Mathematics and Statistics for Data Science
Lecture 1
Vectors and Norms: Fundamental Properties

Prof. Dr.-Ing. Paolo Mercorelli

October 22, 2020

A point in a n-dimensional space can be denoted by an n-tuple of real numbers.

$$A = (a_1, ..., a_i, ..., a_n)$$

A is called an n-dimensional vector. a_i are called scalars.
The set of all n-dimensional vectors is called the n-space or the vector space of n-tuples of real numbers. It is denoted by $\mathbb{R}^n$

Definitions 1

1. Equality: $A = B \Leftrightarrow a_i = b_i \ \forall i$
2. Addition: $A + B = C \Leftrightarrow a_i + b_i = c_i \ \forall i$
3. Multiplication by scalar $\alpha \in \mathbb{R}$: $B = \alpha A = A\alpha \Leftrightarrow b_i = \alpha a_i \ \forall i$

If $A = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$; $B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ Then:

$$A + B = \begin{bmatrix} 1+1 \\ 2+2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$\alpha B = \alpha \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \alpha = \begin{bmatrix} \alpha \\ 2\alpha \end{bmatrix}$$

Properties

1. Commutative: $A + B = B + A$
2. Associative: $A + (B + C) = (A + B) + C = (A + C) + B$
3. Distributive: $\alpha(A + B) = \alpha A + \alpha B$
4. Associative in case of multiplication by scalar: $\alpha(\beta A) = (\alpha\beta)A$
5. Distributive in case of multiplication by scalar: $(\alpha + \beta)A = \alpha A + \beta A$

Note that some operators on either sides of these equations do not have the same meanings. For example, the $+$ on the left of 5 denotes additions of real numbers while the $+$ on the right denotes vector addition.

Remark

1. *The zero vector* $0 = (0, ..., 0)$ *is the identity element of addition i.e.* $A + 0 = 0 + A = A \qquad \forall A$
2. *The vector* $(-1)A = -A$ *is called the negative of* A *i.e.* $A - A = 0$
3. $0A = 0$ *and* $1A = A$

$$\overline{AB} = A - B = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$\overline{A0} = A - 0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Definitions 2

Two vectors A and B are

1. in the same direction if $\exists\ \alpha > 0$ with $B = \alpha A$
2. in the opposite direction if $\exists\ \alpha < 0$ with $B = \alpha A$
3. parallel if $\exists\ \alpha \neq 0$ with $B = \alpha A$

LEUPHANA
UNIVERSITY OF LÜNEBURG

The dot (inner or scalar) product

The dot product between any vector $(A, B) \in \mathbb{R}^n$ is defined as

$$\underbrace{A \cdot B}_{Dot} = \sum_{i=1}^{n} a_i b_i$$

$\forall \quad A, B, C \in \mathbb{R}^n$ and scalar α we have

1. Commutative law: $A \cdot B = B \cdot A$
2. Distributive law: $A \cdot (B + C) = A \cdot B + B \cdot C$
3. Homogenity: $\alpha(A \cdot B) = (\alpha A) \cdot B = A \cdot (\alpha B) = A \cdot B\alpha$
4. Positivity: $A \cdot A \geq 0$ (equal applies if and only if $A = 0$)

Cauchy-Schwarz Inequality

Theorem:

$$(A \cdot B)^2 \leq (A \cdot A)(B \cdot B)$$

or

$$(\sum_{i=1}^{n} a_i b_i)^2 \leq (\sum_{i=1}^{n} a_i^2)(\sum_{i=1}^{n} b_i^2)$$

equal applies if and only if A and B are parallel

Proof: Assuming $A \neq 0$ and $B \neq 0$ (for $A = B = 0$ the Theorem holds). Consider $C = (xA - yB)$ with x and y scalars with x and y and we can choose

$x = B \cdot B$

$y = A \cdot B$ so that

$$\begin{aligned}
C \cdot C &= (xA - yB) \cdot (xA - yB) \\
&= x^2(A \cdot A) + y^2(B \cdot B) - 2xy(A \cdot B) \\
&= (B \cdot B)^2(A \cdot A) + (A \cdot B)^2(B \cdot B) - 2(B \cdot B)(A \cdot B)^2 \\
&= (B \cdot B)^2(A \cdot A) - (B \cdot B)(A \cdot B)^2 \\
&= (B \cdot B)[(B \cdot B)(A \cdot A) - (A \cdot B)^2] \geq 0 \\
&\rightarrow (A \cdot B)^2 \leq (B \cdot B)(A \cdot A)
\end{aligned}$$

LEUPHANA
UNIVERSITY OF LÜNEBURG

Remark

$$If\ C = 0 \rightarrow xA = yB$$
$$A = \frac{y}{x}B$$

$x, y \neq 0$ as $A, B \neq 0$

Length or norm of a vector

The length of norm $\|A\|$ of a vector $A \in \mathbb{R}^n$ is defined as:
$\|A\| \equiv \sqrt{A \cdot A} = \sqrt{\sum_{i=1}^{n} a_i^2}$
$\|A\| \rightarrow$ scalar, $(A \cdot B) \rightarrow$ Scalar

Properties:

1. $\|A\| \geq 0$, zero if and only if (iff) $a_i = 0 \ \forall i$
2. Homogenity: $\|cA\| = |c| \cdot \|A\|$, where c is a scalar
3. Triangular Inequality: $\|A + B\| \leq \|A\| + \|B\|$
   or $\|A + B\|^2 \leq (\|A\| + \|B\|)^2$

   equal applies if only if $A = 0$

Proof of property 2:

$$\|cA\|^2 = (c \cdot A) \cdot (c \cdot A) = c^2 \cdot (A \cdot A) = c^2 \|A\|^2 \quad \forall c \in \mathbb{R}^n$$

Proof of property 3:

$$\begin{aligned}
\|A+B\|^2 &= (A+B) \cdot (A+B) \\
&= A \cdot A + B \cdot B + A \cdot B + B \cdot A \\
&= A \cdot A + B \cdot B + 2A \cdot B \\
&= \|A\|^2 + \|B\|^2 + 2A \cdot B \\
(\|A\| + \|B\|)^2 &= \|A\|^2 + \|B\|^2 + 2\|A\| \|B\|
\end{aligned}$$

Thus $\|A+B\|^2 \leq (\|A\|+\|B\|)^2$ is true if $\|A\|\,\|B\| \geq A \cdot B$

Considering the Schwarz inequality:

$$(A \cdot B)^2 \leq (A \cdot A)(B \cdot B) = \|A\|^2 \|B\|^2$$

$\square$

Does triangular inequality **imply** Schwarz inequality?

Properties

- Schwarz inequality $\leftrightarrow$ Triangular inequality
- Schwarz inequality $\rightarrow$ Triangular inequality (already seen)
- Triangular inequality $\rightarrow$ Schwarz inequality?

$$\|A+B\|^2 \leq (\|A\| + \|B\|)^2$$

$$\|A+B\|^2 \leq (\|A\| + \|B\|)^2$$
$$\|A\|^2 + \|B\|^2 + 2A \cdot B \leq \|A\|^2 + \|B\|^2 + 2\,\|A\|\,\|B\|$$

Definition:
Two non-zero vectors $A, B \in \mathbb{R}^n$ are called perpendicular (or orthogonal)
if $A \cdot B = 0$.

Remark
$\|A + B\|^2 = \|A\|^2 + \|B\|^2 + 2A \cdot B$
$if\ A \perp B \rightarrow \|A + B\|^2 = \|A\|^2 + \|B\|^2$

LEUPHANA
UNIVERSITY OF LÜNEBURG

Definition

$\forall A, B$ non-zero vectors $\in \mathbb{R}^n$
The vector $\alpha B = \frac{B \cdot A}{\|B\|^2} B$ is called projection of A along B.
$\left(\alpha = \frac{B \cdot A}{\|B\|^2}\right)$
The vector $\beta A = \frac{B \cdot A}{\|A\|^2} A$ is called projection of B along A.
$\left(\beta = \frac{B \cdot A}{\|A\|^2}\right)$

Remark

Consider two vector $A, B \in \mathbb{R}^2$ *(non-zero vectors) making an angle* $0 < \theta < \frac{\pi}{2}$

Considering $A = C + \alpha B$ *with* $C \cdot B = 0$

$$\begin{aligned} B \cdot A = B \cdot (C + \alpha B) &\rightarrow B \cdot A = B \cdot C + \alpha B \cdot B \\ &\rightarrow B \cdot A = \alpha \left\| B \right\|^2 \end{aligned}$$

$\rightarrow \alpha = \frac{B \cdot A}{\left\| B \right\|^2}$

$$\|A\| \, cos\ \theta = \alpha \, \|B\| \to cos\ \theta = \frac{|\alpha| \, \|B\|}{\|A\|} = \frac{|A \cdot B| \, \|B\|}{\|B\|^2 \, \|A\|} = \frac{|A \cdot B|}{\|A\| \, \|B\|}$$

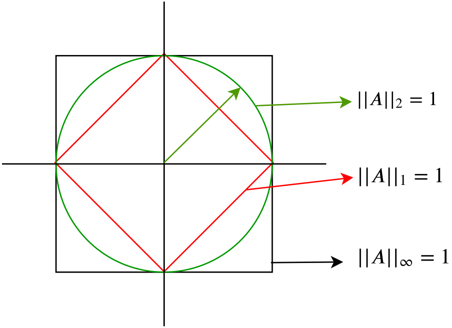
$\to cos\ \theta = \frac{A \cdot B}{\|A\| \|B\|}$

$$cos \theta = \frac{A \cdot B}{\|A\| \, \|B\|} \to \|A\| \, \|B\| \, cos\ \theta = A \cdot B$$

Considering $|cos\ \theta| \leq 1 \to \|A\| \, \|B\| \geq |A \cdot B|$.

Different types of norms

- L_2-norm: $\|A\|_2 = \sqrt{(A \cdot A)} = \sqrt{\sum_{i=1}^{n} a_i^2}$
- L_1-norm: $\|A\|_1 = \sum_{i=1}^{n} |a_i|$
- L_∞-norm: $\|A\|_\infty = \max_{a_i}(|a_i|)$



Seminorm: Change L so that $\|A\| \geq 0$ still holds, but now $\|A\| = 0$ does not have to mean that all $a_i = 0$.

- L_s-seminorm: $\|A\|_s = |\sum_{i=1}^{n} a_i|$