Lecture 4

Norms & Distances.

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

$$\|x\|_2 = \|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

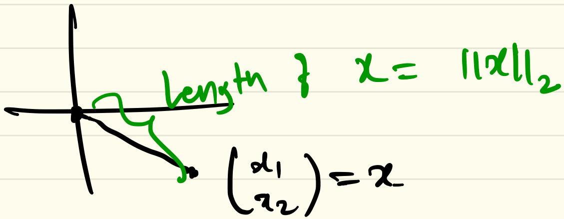
Euclidean-norm or 2-norm of $x \in \mathbb{R}^n$

[Generalization of absolute value for $x \in \mathbb{R}$

$\|x\|_2 = \sqrt{x^2} = |x|$]

For $n = 2$, geometric meaning of $\|x\|_2 = \sqrt{x_1^2 + x_2^2}$

= length of x.



Properties of 2- norm.

- Non-negative Homogeneity: $\forall \alpha \in \mathbb{R}, x \in \mathbb{R}^n$ ✓

$$\|\alpha x\|_2 = |\alpha| \, \|x\|_2$$

- Triangle inequality: $x, y \in \mathbb{R}^n$ (*)

$$\|x+y\|_2 \leq \|x\|_2 + \|y\|_2$$

- Non- negativity: $\forall x \in \mathbb{R}^n$; $\|x\|_2 \geq 0$ ✓

- Definiteness: $\|x\|_2 = 0 \Leftrightarrow x = 0$ ✓

$\Rightarrow$

$$\|x\|_2 = 0 \Rightarrow \|x\|_2^2 = 0 \Rightarrow x_1^2 + x_2^2 + \cdots + x_n^2 = 0$$

$$\Rightarrow x_i^2 = 0 \quad \forall i \in \{1, \cdots, n\}$$

$$\Rightarrow x_i = 0 \quad \forall i \in \{1, 2, \cdots, n\}$$

scalar equality

$$\Rightarrow x = 0$$

vector equality.

Triangle inequality consequence of C-S inequalities.

<u>observe:</u> $x, y \in \mathbb{R}^n$

$\|x+y\|_2 = ??$

$$\left[\begin{array}{l} \|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{x^T x} = \sqrt{\langle x, x\rangle} \\ \|x\|_2^2 = x^T x \end{array}\right] *$$

$$\|x+y\|_2^2 = (x+y)^T (x+y)$$

$$= x^T x + x^T y + y^T x + y^T y$$

$$= x^T x + 2x^T y + y^T y$$

$$\|x+y\|_2^2 = \|x\|_2^2 + 2x^T y + \|y\|_2^2$$

$$\Rightarrow \quad \|x+y\|_2 = \sqrt{\|x\|_2^2 + 2x^T y + \|y\|_2^2}$$

Cauchy - Schwartz inequality (C-S inequality)

$\forall\ x, y \in \mathbb{R}^n$

$$|x^T y| \leq \|x\|_2 \|y\|_2$$

$$|x^T y| \leq \left(\sqrt{x_1^2 + \cdots + x_n^2}\right)\left(\sqrt{y_1^2 + \cdots + y_n^2}\right)$$

Proof: If $x = 0$ or $y = 0$, then equality holds & nothing to prove.

$\alpha = \|x\|_2$ & $\beta = \|y\|_2$ & consider the vector

$$\alpha y - \beta x \in \mathbb{R}^n$$

Then

$$0 \leq \|\alpha y - \beta x\|_2^2 \quad \text{—} (*)$$

$$= (\alpha y - \beta x)^T (\alpha y - \beta x)$$

$$= \alpha^2 y^T y - \alpha \beta y^T x - \beta \alpha x^T y + \beta^2 x^T x$$

$$0 \le \alpha^2 \|y\|_2^2 + \beta^2 \|x\|_2^2 - 2\alpha\beta\, x^T y$$

$$0 \le \|x\|_2^2 \|y\|_2^2 + \|y\|_2^2 \|x\|_2^2 - 2\|x\|_2 \|y\|_2 \; x^T y$$

$$\Rightarrow \quad x^T y \le \|x\|_2 \|y\|_2 \qquad -(*)$$

If we replace x by $-x$

$$\Rightarrow \quad -x^T y \le \|x\|_2 \|y\|_2 \qquad -(**)$$

$$(*) \,\&\, (**) \Rightarrow \quad |x^T y| \le \|x\|_2 \|y\|_2$$

$\boxed{}$

Note: Equality holds true when $\alpha y = \beta x$

$$\Rightarrow \quad y = \beta/\alpha \; x \qquad \text{if } \alpha \neq 0$$

Proof of triangle inequality:

$$\|x+y\|_2^2 = \|x\|_2^2 + 2x^Ty + \|y\|_2^2$$

CS inequality

$$\leq \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2$$

$$= \left(\|x\|_2 + \|y\|_2\right)^2$$

$$\Rightarrow \quad \|x+y\|_2 \leq \|x\|_2 + \|y\|_2$$

Defn:
General norms:
Any function $\|\cdot\| : \mathbb{R}^n \longrightarrow \mathbb{R}$ which satisfies

i) non-negtive homogeneity

ii) triangle inequality

iii) non-negativity

iv) definiteness

Then $\|\cdot\|$ is called as norm.

Ex: 1) $\|\cdot\|_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ $\quad x \longmapsto \sqrt{x_1^2 + \cdots + x_n^2}$

2) $\|\cdot\|_p : \mathbb{R}^n \longrightarrow \mathbb{R}$ $\quad \|x\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}$

$p=2$ is a norm for $p \geq 1$.

$p=1$, absolute sum norm. $p=2$ spectral norm.

3) ∞-norm / sup norm / max norm.

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

- Root-mean square value. (RMS) value.

$$\mathrm{rms}(x) = \sqrt{\boxed{\frac{x_1^2 + \cdots + x_n^2}{n}}} = \frac{\|x\|_2}{\sqrt{n}}$$

mean square (x)

||

typical value of $|x_i|$

$$\mathrm{rms}(\mathbf{1}_n) = \sqrt{\frac{1 + \cdots + 1}{n}} = 1 \quad \forall n$$

$$\|\mathbf{1}_n\|_2 = \sqrt{n}$$

$\|\mathbf{1}_{10}\|_2 = \sqrt{10}$ $\quad \|\mathbf{1}_{100}\|_2 = \sqrt{100} = 10$

- <u>Chebyshev inequality:</u>

Let $x \in \mathbb{R}^n$. Let there be k entries in x such that

$|x_i| \geq a$ where $a > 0$.

$\Rightarrow \quad x_i^2 \geq a^2$

$\|x\|_2^2 = x_1^2 + x_2^2 + \cdots + x_n^2$

$\|x\|_2^2 \geq ka^2$

In terms of $\mathrm{rms}(x)$

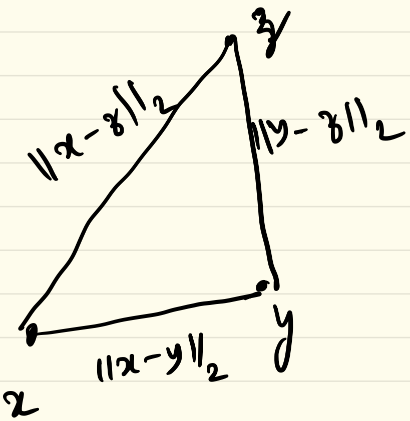
$$\boxed{\frac{k}{n} \leq \left(\frac{\mathrm{rms}(x)}{a}\right)^2} \quad \text{---} (*)$$

Distance:

Euclidean distance: $\forall x, y \in \mathbb{R}^n$

$$d(x,y) = \|x-y\|_2$$

Triangle inequality:



$$\text{dist}(x,y) = \|x-y\|_2$$

$$\text{dist}(y,z) = \|y-z\|_2$$

$$\text{dist}(x,z) = \|x-z\|_2$$

$$\text{dist}(x,z) \leq \text{dist}(x,y) + \text{dist}(y,z)$$

$$\|x-z\|_2 = \|x-y+y-z\|_2$$

$$= \|(x-y)+(y-z)\|_2$$

$$\leq \|x-y\|_2 + \|y-z\|_2$$

Examples:

1) Feature distance: $x, y \in \mathbb{R}^n$ represent features of two objects.

$\|x - y\|_2$

2) RMS prediction error: $y \in \mathbb{R}^n$: observed time-series.

$\hat{y} \in \mathbb{R}^n$: predicted time series.

$y - \hat{y}$: prediction error.

$$\frac{\|y - \hat{y}\|_2}{\sqrt{n}} = \text{rms}(y - \hat{y})$$

3) Nearest neighbour.

Let $z_1, z_2, \ldots, z_k \in \mathbb{R}^n$

Let $x \in \mathbb{R}^n$

Then we say z_j is the nearest neighbour of x

if $\|z_j - x\|_2 \leq \|z_i - x\|_2$ for $i = 1, 2, \ldots, k$.

Units for heterogeneous vector entries:

x, y, z : feature vectors

$x = \begin{pmatrix} \\ \end{pmatrix}, \quad y = \begin{pmatrix} \\ \end{pmatrix}, \quad z = \begin{pmatrix} \\ \end{pmatrix}$

$x = \begin{pmatrix} 1500 \\ 2 \end{pmatrix}, \quad y = \begin{pmatrix} 1600 \\ 2 \end{pmatrix}, \quad z = \begin{pmatrix} 1600 \\ 5 \end{pmatrix}$

Norms on $\mathbb{R}^n$

- $\| \ \|_p$, $\| \ \|_1$, $\| \ \|_2$, $\| \ \|_\infty$
- four properties (Defining properties)

  triangle inequality (C-S inequality in case of 2-norm)
- rms(x)
- Chebyshev inequality.

Distance on $\mathbb{R}^n$

- Ex.
- weighted norm/distance (for applications)