CSE 840: Computational Foundations of Artificial Intelligence 09 13, 2023

Diagonalization, Triangular Matrices, Metric Spaces, Normed Spaces;p-norms

Instructor: Vishnu Boddeti *Scribe: Nicole Schneider - Mahdi Masmoudi*

1 Diagonalization

Definition 1 *An operator $T \in \mathcal{L}(v)$ is diagonalizable if there exists a basis of V such that the corresponding matrix is diagonal:*

$$M(T) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Nice property: Diagonal form is the best since we have the eigenvectors as the basis.

Proposition 2 *Let V be a finite-dimension vector space. $A \in \mathcal{L}(v)$. Then the following statements are equivalent:*

(P_1) *A is diagonalizable*

(P_2) *The characteristic polynomial P_A can be decomposed into linear factors **AND** The algebraic multiplicity of the roots of P_A are equal to the geometric multiplicity*

(P_3) *If $\lambda_1, \dots \lambda_k$ are the pairwise distinct eigenvalues of A, then*

$$V = E\left(A, \lambda_1\right) \oplus E\left(A, \lambda_2\right) \dots \oplus E\left(A, \lambda_k\right)$$

2 Triangular Matrices

Definition 3 *A matrix is called upper triangular if it has the form*

$$M(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Proposition 4 $T \in \mathcal{L}(\mathrm{v}), \Phi = \{v_1, v_2 \dots v_n\}$ *a basis, then following are equivalent:*

(P_1) *$M(T, D)$ is upper triangular*

$(P_2)\, T v_j \in \operatorname{span}\{v_1, v_2 \dots v_j\} \,\forall j = 1, 2, \dots n$

$$Tv_1 = \begin{pmatrix} \lambda_1 & a_{12} & a_{13} \\ 0 & \lambda_2 & a_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 0 \\ 0 \end{pmatrix} = \lambda_1 \cdot v_1$$

$$Tv_2 = \begin{pmatrix} \lambda_1 & a_{12} & a_{13} \\ 0 & \lambda_2 & a_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{12} \\ \lambda_2 \\ 0 \end{pmatrix} = a_{12} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in$$

$$\in \text{span}\,(v_1, v_2)$$

Proposition 5 *V complex, finite-dim VS, $T \in \mathcal{L}(V)$. Then $M(T)$ has an upper triangular form for some basis.*

$\rightarrow$ *If we are in the complex field, every matrix com be expressed as an upper triangular matrix.*

Proposition 6 *Suppose $T \in \mathcal{L}(v), V$ any finite-dim VS, has an upper triangular form. Then the entries on the diagonal are precisely the eigenvalues of T.*

3 Metric Space

Metric spaces $\rightarrow$ Normed spaces $\rightarrow$ inner product spaces $\rightarrow$ Hilbert spaces

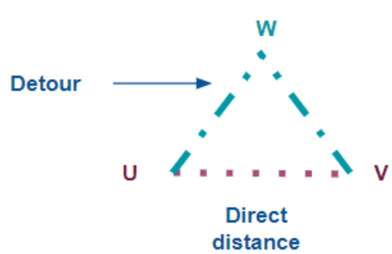
K-N N $\rightarrow$ Metric

Definition 7 *Let x be a set. A function $d : x \times x \rightarrow \mathbb{R}$ is called a metric if the following conditions hold. $\forall u, v, w \in X$:*

(P_1) $d(u,v) > 0$ *if* $u \neq v$ ***and*** $d(u,u) = 0$

(P_2) $d(u,v) = d(v,u)$ *(symmetry)*

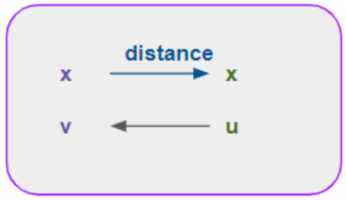
(P_3) $d(u,v) \leqslant d(u,w) + d(w,v)$



Example: *assymetric measures*

(i) friendship graph

(ii) $KL(p\|q) \neq KL(q\|p)$



<u>Notation:</u> Sequence: $(x_1, x_2 \ldots) \to (x_n)_{n\in\mathbb{N}}$

Definition 8 *Sequence:* $(x_1, x_2 \ldots) \to (x_n)_{n\in\mathbb{N}}$
A sequence $(x_n)_{n\in\mathbb{N}}$ *in a metric space* (x, d) *is called a Cauchy Sequence*
if $\forall \varepsilon > 0 \exists N \in \mathbb{N}, \forall n, m > N, d(x_n, x_m) < \varepsilon$

Definition 9 *A sequence* $(x_n)_{n\in\mathbb{N}}$ *converges to* $x \in X$ *if* $\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n > N, \quad d(x_n, x) < \varepsilon$

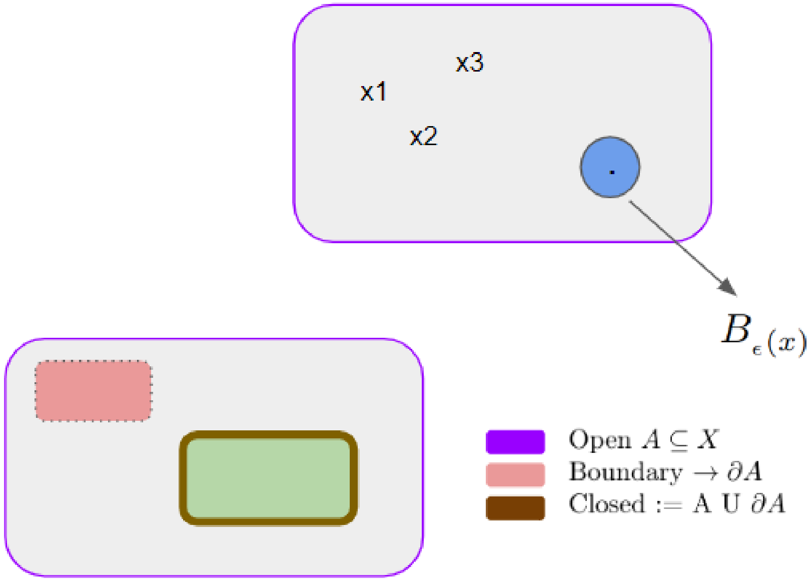
Notation: $x_n \to x, \lim_{n\to\infty} x_n = x$

sequence $(x_n)_{n\in\mathbb{N}} = 1/n$ *on* $x = (0, 1)$

Here $(x_n)_{n\in\mathbb{N}}$ *is a Cauchy seq. But does not converge.*

Sequence $(x_n)_{n\in\mathbb{N}} = 1/n$ *on* $\tilde{x} = [0, 1]$*. Here* $(x_n)_{n\in N}$ *is a cauchy sequence that converses on* $\tilde{x}$ *to 0*
.

Definition 10 *A metric space is called <u>complete</u> if every Cauchy sequence converges.*



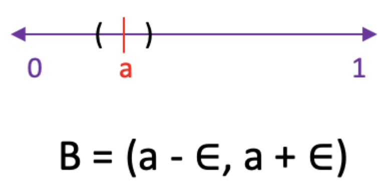
Notation: $B_\epsilon(u) := \{x \in X \mid d(x,u) \leqslant \varepsilon\} \to \varepsilon$ *- ball*

Definition 11 *A set $A \subseteq x$ is called closed if all Cauchy sequences converge and have their limit point of A.*

Definition 12 *A set $A \subseteq x$ is called open if:*

$$\forall a \in A \quad \exists \varepsilon > 0 : \quad B_\epsilon(a) \subset A$$

- *Set $[0,1]$ is closed.*
- *Set $(0,1)$ is open.*



- *A set A can be neither open nor closed. e.g. $[0,1)$.*

Definition 13 *A point $a \in A$ is an* <u>*interior point*</u> *of A if $\exists \varepsilon > 0$, s.t. $B_\epsilon(a) \subset A$.*

- *e.g. $A = [0,1]$, then $x \in (0,1)$ are interior points.*

Definition 14 *The (topological) closure of a set A is defined as the set of points that can be approximated by Cauchy sequences in A :*

$$\omega \in \bar{A} \Leftrightarrow \forall \varepsilon > 0 \exists z \in A : d(\omega, z) < \varepsilon$$

Notation: *$\bar{A}$ is the closure of A. $A \cup dA$ (always closed!)*

Definition 15 *The (topological)* <u>*interior*</u> *of a set A is defined as the set of interior points of A.*

Notation: A^0

Definition 16 *The (topological)* <u>*boundary*</u> *of a set A is defined as the set $\bar{A} \backslash A^\circ$.*

$$\begin{aligned} x &= [0,1) \quad && \text{sometimes} \\ \bar{x} &= [0,1] && \partial x = x \backslash x^0 \\ x^0 &= (0,1) && \quad\;\; = \{0\} \end{aligned}$$

$$=> boundary \partial x = \bar{x} \backslash x^0 = \{0, 1\}$$

Definition 17 *A set A is* <u>*dense*</u> *in X if we can approximate every $x \in X$ by a sequence in A. Formally, $\forall x \in X \quad \forall \in > 0, B_\epsilon(x) \cap A \neq \varnothing$.*

<u>*Example:*</u> *$\mathbb{Q} \subset \mathbb{R}$ is dense*

Definition 18 *A set $A \subset X$ is* <u>*bounded*</u> *if there exists $D > 0$ such that $\forall u, v \in A$ $d(u,v) < D$.*

4 Norms

Definition 19 *Let V be a vector space. A* <u>*norm*</u> *on V is a function $\| \cdot \| : V \to \mathbb{R}$ such that $\forall x, y \in V, \lambda \in F$, the following conditions hold:*

(P_1) $\|\lambda x\| = |\lambda| \|x\|$ *(homogeneous)*

(P_2) $\|x + y\| \leqslant \|x\| + \|y\|$ *(triangle inequality)*

(P_3) $x = 0 \Rightarrow \|x\| = 0$

(P_4) $\quad \|x\| = 0 \Rightarrow x = 0$

$\| \cdot \|$ is a semi-norm if $(P_1) - (P_3)$ are satisfied.

Intuition: norm(x) = *"length of x"*

$= \text{distance}(x, 0)$

Examples:

- *Euclidean norm on* $\mathbb{R}^d$: $\|x\| = \left(\sum_{i=1}^{d} x_i^2\right)^{1/2}$
- *Manhattan distance:* $\|x\| = \left(\sum_{i=1}^{d} |x_i|\right)$

5 p-Norm

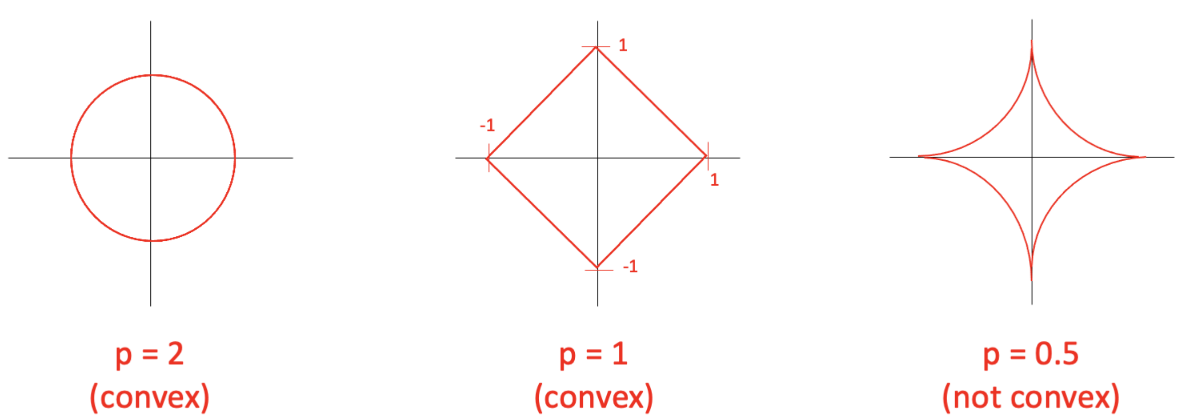
Consider $V = \mathbb{R}^d$. *Define* $\|\cdot\|_P : \mathbb{R}^d \to \mathbb{R}$

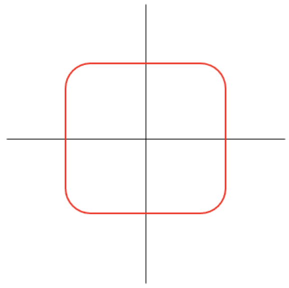
$$\|x\|_p := \left(\sum_{i=1}^{d} |x_i|^p\right)^{1/p} \quad \textit{for } 0 < p < \infty$$

- $\|\cdot\|_p$ *is a norm if* $p \geqslant 1$
- *Unit balls: the unit ball of a norm is the set of points such that norm* $\leqslant 1$ *:*

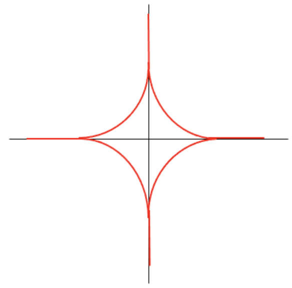
$$B_p := \left\{x \in \mathbb{R}^2 \mid \|x\|_p \leqslant 1\right\}$$

Examples: $\mathbb{R}^2$

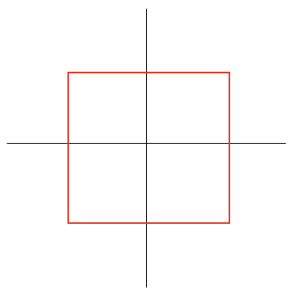




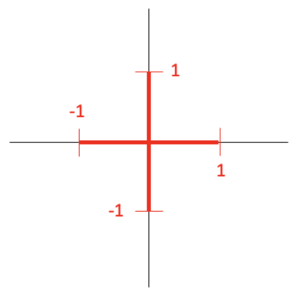
p = 5



p = 0.1



p = ∞



p = 0

Definition 20 $\|x\|_\infty := \max |x_i|$ *(is a norm)*
$\|x\|_0 :=$ *number of non-zero coordinates* $= \sum_{i=1}^{d} \mathbb{1}\{x_i \neq 0\}$ $\|x\|_0$ *is not a norm*

$$x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \|x\|_0 = 1; \lambda x = \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \|\lambda x\|_0 = 1$$

$$\lambda = 5 \qquad \neq 5.1$$