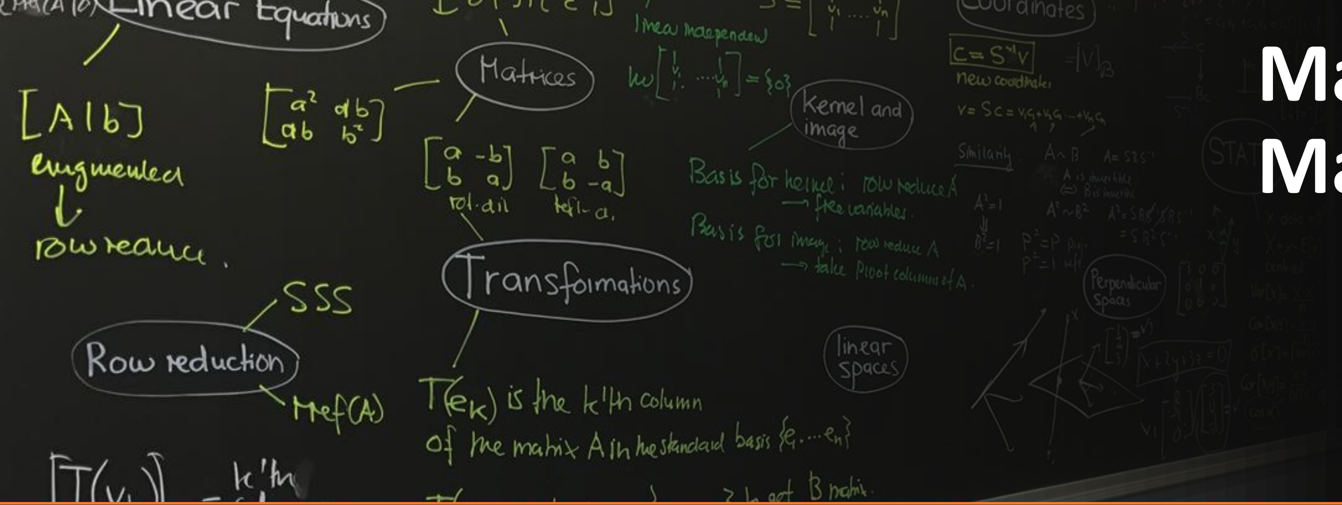
Vectors and Operations on Vectors

Zubair Khalid

Department of Electrical Engineering
School of Science and Engineering
Lahore University of Management Sciences

https://www.zubairkhalid.org/ee212_2021.html

LUMS
A Not-for-Profit University



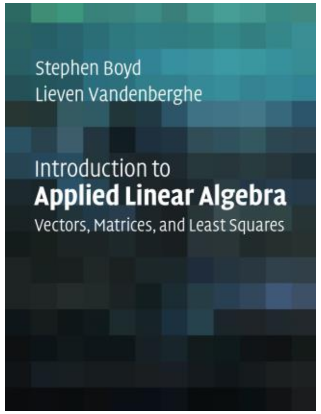
Mathematical Foundations for Machine Learning and Data Science

Vectors – Notation and Basic Operations

LUMS
A Not-for-Profit University

Outline

- *Vectors notation*
- *Applications*
- *Basic vector operations*



Chapter 1

ML Example

Problem:

Given a photo of a person,
- she/he is **smiling** or not (emotion detection)
- she/he is **sleeping** or not (eyes open/close)

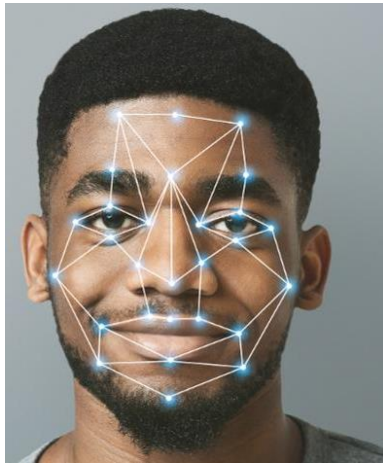
Machine Learning Problem:

Step 1:
Gather data and label it

Step 2:
Extract useful features

Step 3:
Develop a model

Step 4:
Train and evaluate model



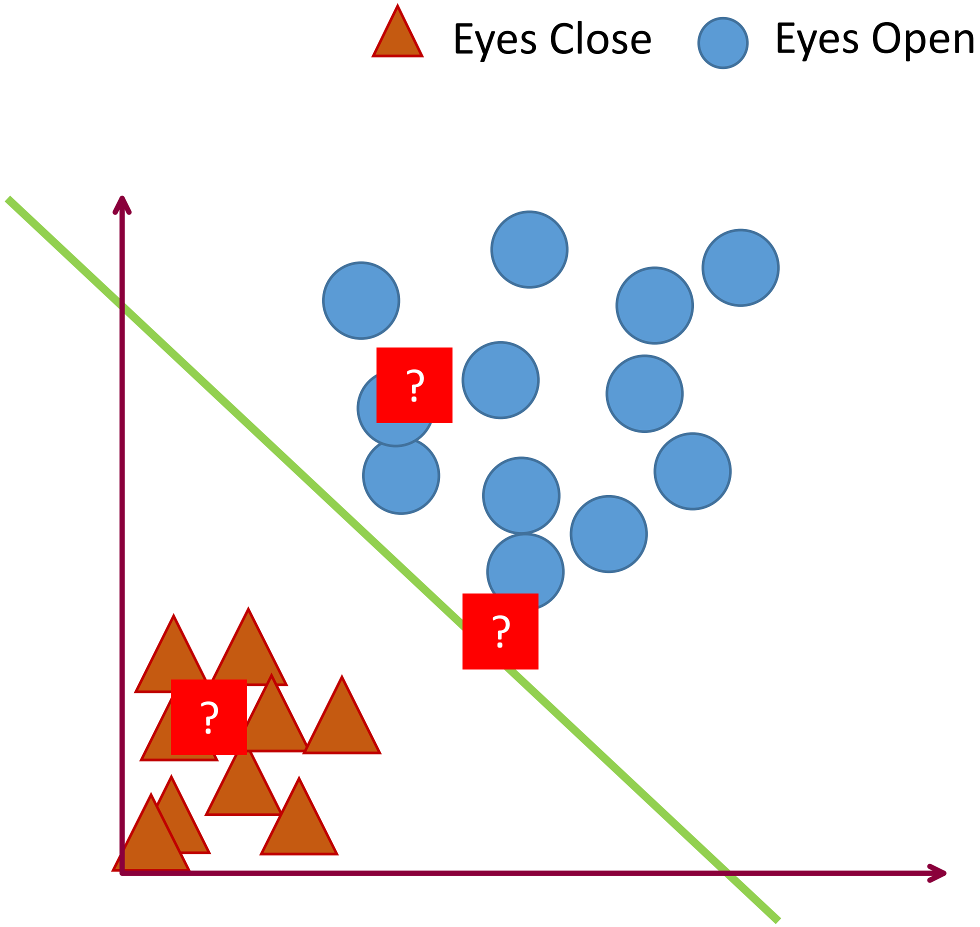
ML Example

Eyes Open/Close Problem:

Step 2:
Extract useful features

Step 3:
Develop a model

Step 4:
Train and evaluate model



Vectors

Definition:

A vector is an ***ordered*** finite list of numbers (real or complex).

Notation:

Usually denoted by a letter symbol; stack the list of numbers in an ordered form.

For example, consider a vector of 4 real numbers given by

$$a = \begin{bmatrix} -1.1 \\ 30.1 \\ 6.1 \\ -2.7 \end{bmatrix} \qquad a = \begin{pmatrix} -1.1 \\ 30.1 \\ 6.1 \\ -2.7 \end{pmatrix} \qquad a = (-1.1,\ 30.1,\ 6.1,\ -2.7)$$

Size of a vector: Number of elements the vector contains (also referred to as length or dimension).

We usually express vector b of size n as $b \in \mathbf{R}^n$ and call it n-vector.

Entry of a vector: b_k - k-th entry of the vector b. For example, $a_2 = 30.1$ for the vector a defined above.

Vectors

Zero vector: A vector with all elements equal to zero.

denoted by $\mathbf{0} \in \mathbf{R}^n$.

One vector: A vector with all elements equal to one.

denoted by $\mathbf{1} \in \mathbf{R}^n$.

Unit vector: A standard unit vector is vector with all elements zero except one element that is equal to one.

denoted by $e_i \in \mathbf{R}^n$ and is defined as

$$(e_i)_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} = \delta_{i,j}$$

k-Sparse vector: A vector with at-most **k** non-zero entries.

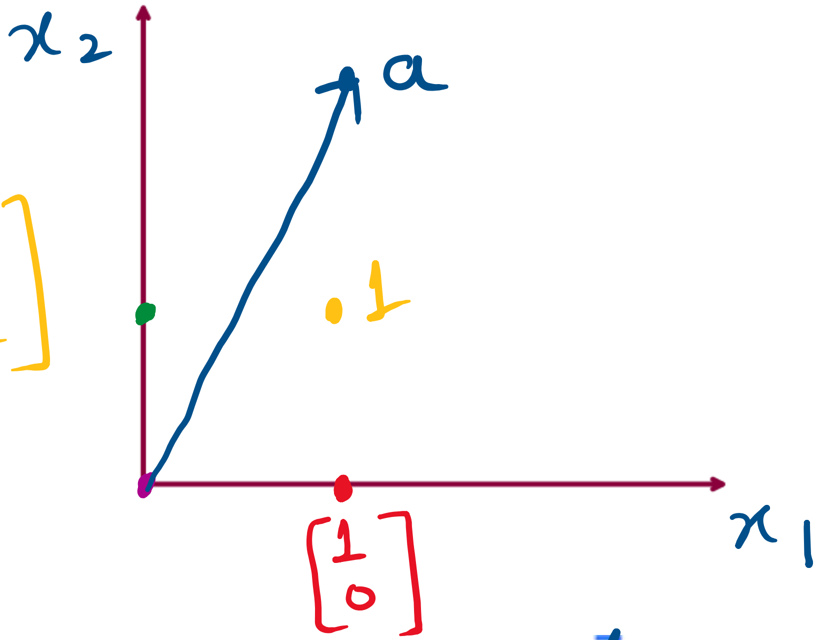
Geometric Interpretation

2-vector:

$e_1 \quad 0$

$e_2 \quad 1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

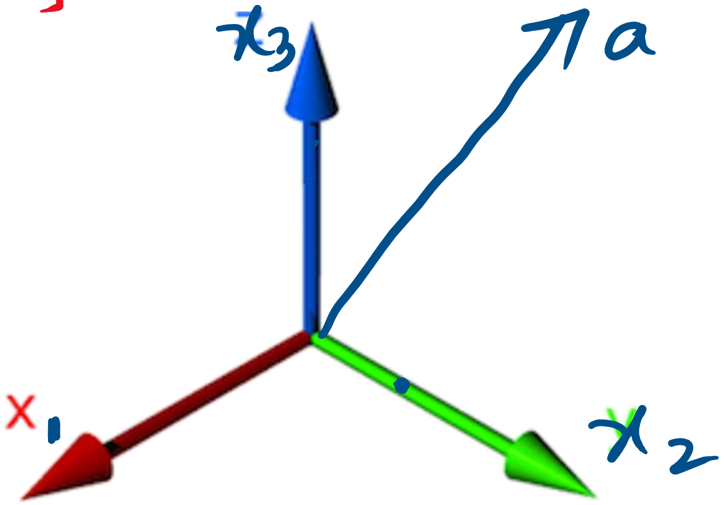
$a = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$



R^2

3_vector

$a = [-2, 1, 2]$



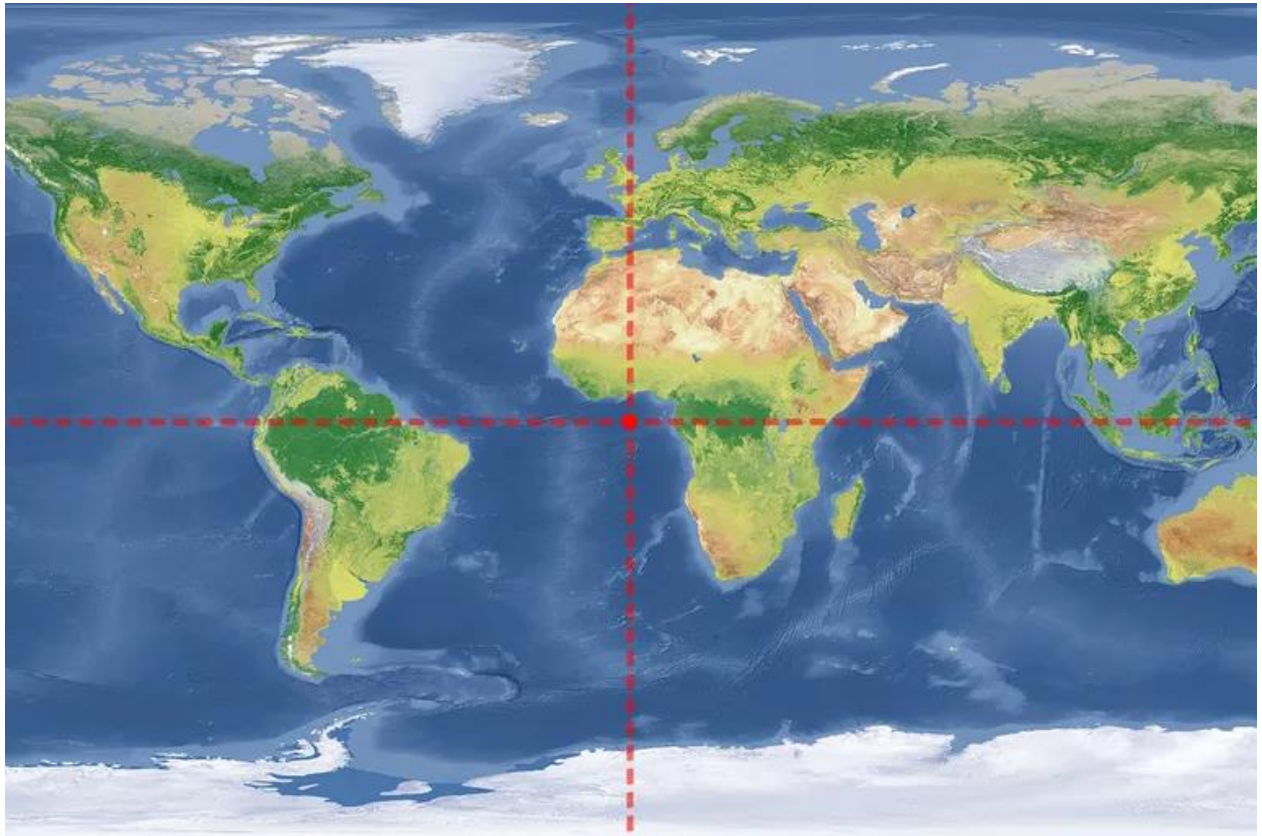
$a \in R^n$

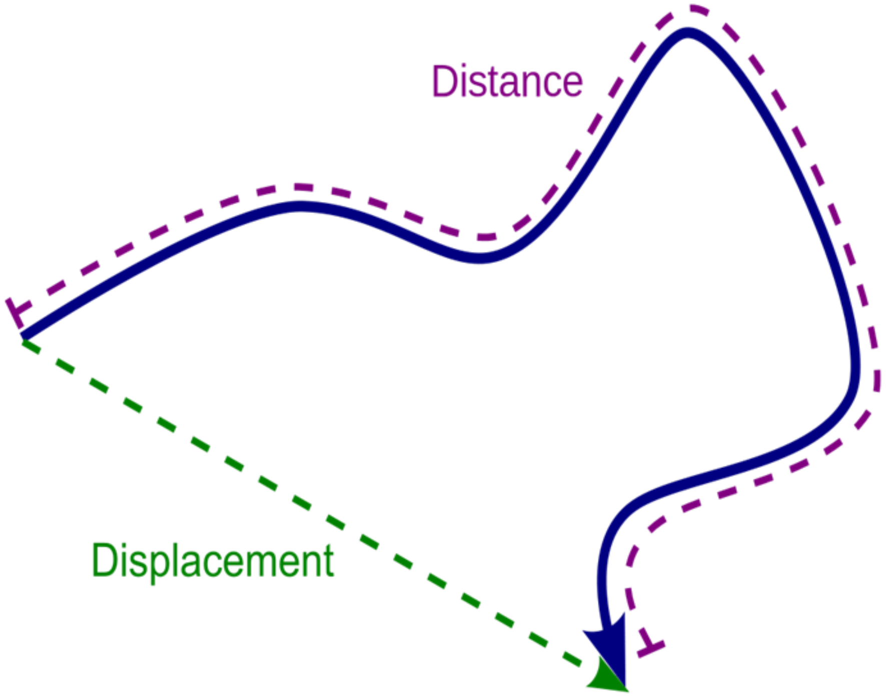
$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = a_1 e_1 + a_2 e_2 + \dots a_n e_n$$

Examples of Vectors - Applications

Location

Displacement, Velocity, Acceleration





Examples of Vectors - Applications

Color

Each color is represented by 3-vector.

Quantities

An n-vector q can represent the amounts or quantities of n different resources or products held (or produced, or required) by an entity such as a company.

For example, n-vector represents the quantity of n products stocked in a warehouse.

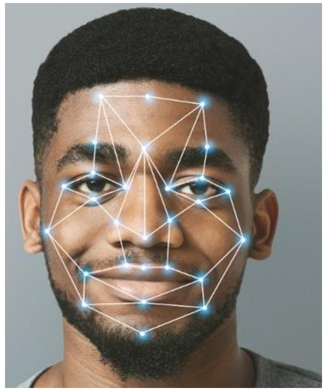
Values across a population

An n-vector can give the values of some quantity across a population of individuals or entities.

For example, an n-vector **a** can represent the blood pressure of a collection of n patients, with a_i the blood pressure of patient i, for i = 1, 2, …, n.

Examples of Vectors - Applications

Image

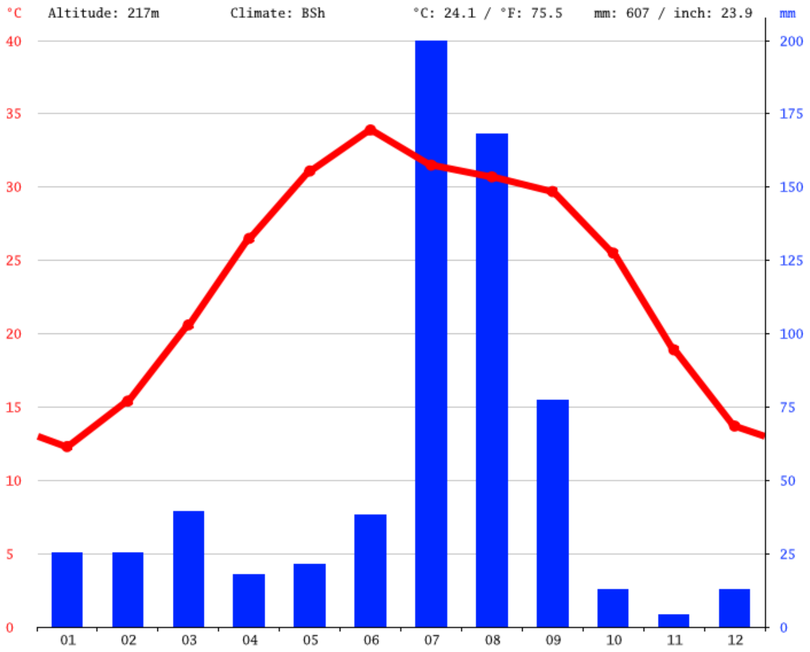


Time series

12-vector can represent the average monthly temperature, rainfall, pressure etc of Lahore.

30-vector can represent the number of expected COVID-19 patients in Pakistan cases over the next 30 days.

Other examples include exchange rate, audio, and, in fact, any quantity that varies over time.

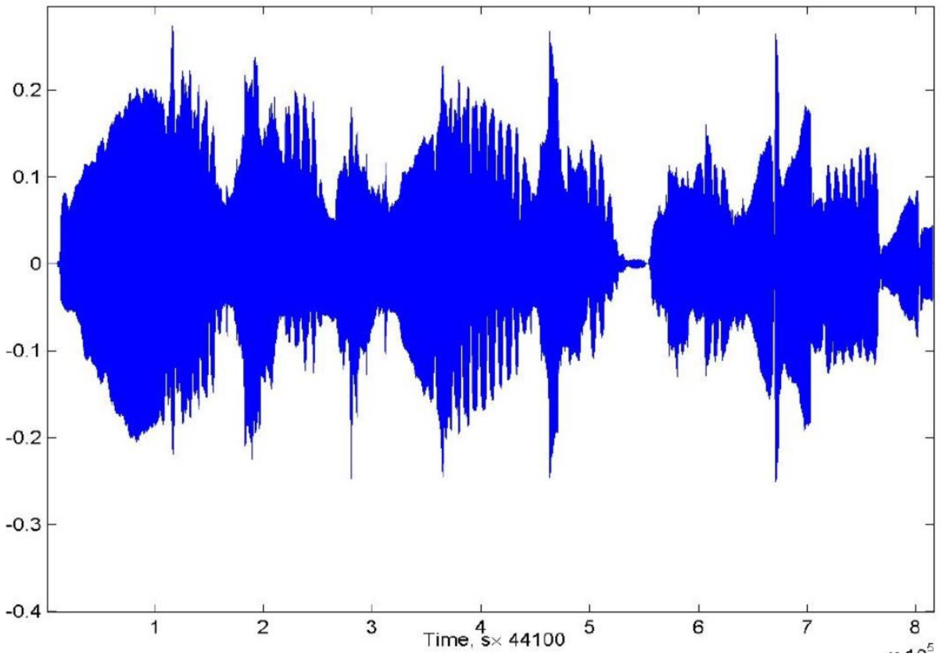


Examples of Vectors - Applications

Audio

Sound (Flute) versus time

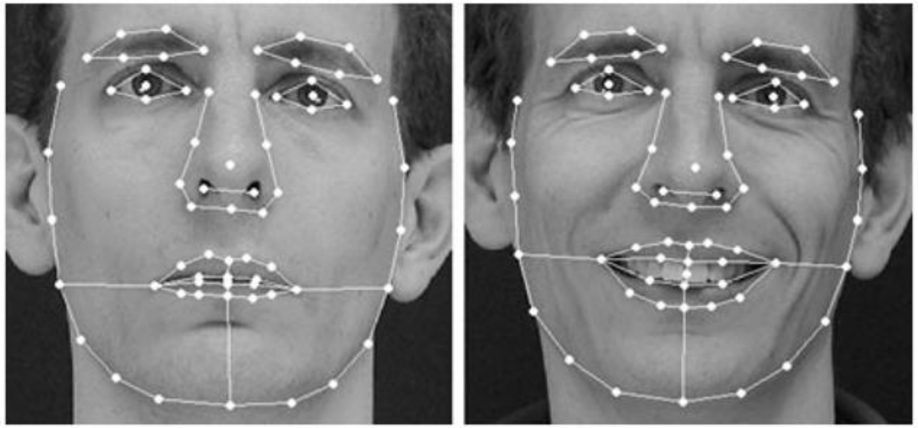
8 second sound = 44100x8-vector



Feature or Attribute

In Machine Learning, classification is mostly carried out collecting features or characteristics derived from the object.

Such a vector is sometimes called a feature vector, and its entries are called the features or attributes.



Emotion Recognition

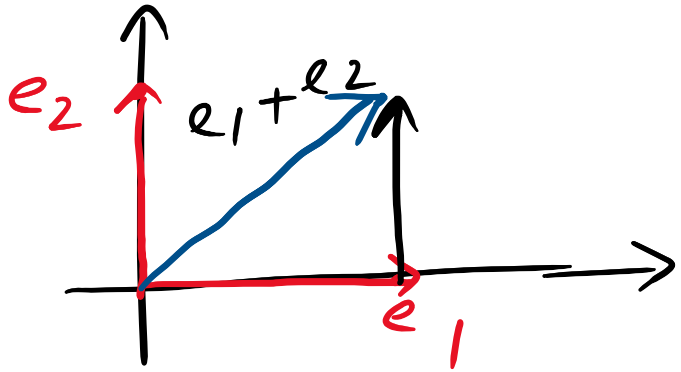
Operations on Vectors

Additivity

$a \in R^n, \quad b \in R^n$

$$c = a + b = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix}$$

$$c_i = a_i + b_i, \quad i = 1, 2, \ldots, n$$



* Commutative $\quad a + b = b + a$

* Associative $\quad a + (b + d) = (a + b) + d$

Operations on Vectors

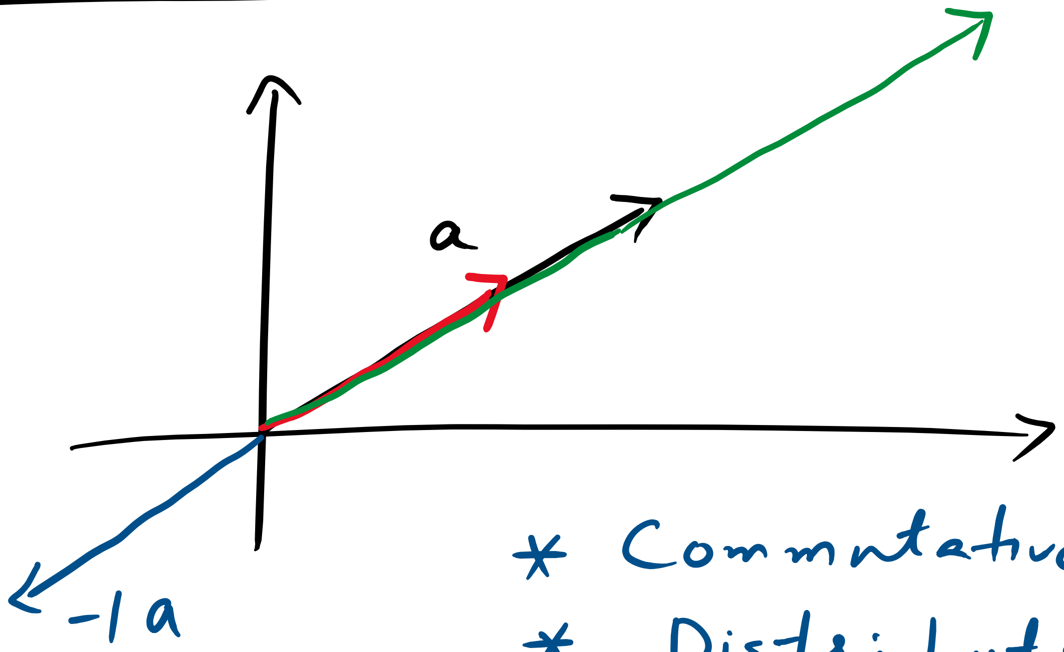
Scaling

$a \in R^n, \quad \beta \in R$

$b = \beta a = [\beta a_1, \beta a_2, \dots, \beta a_n]$

$\boxed{b_i = \beta a_i}$

*



R^2

$|\beta| < 1$
shrinking

$|\beta| > 1$
expansion

$\beta \rightarrow -ve$
change in direction

* Commutative
* Distributive

Operations on Vectors

Linear Combination

$a_1, a_2 \dots a_m \in R^n, \quad \beta_1, \beta_2 \dots \beta_m \in R$

$\beta_1 a_1 + \beta_2 a_2 + \dots + \beta_m a_m \in R^n$

* Affine Combination ; $\beta_1 + \beta_2 + \dots \beta_m = 1$

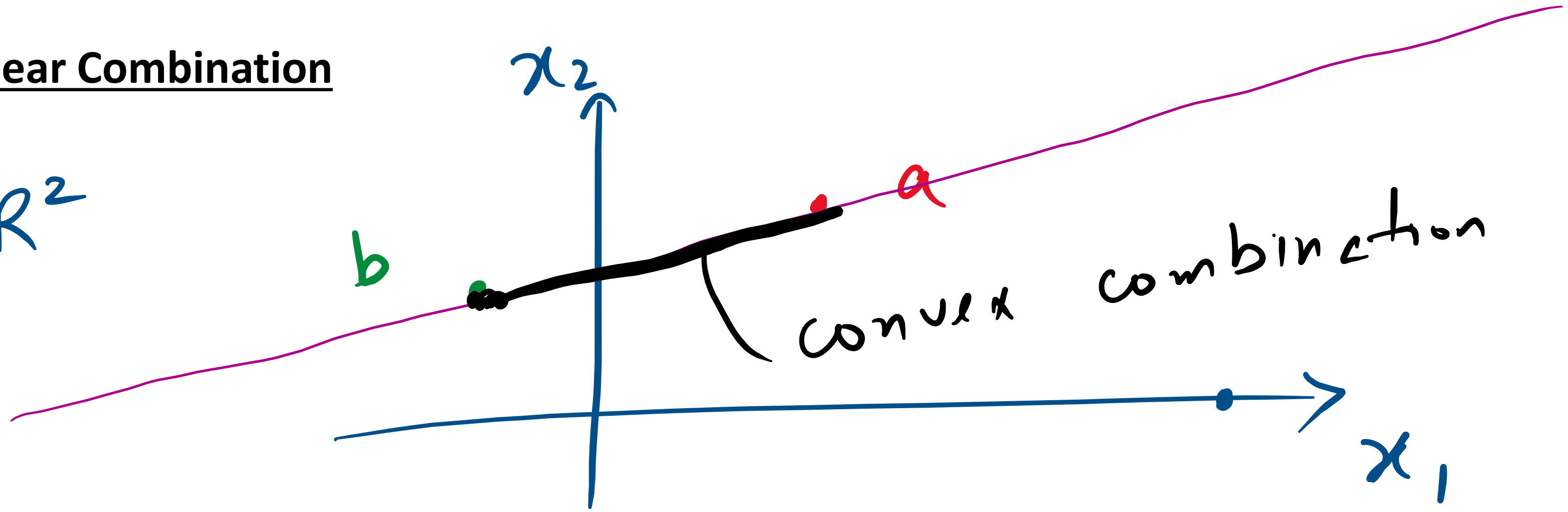
* Convex Combination ; $\beta_1 + \beta_2 + \dots \beta_m = 1$

$0 \leq \beta_i \leq 1, \quad i = 1, 2, \dots m$

Operations on Vectors

Linear Combination

R^2



Affine Combination

$$\beta a + (1-\beta) b = \beta a + b - \beta b$$
$$= b + \beta (a - b)$$

Operations on Vectors

Inner Product $a, b \in R^n$

* $\langle a, b \rangle = a^T b = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$
* $\langle a, b \rangle = \langle b, a \rangle$
* Projection of a on b.

$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \quad b = e_2$

$b = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$

* $a^T b = \langle a, b \rangle = a_2$

Operations on Vectors

Inner Product - Applications

* Sum of elements

$a \in R^n$

$$\langle a, 1 \rangle = a_1 + a_2 + \cdots + a_n$$

* Average of a, $\bar{a}$, $avg(a) = \langle 1(\frac{1}{n}), a \rangle$

$$= \frac{a_1 + a_2 + \cdots + a_n}{n}$$

* $\langle a, a \rangle = a_1^2 + a_2^2 + \cdots a_n^2$

* Weighted-sum $\langle w, a \rangle$

Operations on Vectors

Inner Product - Applications

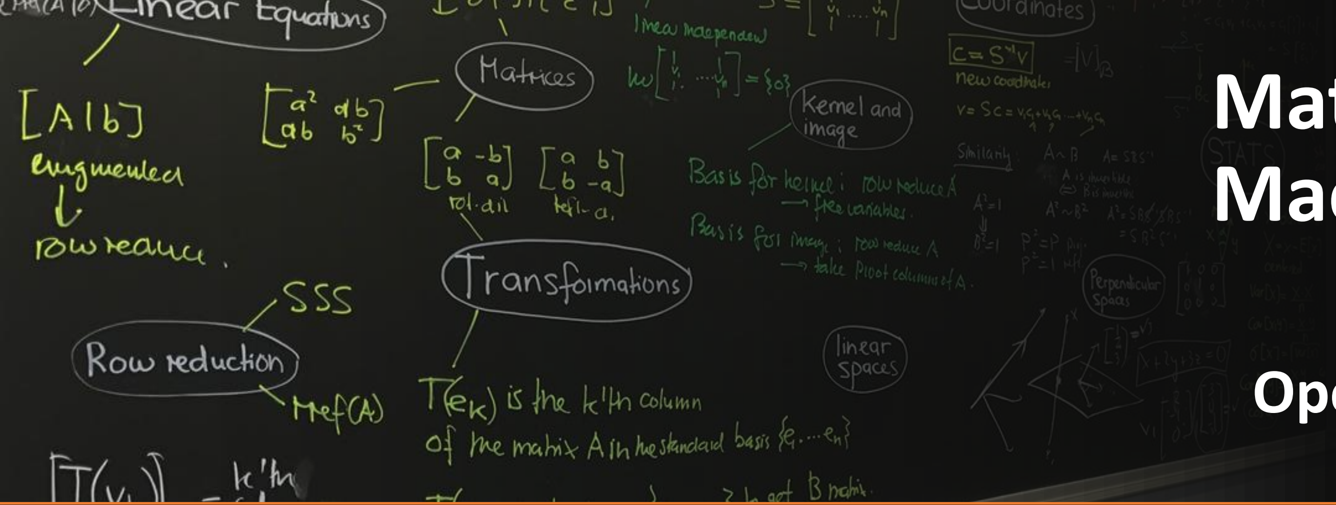
* $a \in R^{51}$

$w = [0\ 0\ 0 \ldots\ 1\ -1\ 0\ 0\ 0\ 0] \in R^{51}$

$\langle w, a \rangle = a_{20} - a_{31}$

* Price-quantity sum

$p \in R^n, \quad q \in R^n \qquad \langle p, q \rangle = p_1 q_1 + p_2 q_2 + \ldots + p_n q_n$



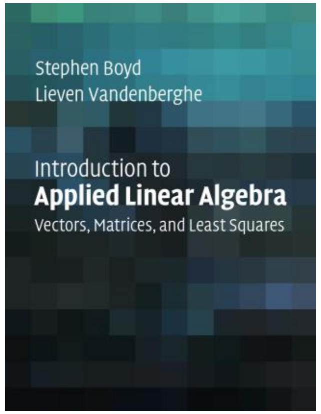
Mathematical Foundations for Machine Learning and Data Science

Operations on Vectors – Distance, Angle and Standard Deviation

LUMS
A Not-for-Profit University

Outline

- *Operations on Vectors*
  - *Norm*
  - *Distance*
  - *Angle*
  - *Standard Deviation*



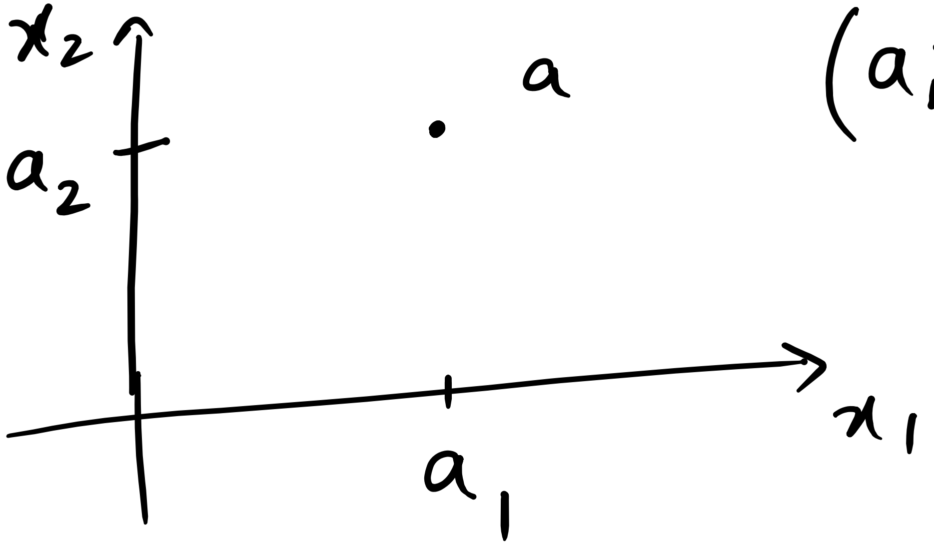
Chapter 3

Operations on Vectors

Norm Measure distance of a vector

* $\alpha \in R$ $\quad |\alpha|$

* R^2



$$(a_1^2 + a_2^2)^{1/2} = \|a\|_2$$

R^n; $\quad b \in R^n$

$$\|b\|_2 = (b_1^2 + b_2^2 + \dots + b_n^2)^{1/2}$$

Euclidean norm
2-norm

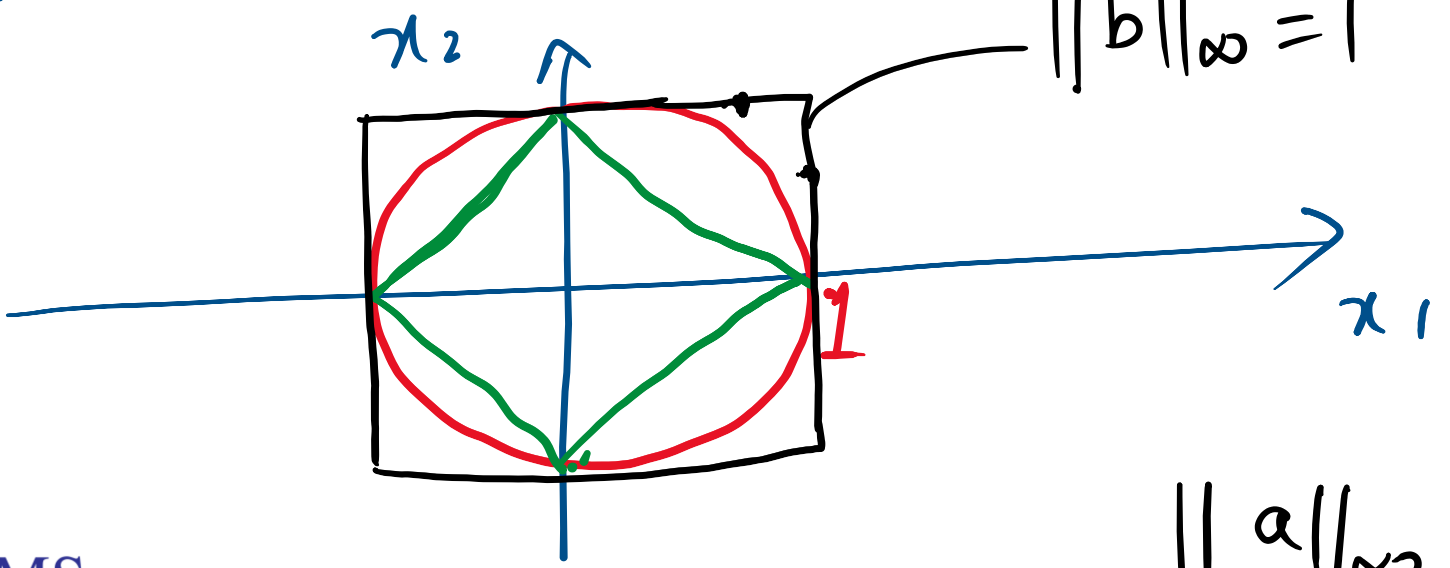
Operations on Vectors

Norm P- norm

For $p \geq 1$, $a \in \mathbb{R}^n$

$$\|a\|_p = \left(|a_1|^p + |a_2|^p + \dots + |a_n|^p\right)^{1/p}$$

$\underline{R^2}$



$\|b\|_\infty = 1$

$\|b\|_2 = 1$

$\|b\|_1 = 1$

$|b_1| + |b_2| = 1$

$$\|a\|_\infty = \max(|a_1|, |a_2| \dots |a_n|)$$

Operations on Vectors

Norm

$$a = \begin{bmatrix} 5 \\ -3 \\ -7 \\ 4 \\ 1 \end{bmatrix}$$

$$\|a\|_2 = (25 + 9 + 49 + 16 + 1)^{1/2} = (100)^{1/2} = 10$$

$$\|a\|_1 = 5 + 3 + 7 + 4 + 1 = 20$$

$$\|a\|_\infty = \max(5, 3, 7, 4, 1) = 7$$

$$\|a\|_p \geq \|a\|_q$$

$$p \leq q$$

Operations on Vectors

Norm $f : R^n \longrightarrow R$

* Non-negativity $\| \ \| \geq 0$
* Definiteness $\|a\| = 0 \iff a = 0$
* Homogenity $\|\alpha a\| = |\alpha| \, \|a\|$
* Triangle Inequality $\|a+b\| \leq \|a\| + \|b\|$

$\|a\| \longrightarrow$ 2-norm

Operations on Vectors

Norm

* $\|a\|^2 = a_1^2 + a_2^2 + \dots + a_n^2$

$= \langle a, a \rangle$

* RMS value $\left(\dfrac{a_1^2 + a_2^2 + \dots + a_n^2}{n}\right)^{1/2} = \dfrac{\|a\|}{\sqrt{n}}$

* Unit vector $\|e_i\| = 1$ $\quad a, \; b = \dfrac{a}{\|a\|}$

Operations on Vectors

Distance $a, b \in \mathbf{R}^n$

$$\mathrm{dist}(a,b) = \|a - b\|$$

Interpretation: *Compare two vectors; close or far*

Example:

$$a = \begin{bmatrix} 1.8 \\ 2.0 \\ -3.7 \\ 4.7 \end{bmatrix} \qquad b = \begin{bmatrix} 0.6 \\ 2.1 \\ 1.9 \\ -1.4 \end{bmatrix} \qquad c = \begin{bmatrix} 2.0 \\ 1.9 \\ -4.0 \\ 4.6 \end{bmatrix}$$

$\|a - b\| = 8.368$

$\|a - c\| = 0.387$

$\|b - c\| = 8.533$

Operations on Vectors

Application Examples

Determine similarity between objects

$a, b \in R^n \qquad \|a - b\|$

Prediction error

$y \qquad \|y - \hat{y}\| = \|e\|$

$\hat{y}$

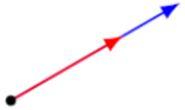
Operations on Vectors

Angle $a, b \in \mathbf{R}^n$

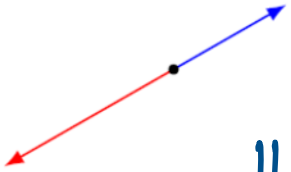
$$\theta = \angle(a, b) = \cos^{-1} \frac{a^T b}{\|a\| \|b\|}$$

$\theta \in [0, \pi]$

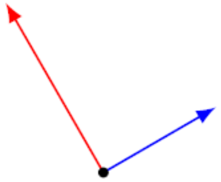
* Aligned $\theta = 0^\circ$



* Anti-aligned $\theta = 180^\circ$



* Orthogonal $\theta = 90^\circ$



$$\cos\theta = \frac{a^T b}{\|a\| \|b\|}$$

$$\|a\| \|b\| \cos\theta = a^T b$$

$$\|a\| \|b\| \geq |a^T b|$$

Cauchy-Schwarz Inequality

Operations on Vectors

Correlation coefficient $a, b \in \mathbf{R}^n$

$$\tilde{a} = a - \mathbf{avg}(a)\mathbf{1}, \qquad \tilde{b} = b - \mathbf{avg}(b)\mathbf{1}$$

(Demeaned vectors)

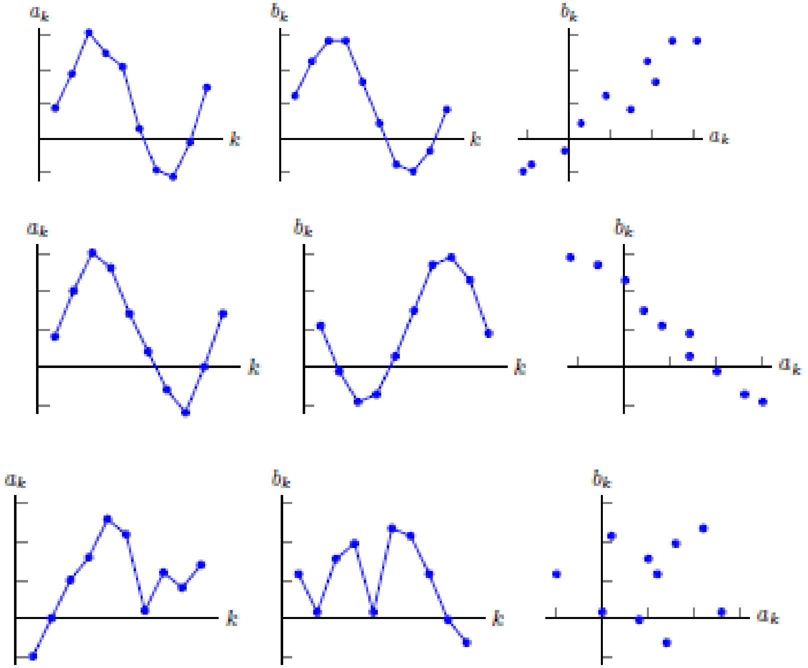
$$\rho = \frac{\tilde{a}^T \tilde{b}}{\|\tilde{a}\| \, \|\tilde{b}\|} = \cos \angle (\tilde{a}, \tilde{b})$$

$-1 \le \rho \le 1$

$\rho \approx 1$

$\rho = -1$

$\rho = 0$



Interpretation:

Operations on Vectors

Standard Deviation

$$\mathbf{std}(x) = \sqrt{\frac{(x_1 - \mathbf{avg}(x))^2 + \cdots + (x_n - \mathbf{avg}(x))^2}{n}}$$

Interpretation: a typical amount by which entries of a vector deviate around average value

$y = x - \text{avg}(x)\,\mathbf{1}$

$$\text{std}(x) = \sqrt{\frac{y_1^2 + y_2^2 + \cdots + y_n^2}{n}} = \frac{\|y\|}{\sqrt{n}}$$

RMS(y)

Operations on Vectors

Standard Deviation

Properties $\quad x \in R^n \quad \alpha \in R$

* $std(x + \alpha 1) = std(x)$

* $std(\alpha x) = |\alpha| \, std(x)$

Relationship between std, avg and rms

$$rms(x)^2 = avg(x)^2 + std(x)^2$$

Error = BIAS + VARIANCE

Operations on Vectors

Standard Deviation

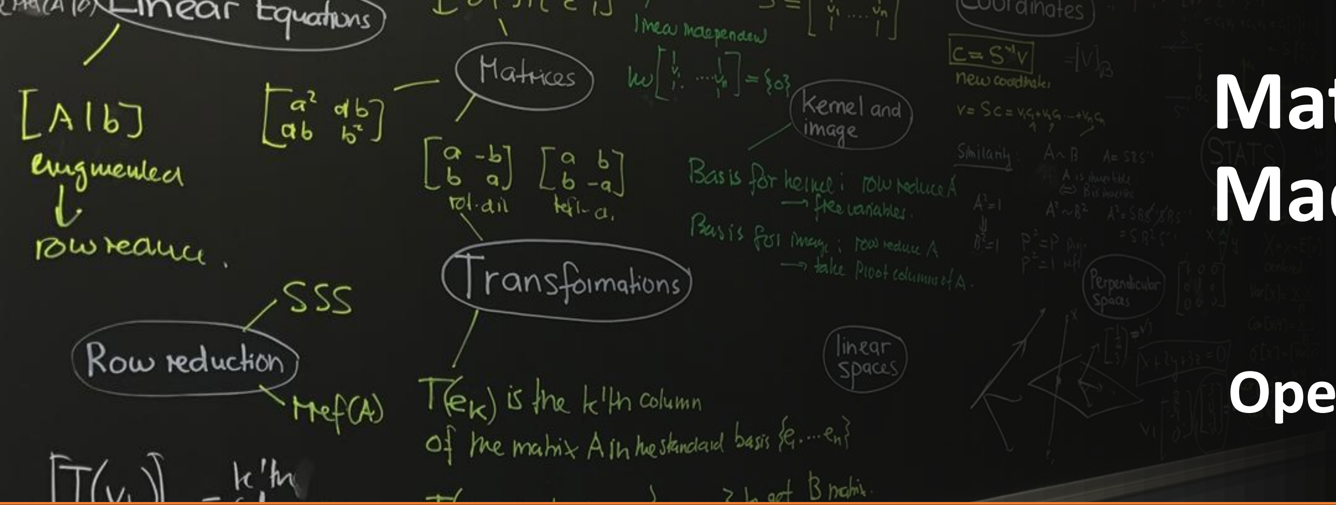
Concept of Standardization

$x \in R^n$

* $y = x - \text{avg}(x)$; $\text{avg}(y) = 0$

* $z = \dfrac{x - \text{avg}(x)}{\text{std}(x)}$; $\left.\begin{array}{l}\text{avg}(z) = 0 \\ \text{std}(z) = 1\end{array}\right]$ Zero mean std-dev unity

* μ, σ $\quad z = \dfrac{x - \mu}{\sigma}$



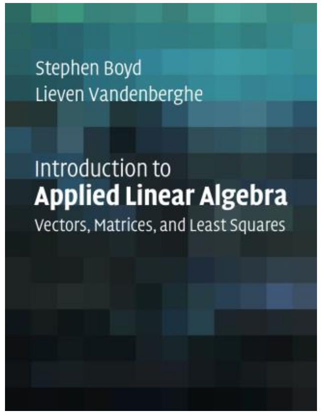
Mathematical Foundations for Machine Learning and Data Science

Operations on Vectors – Linear Independence, Basis and Orthonormal Vectors

LUMS
A Not-for-Profit University

Outline

- *Operations on Vectors*
  - *Span*
  - *Linear Independence*
  - *Basis*
  - *Orthonormal Vectors*



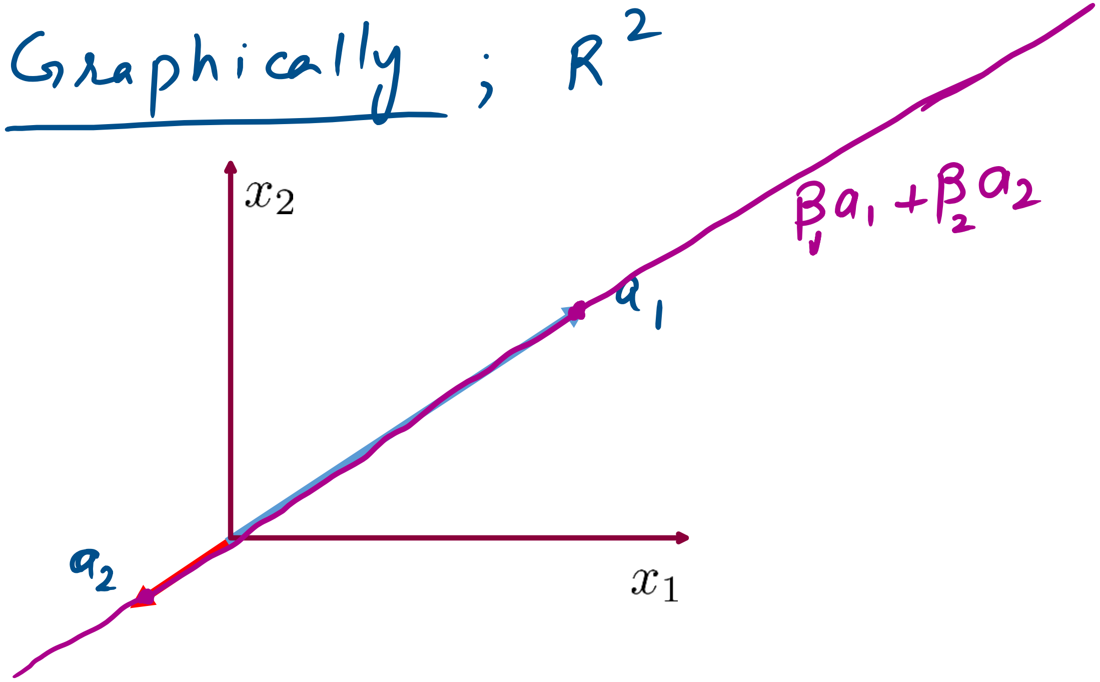
Chapter 5

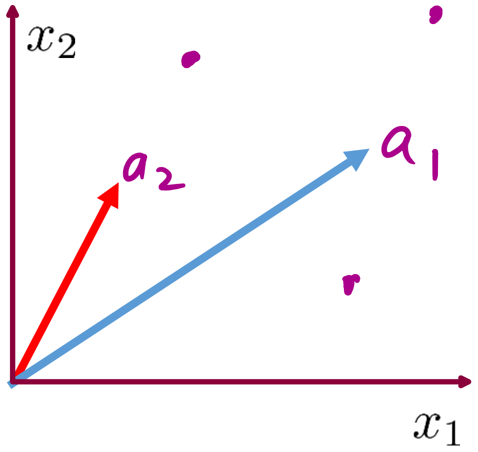
Span

Linear Combination (Review) $a_1, a_2 \in R^n \quad \beta_1, \beta_2 \in R$

$$\beta_1 a_1 + \beta_2 a_2 \in R^n$$

Graphically ; R^2





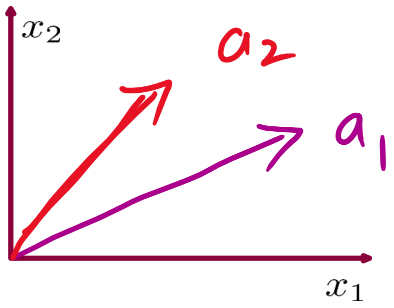
Span

Definition: $a_1, a_2 \ldots, a_K \in R^n$

$\beta_1, \beta_2 \ldots \beta_K \in R$

$$\text{span}(a_1, a_2, \ldots, a_K) = \{\beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_K a_K\}$$

* Span of a collection of vectors is a set of all linear combination of vectors



R^3; 2 vectors

* Plane

3 vectors $\rightarrow R^3$

Examples:

LUMS
A Not-for-Profit University

Linear Independence

Definition:

$a_1, a_2 \ldots a_k \in R^n$ (Linearly independent)

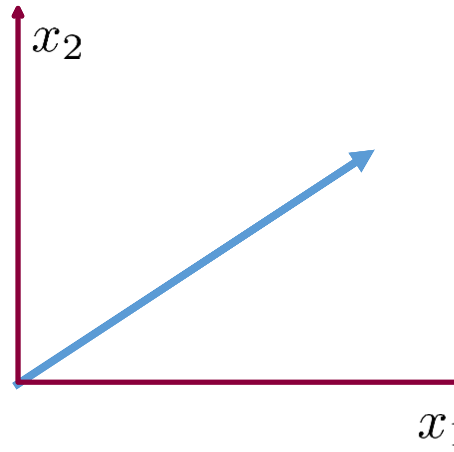
$\beta_1, \beta_2 \ldots \beta_k \in R$

$$\boxed{\beta_1 a_1 + \beta_2 a_2 + \cdots \beta_k a_k = 0}$$

iff $\beta_1 = \beta_2 = \cdots \beta_k = 0$

$$a_1 = -\left(\frac{\beta_2}{\beta_1} a_2 + \cdots \frac{\beta_k}{\beta_1} a_k \right)$$

* a_1 * 0-vector in any list (linearly dependent)



Examples:

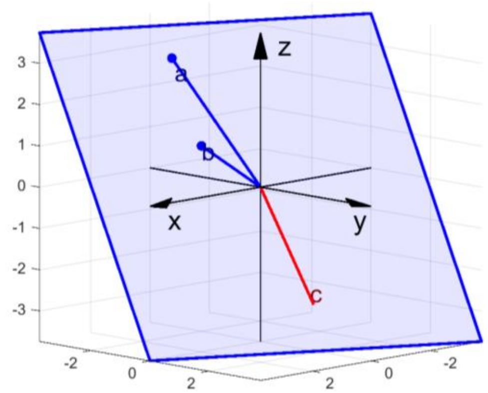
Linear Independence

Examples

$$a = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \qquad c = \begin{bmatrix} -0.3 \\ 1.5 \\ -2.7 \end{bmatrix}$$

$$\underline{1.2a - 0.9b + c = 0}$$

linearly dependent



Examples:

LUMS
A Not-for-Profit University

Linear Independence

Implications:

* List is linearly dependent
  → Superset of list linearly dependent

* List – linearly independent
  → subset of list is also linearly independent

* $b \in R^n$ $\quad b = \beta_1 a_1 + \beta_2 a_2 + \dots + \beta_k a_k$

  $a_1, a_2 \dots a_k \in R^n$ $\quad \beta_1, \beta_2 \dots \beta_k$ (Unique)

Examples:

Basis

Independence-Dimension Inequality

(Connection with linear independence):

If $a_1, a_2, \ldots, a_k \in R^n$ (Linearly Independent)

$k \leq n$

* A list of vectors in R^n can have at most n elements

Basis

Def: A collection of n linearly independent n-vectors is called basis.

If $a_1, a_2 \ldots a_n \in \mathbb{R}^n$ linearly independent

* Any vector $b \in \mathbb{R}^n$ can be represented (uniquely) as

$$b = \beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_n a_n$$

$$b = b_1 e_1 + b_2 e_2 + \cdots + b_n e_n$$

LUMS
A Not-for-Profit University

Orthonormal Vectors

* Orthogonal vectors

a_1, a_2 are orthogonal if

$\langle a_1, a_2 \rangle = a_1^T a_2 = 0 \qquad a_1 \perp a_2$

* A list of vectors $a_1, a_2, \ldots a_k$ are mutually orthogonal if

$\langle a_i, a_j \rangle = 0 \qquad i \neq j$

Orthonormal Vectors

$$\underbrace{\text{orthogonal} + \text{unit norm}}_{\text{orthonormal}}$$

* $a_1, a_2, \ldots, a_k$

$$\langle a_i, a_j \rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

* orthonormal vectors are linearly independent

Orthonormal Vectors

Linear Combination of Orthonormal Vectors

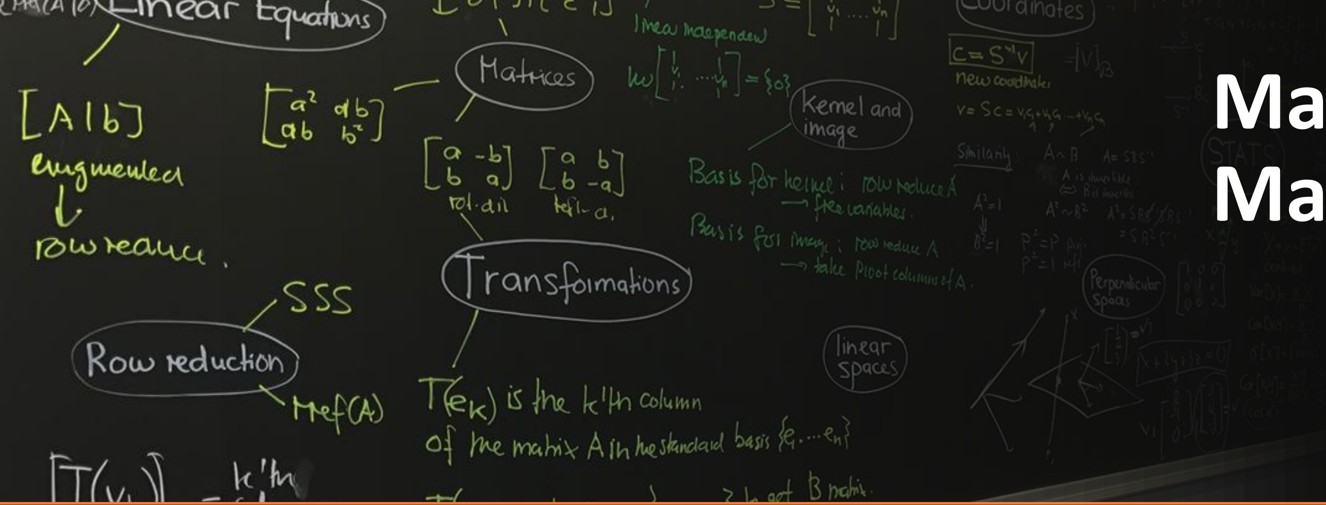
$$b = \beta_1 a_1 + \beta_2 a_2 + \dots \beta_K a_K$$

* How to find β_i ?

$$\langle b, a_i \rangle = \beta_i = \langle \beta_1 a_1 + \beta_2 a_2 + \dots \beta_K a_K , a_i \rangle$$

$$= \beta_i a_i^T a_i = \beta_i$$

$K = n$; $a_1, a_2, \dots a_n$ (Orthonormal Basis)



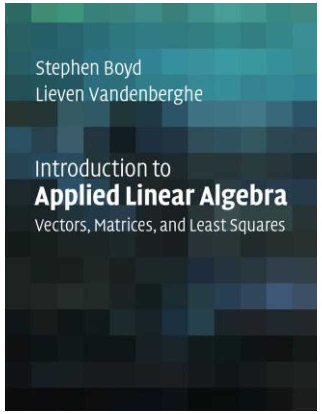
Mathematical Foundations for Machine Learning and Data Science

Operations on Vectors – Gram-Schmidt Orthogonalization Algorithm

LUMS
A Not-for-Profit University

Outline

- *Gram-Schmidt Orthogonalization*
  - *Overview*
  - *Algorithm*
  - *Example*



Chapter 5

LUMS
A Not-for-Profit University

Gram-Schmidt Orthogonalization

Overview: **Input:** Given a list of vectors $a_1, a_2, \ldots, a_k \in R^n$

If linearly independent

Orthonormal vectors $q_1, q_2, \ldots, q_k \in R^n$

* For each i, a_i is a linear combination of $q_1, q_2, \ldots, q_i$ orthonormal vectors

* If $a_1, a_2 \ldots a_j$ linearly independent

$a_1, a_2 \ldots a_j, a_{j+1}$ linearly dependent

Gram-Schmidt Orthogonalization

Algorithm:

Input: Given a list of vectors $a_1, a_2, \ldots, a_k \in R^n$

Output:

1. Test linear independence
2. Orthonormal vectors $q_1, q_2, \ldots, q_k \in R^n$

Algorithm

for $i = 1, 2, \ldots, k$

$$\tilde{q}_i = a_i - \sum_{j=1}^{i-1} (q_j^T a_i)\, q_j$$

Orthogonalization

$$q_i = \frac{\tilde{q}_i}{\|\tilde{q}_i\|}$$

Normalization

end

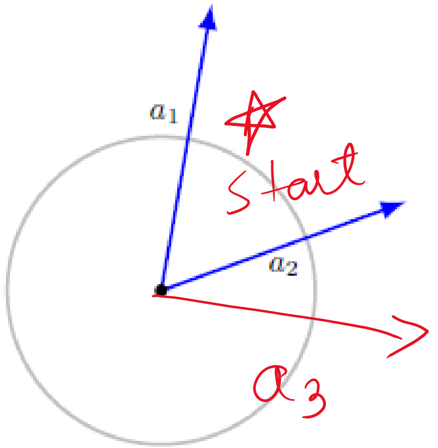
Example for $k = 3$

- i=1 $\quad \tilde{q}_1 = a_1$
- i=2 $\quad \tilde{q}_2 = a_2 - (q_1^T a_2)\, q_1$
- i=3 $\quad \tilde{q}_3 = a_3 - \left((q_1^T a_3)\, q_1 + (q_2^T a_3)\, q_2 \right)$

If $\tilde{q}_i = 0$, we detect linear dependence.

Gram-Schmidt Orthogonalization

Interpretation

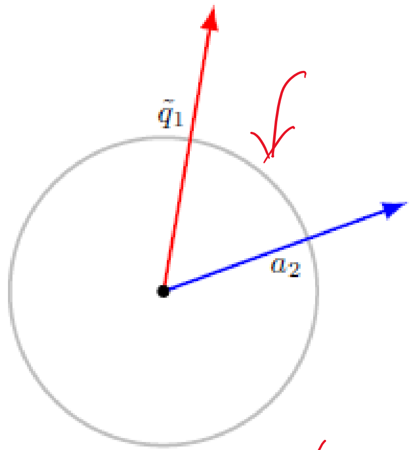


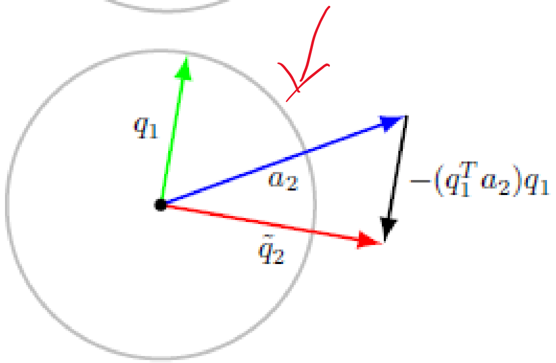
$$\tilde{q}_1 = a_1$$

$$\tilde{q}_2 = a_2 - (q_1^T a_2)\, q_1$$

Orthogonalization

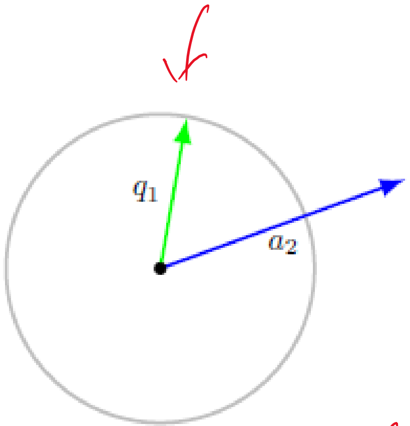
$$\tilde{q}_i = a_i - \sum_{j=1}^{i-1} (q_j^T a_i)\, q_j$$

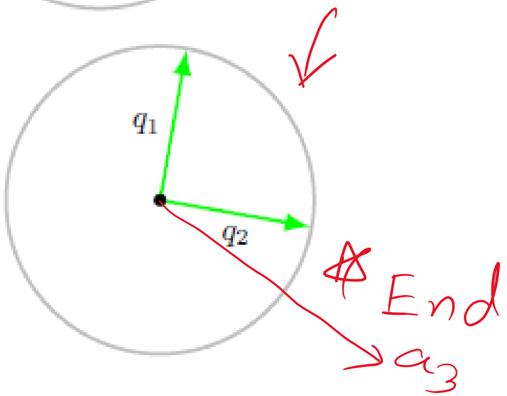




Normalization

$$q_i = \frac{\tilde{q}_i}{\|\tilde{q}_i\|}$$





$\tilde{q}_3 = a_3 - (a_3)$
$= 0$

Iteration 1

Iteration 2

Gram-Schmidt Orthogonalization

Example: $a_1 = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \quad a_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

$$\tilde{q}_i = a_i - \sum_{j=1}^{i-1} (q_j^T a_i)\, q_j$$

$$q_i = \frac{\tilde{q}_i}{\|\tilde{q}_i\|}$$

$\underline{i=1}$

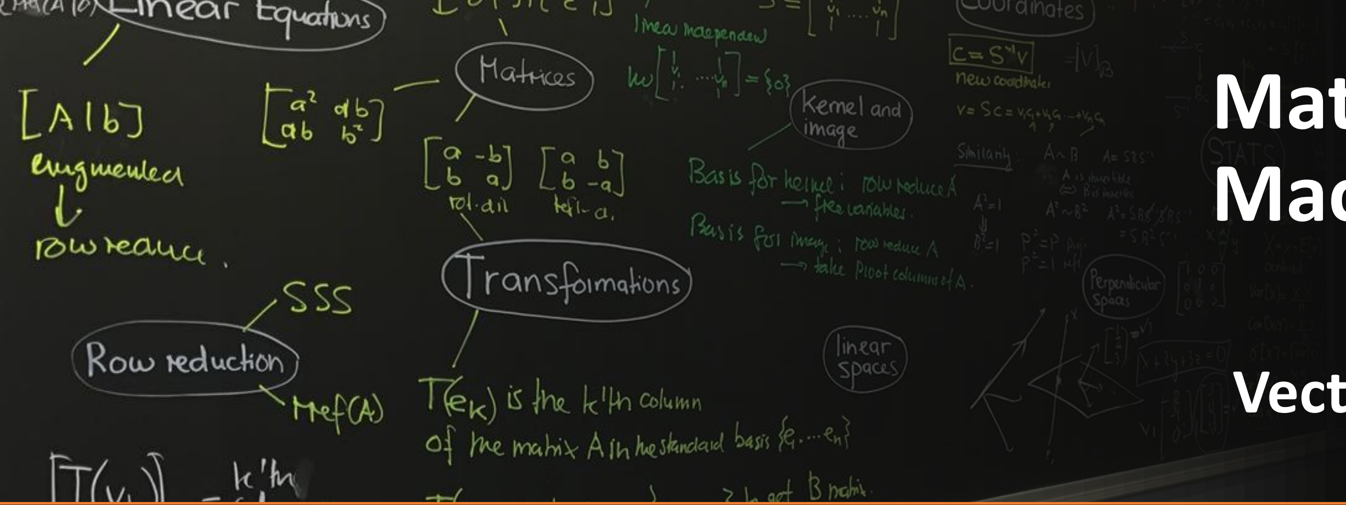
$\tilde{q}_1 = a_1 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \quad q_1 = \frac{\tilde{q}_1}{\|\tilde{q}_1\|} = \frac{1}{5}\begin{bmatrix} 4 \\ 3 \end{bmatrix}$

$\underline{i=2}$

$\tilde{q}_2 = a_2 - (q_1^T a_2)\, q_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix} + \frac{1}{5}\begin{bmatrix} 4 \\ 3 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} -6 \\ 8 \end{bmatrix}$

$\frac{1}{5}(-8+3) = -1$

$q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|} = \frac{1}{5}\begin{bmatrix} -3 \\ 4 \end{bmatrix}$



Mathematical Foundations for Machine Learning and Data Science

Vector Spaces and Subspaces – Linear Algebra

LUMS
A Not-for-Profit University

Outline

- *Vector Spaces*
- *Subspaces*
- *Dimension of subspace*

Vector Space

Definition

Let V be a set of vectors.

Two operators:

($+$) Addition of vectors
($\cdot$) multiplication with scalar

For $u, v, w \in V$ and $a, b \in \mathbf{R}$, following axioms hold

1. $u + v \in V$
2. $u + v = v + u$
3. $(u + v) + w = v + (u + w)$
4. $0 \in V : \; v + 0 = v$
5. $\exists -u \in V : \; u + (-u) = 0$
6. $a\,u \in V$
7. $a\,(u + v) = a\,u \, + \, a\,v$
8. $(a + b)\,u = a\,u \, + \, b\,u$
9. $(ab)\,u = a\,(b\,u)$
10. $(1)\,u = u$

Subspace

Definition

- V is a vector space
- W is a subset of V
- If W is also a vector space under the addition and scalar multiplication that is defined on V

  then W is a subspace

Alternatively

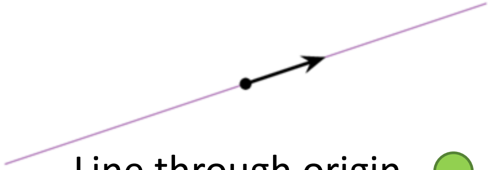
W, a subset of V, is a subspace if

- $0 \in W$ Non-emptiness
- For $u, v \in W$, $u + v \in W$ Closure under addition
- For $a \in R$ and $u \in W$, $au \in W$ Closure under multiplication

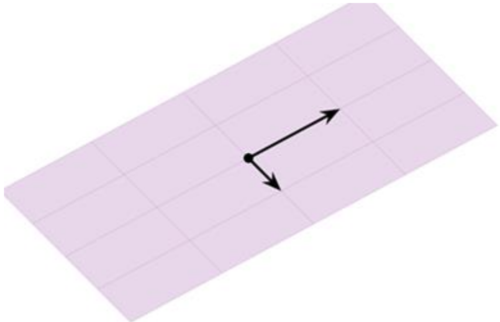
- Every vector space V has at least two subspaces. Namely, V itself and $\{0\}$ (the zero subspace).

Subspace

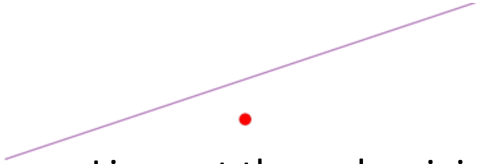
Examples (Subspace or not):



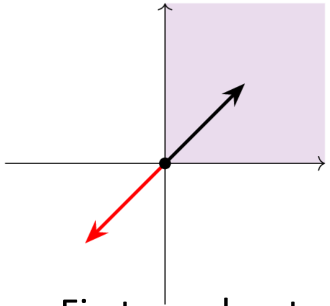
Line through origin



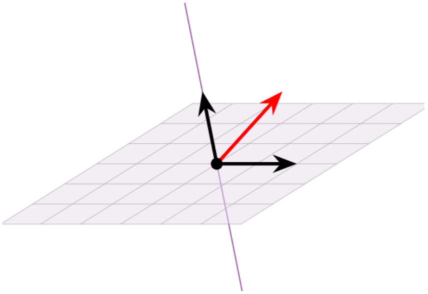
Plane through origin



Line not through origin



First quadrant



Line union plane

Subspace

Examples:

(a) Let W be the set of all points, (x, y), from $\mathbb{R}^2$ in which $x \geq 0$. Is this a subspace of $\mathbb{R}^2$?

(b) Let W be the set of all points from $\mathbb{R}^3$ of the form $(0, x_2, x_3)$. Is this a subspace of $\mathbb{R}^3$?

Dimension of Subspace

The number of vectors in any basis of V is called the ***dimension*** of V.

Expressed as dim(V)

Examples

- line through origin in $\mathbf{R}^4$ dim(V) = 1

- plane through origin in $\mathbf{R}^3$ dim(V) = 2