Linear Algebra & Convex Optimization – Lecture 1

Dr. Viswanath Gopalakrishnan
IIIT-B

Course Outline

Lecture Topics (14 Lectures) :

- Vector Operations,, Norms, Distances
- Clustering, Linear Dependence, Basis, Orthogonality
- Matrices, Matrix -Vector Product, Matrix Applications
- Matrix Inverses , Solving Linear Equations, Projection Spaces
- Least Squares, Data Fitting, Classification
- Eigen Analysis, PCA, Positive/Negative Definiteness
- Singular Value Decomposition and Applications
- Introduction to Functions, Derivatives and Matrix Calculus
- Convex Functions and Optimization Problems
- Optimality Criteria, Equivalent Convex Problems
- Lagrange Duality, Complementary Slackness, KKT conditions
- Constrained Optimization , Application to PCA
- Un-constrained Optimization, Gradient Descent Methods

Tutorials (2 Lectures):

Python Exercises (Vectors, Matrices, Least Squares)

Texts:

1. Introduction to Applied Linear Algebra, *Stephen Boyd & Lieven Vandenberghe*
2. Convex Optimization, *Stephen Boyd*

Evaluation Scheme – Part 1

Total (50)	
25 Marks	Mid-term Exam (Multiple Choice, 25 Questions of 1 mark each)
10 Marks	Assignment 1 (Pen and Paper)
10 Marks	Assignment 2 (Pen and Paper)
5 Marks	Attendance + Class Participation(Polls)

Outline

- Vector Representation
- Vector Operations
- Norms, Distances
- Vector Standardization

Textbook: Introduction to Applied Linear Algebra, S. Boyd: Chapters 1,2,3.

Vector Representation : Algebraic

Vector : Ordered set of numbers

$$\begin{bmatrix} -1.1 \\ 0.0 \\ 3.6 \\ -7.2 \end{bmatrix} \quad \text{or} \quad \begin{pmatrix} -1.1 \\ 0.0 \\ 3.6 \\ -7.2 \end{pmatrix} \quad \text{or} \quad (-1.1, 0.0, 3.6, -7.2).$$

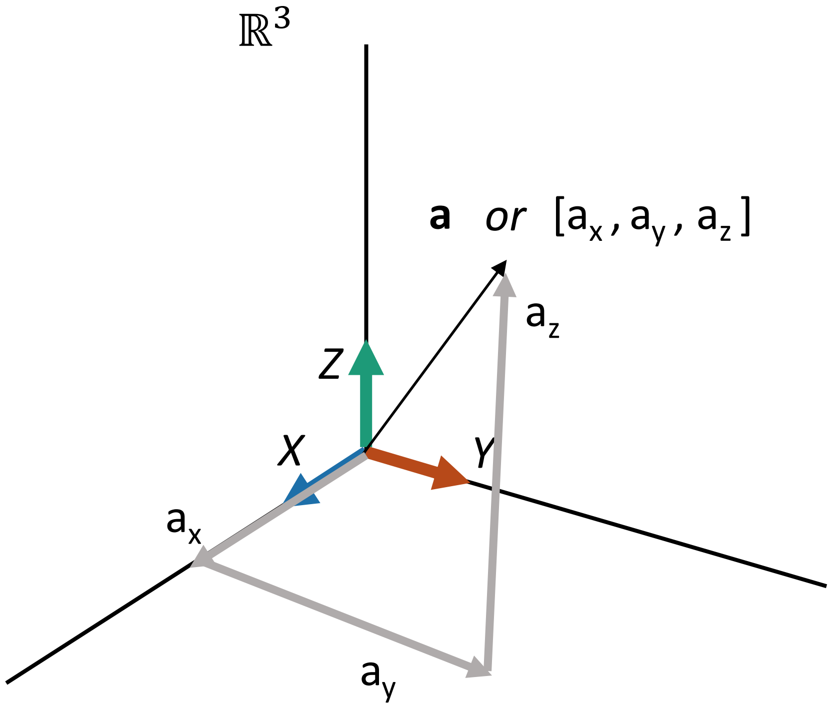
Indexing:

n dimensional vector indexed from $\mathrm{i} = 1$ to $\mathrm{i} = n$

elements in a single vector : $a = [a_1, a_2 \dots, a_n]$

elements in a group of vectors: j th element in i th vector : $(a_\mathrm{i})_j$

Vector Representation : Geometric



$a \in \mathbb{R}^3$

[a_x , a_y , a_z] ➡ Point in 3- dimensional Space /
3-dimensional Vector / Data

Vector Representations

Stacked Vectors :

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

$b \in \mathbb{R}^m$

$\mathrm{c} \in \mathbb{R}^n$

$\mathrm{d} \in \mathbb{R}^p$

$a \in \mathbb{R}^{(m+n+p)}$

$$\underset{m \times 1}{b} = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ \\ b_m \end{bmatrix} \qquad \underset{n \times 1}{c} = \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ \\ c_n \end{bmatrix} \qquad \underset{p \times 1}{d} = \begin{bmatrix} d_1 \\ d_2 \\ . \\ . \\ \\ d_p \end{bmatrix} \qquad \underset{(m+n+p) \times 1}{a} = \begin{bmatrix} b_1 \\ b_2 \\ . \\ b_m \\ c_1 \\ . \\ c_n \\ d_1 \\ . \\ d_p \end{bmatrix}$$

Transpose of Stacked Vectors :

$$a^T = [b^T \; c^T \; d^T]$$

$$\underset{1 \times (m+n+p)}{a^T} = [b_1 \; b_2 \; \dots b_m \; c_1 \dots c_n \, d_1 \dots d_p]$$

Slicing Vectors :

$$a_{r:s} = (a_r, \ldots, a_s) \qquad b = a_{1:m} \qquad c = a_{(m+1):(m+n)} \qquad d = a_{(m+n+1):(m+n+p)}$$

Unit Vectors

Unit Vector : All elements equal to zero except one element equal to one

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

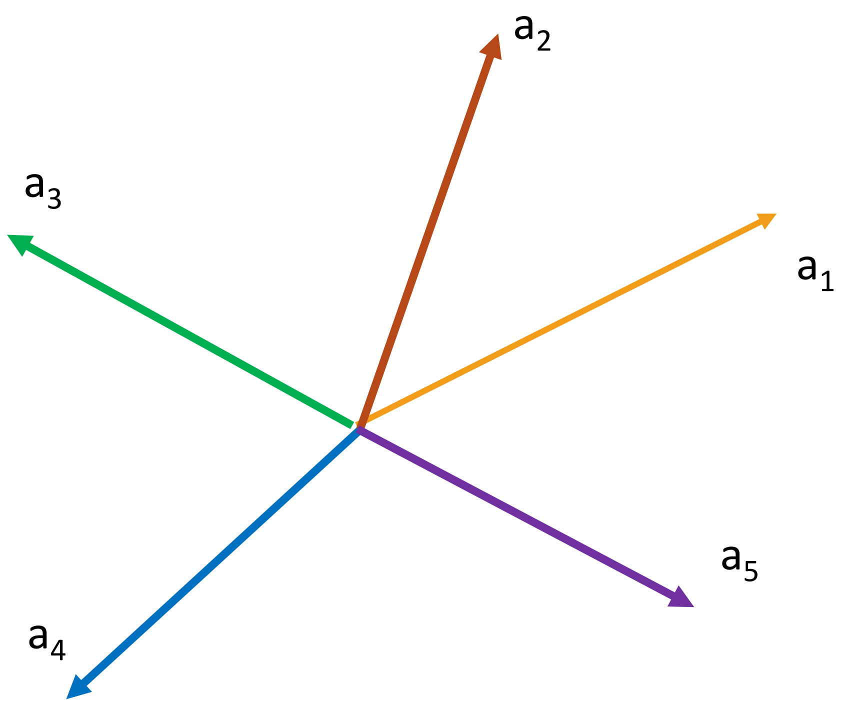
j th element in i th unit vector :

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

Examples : Data in Multi-Dimensional SPace

5 –Dimensional Space ($\mathbb{R}^5$)

(Do not bother to Imagine !
Look only Algebraic way)



$$a_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad a_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad a_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

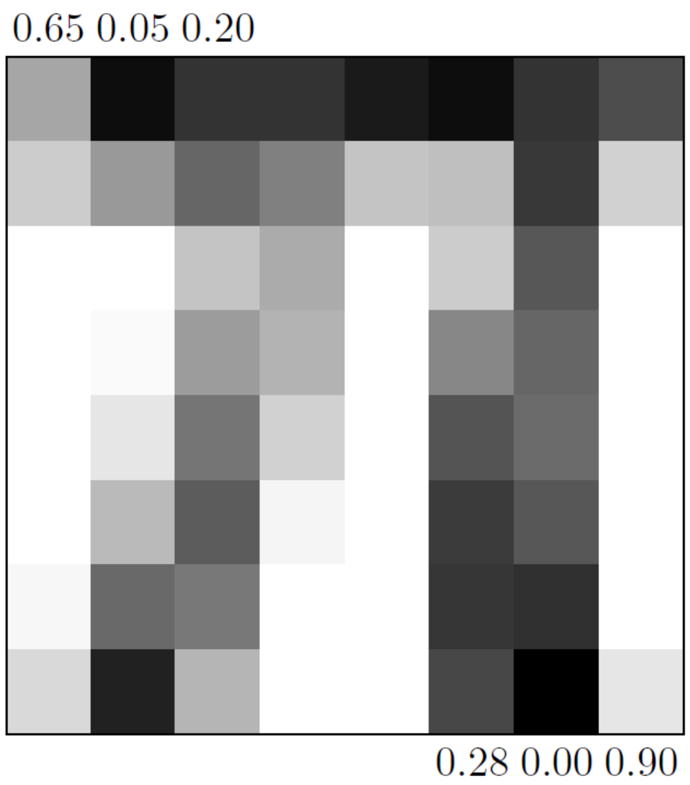
$$a_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad a_5 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Dimension	Meaning	Value
a_1	Height (ft)	6
a_2	Age	30
a_3	Weight (kg)	70
a_4	Waist-Size(in)	32
a_5	Gender	1

5- Dimensional Data

$$\mathbf{a} = [\,6\,,30\,,70\,,32,\,1\,] \Rightarrow \mathbf{a} = 6\mathbf{a_1} + 30\mathbf{a_2} + 70\mathbf{a_3} + 32\mathbf{a_4} + \mathbf{a_5}$$

Example : Image Representation



8×8 image

8×8 image can be represented as 64 dimensional vector

$x = [0.65\ \ 0.05\ 0.20\ \ \dots\dots\dots.0.28\ 0.00\ 0.90\ ]$

x is a point in 64-dimension space, $x \in \mathbb{R}^{64}$

Videos:

k frames in a video of resolution $m \times n$ can be represented as as $m \times n \times k$ vector

$x \in \mathbb{R}^{m\times n \times k}$

Poll : $x \in \mathbb{R}^{m\times n \times k}$ & $x \in \mathbb{R}^{m\times k \times n}$ represents the same video ?

Example : Document Representation

Word Count Histogram:

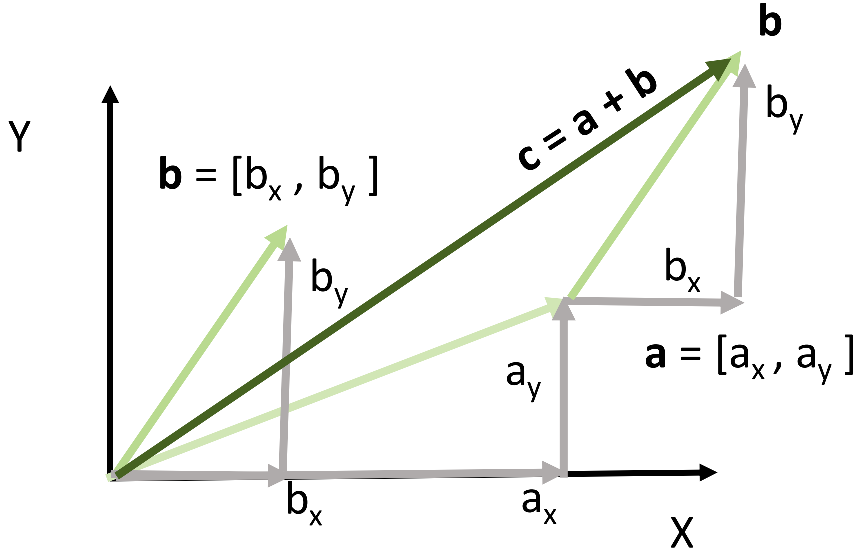
$$
\begin{array}{l}
\text{word} \\
\text{in} \\
\text{number} \\
\text{horse} \\
\text{the} \\
\text{document}
\end{array}
\qquad
\begin{bmatrix} 3 \\ 2 \\ 1 \\ 0 \\ 4 \\ 2 \end{bmatrix}
$$

Word –Preprocessing:

1) Tokenization
2) Normalization
3) Stop-word removal
4) Lemmatization

Vector Addition

How Vector Addition Looks Geometrically ?



For any 2D vectors a, b

$$\boldsymbol{a} + \boldsymbol{b} = [\, a_x + b_x \;\;,\;\; a_y + \; b_y \,]$$

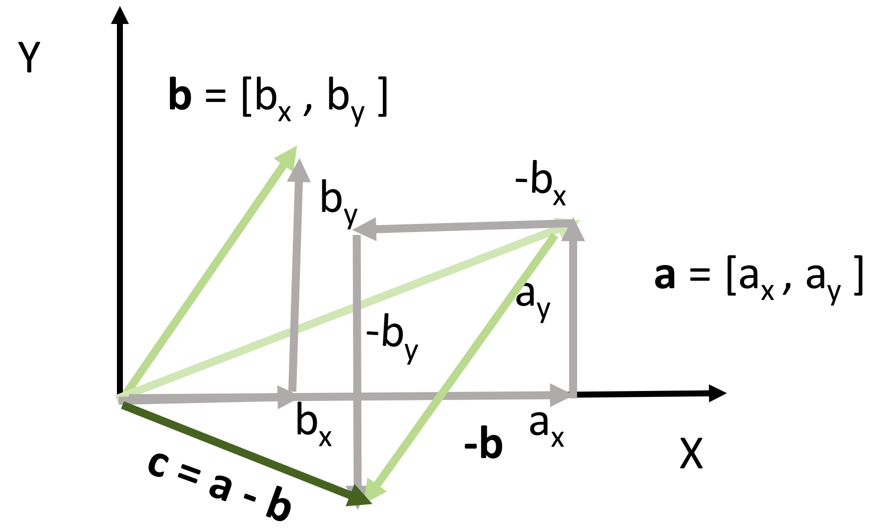
Algebraic Definition

For Multi-dimensional vectors $\boldsymbol{a} = [a_1, a_2, \cdots, a_n]$ & $\boldsymbol{b} = [b_1, b_2, \cdots, b_n\,]$

$$\boldsymbol{a} + \boldsymbol{b} = [a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n]$$

Vector Subtraction

How Vector Subtraction Looks Geometrically ?



For any 2D vectors a, b

$$\boldsymbol{a} - \boldsymbol{b} = [\, a_x - b_x \;\;,\;\; a_y - \, b_y \,]$$

Algebraic Definition

For Multi-dimensional vectors $\boldsymbol{a} = [a_1, a_2, \cdots, a_n]$ & $\boldsymbol{b} = [b_1, b_2, \cdots, b_n\,]$

$$\boldsymbol{a} - \boldsymbol{b} = [a_1 - b_1, a_2 - b_2, \dots, a_n - b_n]$$

Vector Addition -Examples

1. Document Addition

x : word histogram of Document 1

y : word histogram of Document 2

$x + y$: word histogram of combined Document (Using Same Dictionary

2. Audio Addition

x : recording of voice between t_1 to t_2

y : recording of music between t_1 to t_2

$x + y$: Recording of Music and Voice together in same time

3. Image Addition

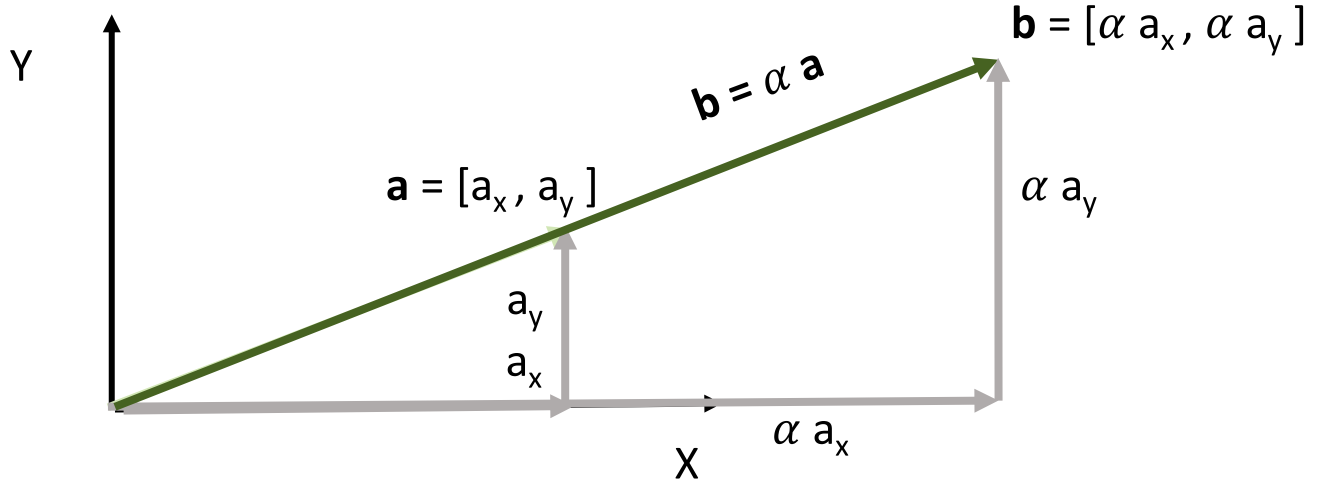
x : 1D representation of image I_1

y : 1D representation of image I_2

$x + y$: I_2 overlaid on I_1

Vector Scaling

$$(-2)\begin{bmatrix}1\\9\\6\end{bmatrix}=\begin{bmatrix}-2\\-18\\-12\end{bmatrix}$$



Linear Combination of unit vectors :

$$b=b_1e_1+\cdots+b_ne_n$$

$$\begin{bmatrix}-1\\3\\5\end{bmatrix}=(-1)\begin{bmatrix}1\\0\\0\end{bmatrix}+3\begin{bmatrix}0\\1\\0\end{bmatrix}+5\begin{bmatrix}0\\0\\1\end{bmatrix}$$

Special Linear Combinations:

Sum

$$\beta_1=\cdots=\beta_m=1 \quad \Rightarrow \quad a_1+\cdots+a_m$$

Average

$$\beta_1=\cdots=\beta_m=1/m \quad \Rightarrow \quad (1/m)(a_1+\cdots+a_m)$$

Affine

$$\beta_1+\cdots+\beta_m=1.$$

Linear Combination: Examples

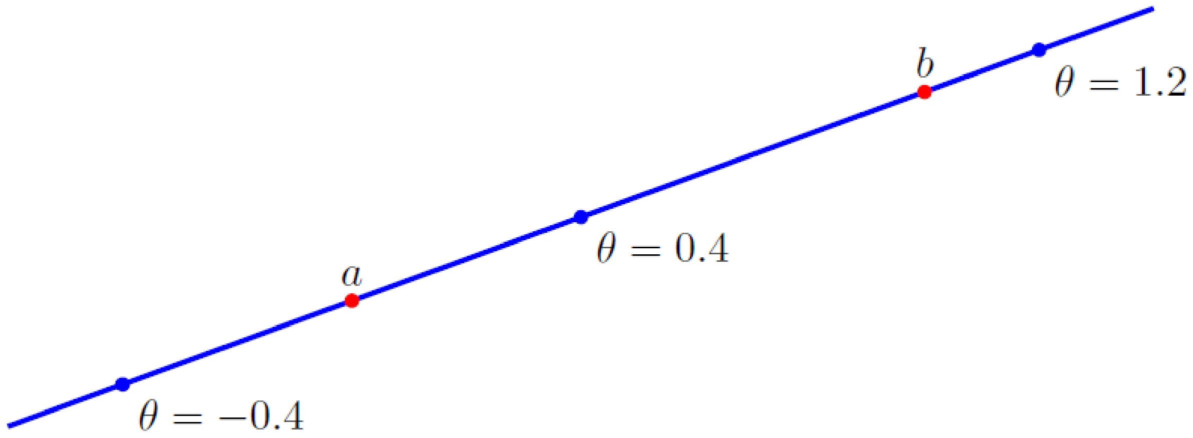
Audio Mixing:

$x_1, x_2 \ldots\ x_n$: n audio tracks $\beta_1, \beta_2 \ldots\ \beta_n$: mixing ratio

$\beta_1 x_1 + \beta_2 x_2 \ \ldots\ldots + \beta_n x_n$: mixed audio signal

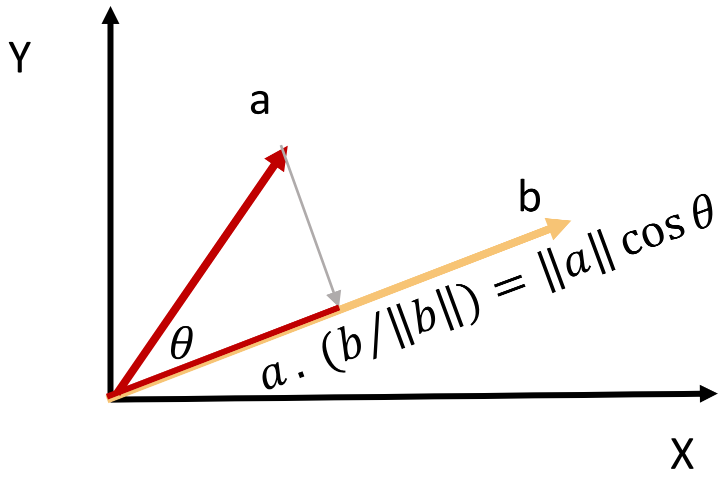
Line and Segment:

$$c = \theta a + (1 - \theta) b$$



Dot Product : Geometric Definition

$\boldsymbol{a}\,.\,\boldsymbol{u}$ **:** Amount of **a** in the direction of unit vector $u = b/\|b\|$



For any 2D vectors a, b

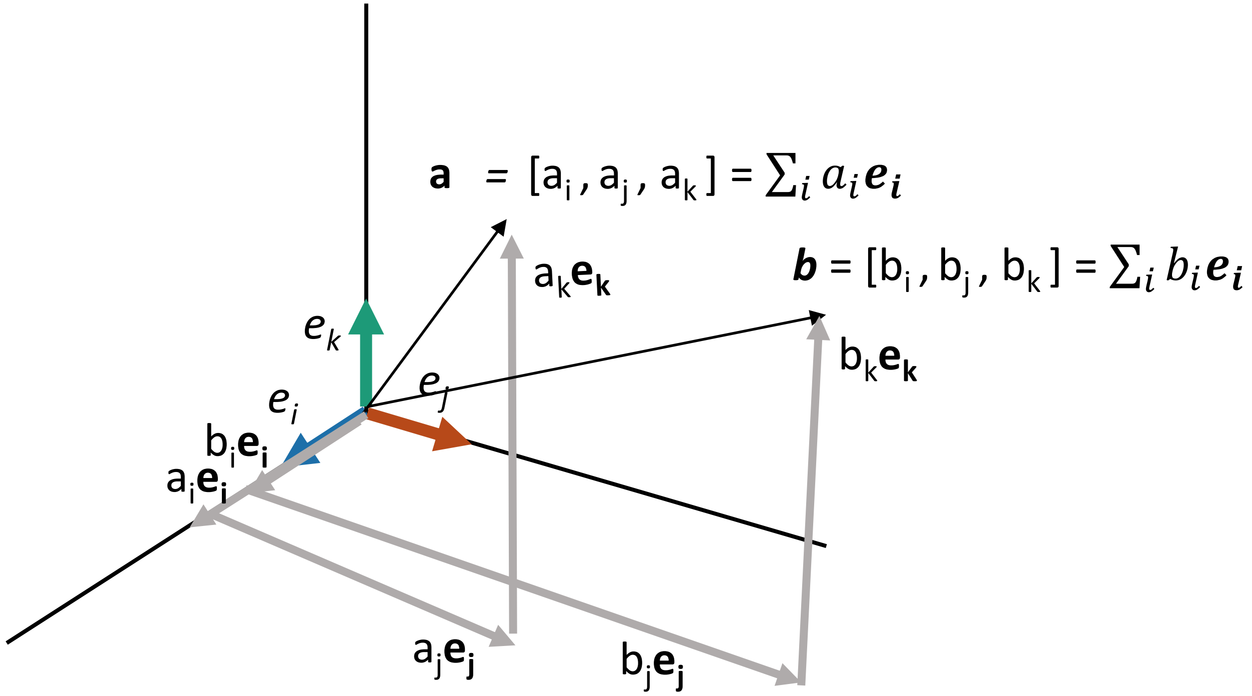
$$a\,.\,b = \|a\|\|b\|\cos\theta$$

$\boldsymbol{a}\,.\,\boldsymbol{b}$ **:** Amount of $\boldsymbol{a}$ in the direction of $\boldsymbol{b}$ scaled by the value of $\boldsymbol{b}$

OR

Amount of $\boldsymbol{b}$ in the direction of $\boldsymbol{a}$ scaled by the value of $\boldsymbol{a}$

Dot Product (Inner Product) : Algebraic Definition



[Amount of **a** in the direction of unit vector]

$$a\,.\,\boldsymbol{e_i} = a_i \qquad a\,.\,\boldsymbol{e_j} = a_j$$

$$a\,.\,b = a\,.\sum_i b_i \boldsymbol{e_i}$$

$$= \sum_i b_i (a \cdot \boldsymbol{e_i})$$

$$= \sum_i b_i a_i \quad = \sum_i a_i b_i$$

Poll: If a_1, b_1, a_2, b_2 are **any** 4 n-vectors, $a_1^T b_1 = a_2^T b_2$ implies the angle between a_1 & b_1 is same as a_2 & b_2 ?

No

Representations:

1. $a^T b$
2. $\langle a, b \rangle$
3. $\langle a | b \rangle$
4. $a.\,b$

Inner Product : Examples

1. Unit Vector : $e_i^T a = a_i$

2. Sum : $\mathbf{1}^T a = a_1 + a_2 + \cdots a_n$ where $\mathbf{1} = [1, 1 \cdot \cdots, 1]$

3. Average : $(\mathbf{1}/n)^T a = (a_1 + a_2 + \cdots a_n)/n$ where $\mathbf{1}/n = [1/\mathrm{n}, 1/\mathrm{n} \cdot \cdots, 1/\mathrm{n}]$

4. Sum of Squares : $a^T a = a_1^2 + a_2^2 + \cdots a_n^2$

5. Block Vectors:

$$a^T b = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}^T \begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix} = a_1^T b_1 + \cdots + a_k^T b_k$$

Inner Product : Applications

1. **Co-occurance**:

a : indicates membership, Set A

b : indicates membership, Set B

$a^T b$: # of elements in A ∩ B

$$a = (0, 1, 1, 1, 1, 1, 1) \quad b = (1, 0, 1, 0, 1, 0, 0) \quad a^T b = 2.$$

2. **Score:**

f : set of features of an object

w : weight of each feature

$w^T f$: score for the object based on the importance of each feature

Example: *Document Sentiment Analysis*

f : histogram from vocabulary of size n

w: $\{-1,0,1\}^n$

3. **Polynomial Evaluation:**

$$p(x) = c_1 + c_2 x + \cdots + c_{n-1} x^{n-2} + c_n x^{n-1}$$

c : coefficients of the polynomial

$$z = (1, t, t^2, \ldots, t^{n-1})$$

$$c^T z = p(t)$$