CSE 840: Computational Foundations of Artificial Intelligence February 3, 2025

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

Instructor: Vishnu Boddeti *Scribe: Shamsvi Balooni Khan*

1 Diagonalization

1.1 Definition

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

$$M(T) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Example: Any diagonal matrix D is diagonalizable because it is similar to itself. For instance,

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = I_3 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} I_3^{-1}.$$

1.2 Nice Property

The diagonal form is the best since we have the eigenvectors as the basis.

1.3 Proposition

Let V be a finite-dimensional vector space. $A \in L(V)$. Then the following statements are equivalent:

(P1) A is diagonalizable.

(P2) The characteristic polynomial P can be decomposed into linear factors, and the algebraic multiplicity of each root of P equals its geometric multiplicity.

(P3) If $\lambda_1, \ldots, \lambda_k$ are the pairwise distinct eigenvalues of A, then

$$V = E(A, \lambda_1) \oplus E(A, \lambda_2) \oplus \cdots \oplus E(A, \lambda_k)$$

Example of a diagonalizable matrix: SYMMETRIC MATRIX (property- all symmetrix matrices are diagonalizable).

*Not all matrices corresponding to linear maps are diagonalizable, in such cases, the next best thing we can hope for is a Triangular Matrix.

2 Triangular Matrices

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

$$\begin{pmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & * \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are diagonal elements, all entries to the left of the diagonal are zeros, and the entries marked with $*$ may be non-zero.

If all the entries both to the left and to the right of the diagonal are zeros, i.e., all the $*$ entries are zero, the matrix reduces to a **diagonal matrix**:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

*A diagonal matrix has non-zero entries only on its diagonal, with all off-diagonal elements equal to zero.

=> A square matrix whose all elements above the main diagonal are zero is called a lower triangular matrix.

=> A square matrix whose all elements below the main diagonal are zero is called an upper triangular matrix.

2.1 Proposition

Let $T \in L(V)$, and let $B = \{v_1, v_2, \ldots, v_n\}$ be a basis. Then the following are equivalent:

(P1) The matrix representation $M(T, \mathcal{B})$ is **upper triangular**.

(P2) For any $j = 1, 2, \ldots, n$,

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

where v_j is a particular vector from the basis $\mathcal{B}$.

If we apply this linear map to v_1, we obtain:

$$T(v_1) = \begin{pmatrix} \lambda_1 & a_{12} & a_{13} \\ 0 & \lambda_2 & a_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 0 \\ 0 \end{pmatrix} = \lambda_1 v_1$$

This result lies in the span of v_1, i.e., $T(v_1) \in \text{span}(v_1)$.

Next, applying T to v_2:

$$T(v_2) = \begin{pmatrix} \lambda_1 & a_{12} & a_{13} \\ 0 & \lambda_2 & a_{23} \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{12} \\ \lambda_2 \\ 0 \end{pmatrix} = a_{12} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = a_{12}v_1 + \lambda_2 v_2$$

This result lies in the span of v_1 and v_2, i.e.,

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

This process allows us to build a linear operator in a step-by-step manner, where each transformed basis vector $T(v_j)$ is expressed as a linear combination of $v_1, v_2, \ldots, v_j$.

$=>$ Upper Triangular Matrix **always** has eigenvalues on its diagonal

2.2 Proposition

Suppose $T \in L(V)$, where V is a finite-dimensional vector space, and T has an upper triangular form. Then M(T) has an **upper triangular form** for some **basis**.

*In a **COMPLEX FIELD**, every matrix can be expressed as an Upper Triangular matrix.

2.3 Proposition

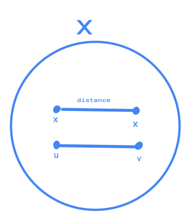
Suppose $T \in L(V)$, where V is a finite-dimensional vector space, and T has an upper triangular form, then the entries on the diagonal are precisely the eigenvalues of T.

3 Metric Spaces

3.1 Definition

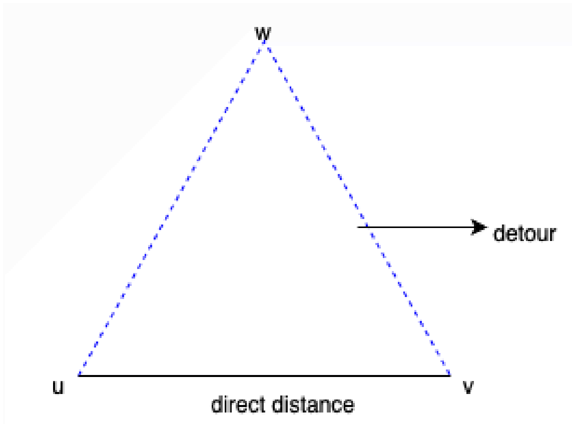
Let X be a set. A function $d : X \times X \to \mathbb{R}$ is called a **metric** if the following conditions hold:

(P1) $d(u, v) > 0$ if $u \neq v$, and $d(u, u) = 0$.



(P2) $d(u, v) = d(v, u)$ (symmetry).

(P3) $d(u, v) \leq d(u, w) + d(w, v)$ (triangle inequality).



d(u,v) is the direct distance.

=> Distance between the points must be positive.

3.2 Definition - Sequences

Sequence is basically an ordered set of elements.

A sequence (x_n) in a metric space (X, d) is called a **Cauchy sequence** if:

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n, m > N, d(x_n, x_m) < \epsilon$$

=> Beyond a certain value of N, looking at the points, those points are going to be no further apart than ϵ, i.e., these points will get closer and closer to each other as we go further in the sequence or increase the index of the sequence.

=> Every convergent sequence is a Cauchy Sequence.

3.3 Definition - Converge

A sequence $(x_n)_{n \in \mathbb{N}}$ **converges** to a point $x \in X$ if

$$\forall\, \epsilon > 0,\ \exists\, N \in \mathbb{N}\,\text{such that}\, \forall n > N,\ d(x_n, x) < \epsilon$$

This means the points in the sequence get arbitrarily close to x as n increases.

*The point to which the sequence is converging also belongs to the set.

Notation:

$$x_n \to x \quad \text{or equivalently} \quad \lim_{n \to \infty} x_n = x$$

Consider the sequence $(x_n)_{n \in \mathbb{N}} = \frac{1}{n}$ on $X = (0, 1)$ (an open set).

Here, $(x_n)_{n \in \mathbb{N}}$ is a **Cauchy sequence**, but it does **not converge** in X.

(Note: All values are between 0 and 1, but 0 and 1 are not included in the set.

*The sequence is approaching 0, which is not in X.)

Now consider the sequence $(x_n)_{n \in \mathbb{N}} = \frac{1}{n}$ on $\bar{X} = [0, 1]$ (a closed set).

Here, $(x_n)_{n \in \mathbb{N}}$ is a **Cauchy sequence** that **converges** in $\bar{X}$ to 0.

3.4 Definition - Complete

A metric space is called **complete** if every Cauchy sequence converges.

Example: The set of real numbers $\mathbb{R}$ with the standard metric is a **complete** set.

3.5 Definition - Epsilon ball

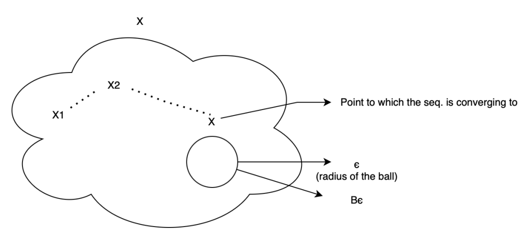
The Epsilon-Ball (ϵ-ball) around a point x is denoted by $B_\epsilon(x)$ and defined as:

$$B_\epsilon(u) := \{x \in X \mid d(x, u) < \epsilon\}$$

Where B_ϵ is the Epsilon-Ball

B_ϵ is defined as a set of points in a set X, such that the distance from x to u is less than (or equal to) ϵ.

This represents the set of all points in X whose distance from u is less than ϵ.



*If we can draw a B_ϵ around a point, then that set is open (does not include the boundary).

*$\epsilon > 0$

***Open Set:** $A \subseteq X$ (A is a subset of X.)

***Boundary:** ∂A (All the points whose distance is less than ϵ)

***Closed Set:** $A \cup \partial A$ (taking an open set and adding the boundary points to it)

3.6 Definition - Closed Set

A set $A \subseteq X$ is called **closed** if every Cauchy sequence in A converges to a limit in A.

**Limit point: the point to which the Cauchy Sequence is converging to

3.7 Definition - Open set

A set $A \subset X$ is called **open** if:

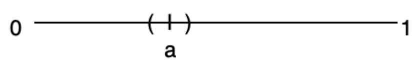
$$\forall a \in A, \exists \epsilon > 0 : B_\epsilon(a) \subset A$$

* Open sets can be bounded too.

3.8 Examples of Sets

=> Closed:[0,1] ; 0,1, set of all real numbers

=> Open: (0,1)



Here, B $= (\varepsilon, a + \varepsilon)$

=> A set can be neither open nor closed, eg: [0,1)

=> A set can be both open and closed

3.9 Definition - Interior Point

Def: A point $a \in A$ is an **interior point** of A if

$$\exists\, \varepsilon > 0, \text{ such that } B_\varepsilon(a) \subset A$$

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

3.10 Definition - Closure

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

$$w \in \overline{A} \iff \forall \varepsilon > 0, \exists z \in A : d(w, z) < \varepsilon$$

Here, we basically take the open set A and its union with all of the boundary points.

Notation: $\overline{A}$ is the closure of A.

* $A \cup \partial A$, is always closed.

3.11 Definition - (Topological)Interior

The **interior** of a set A is defined as the set of interior points of A:

Notation: A°

3.12 Definition - (Topological) Boundary

The **boundary** of a set A is defined as the set:

$$\overline{A} \setminus A^\circ$$

* We simply take the closure and remove the interior.

Examples:

$$X = \{0, 1\}$$
$$\overline{X} = [0, 1] \quad \text{(closure)}$$
$$X^\circ = (0, 1) \quad \text{(interior)}$$
$$\Rightarrow \text{boundary} \quad \partial X = \overline{X} \setminus X^\circ = \{0, 1\}$$
$$\text{sometimes} \quad \partial X = X \setminus X^\circ = \{0\}$$

3.13 Definition - Dense

A set A is **dense** in X if we can approximate **every** $x \in X$ by a sequence in A.

$$\text{Formally,} \quad \forall x \in X,\ \forall \varepsilon > 0,\ B_\varepsilon(x) \cap A \neq \emptyset$$

Example: $\mathbb{Q} \subset \mathbb{R}$ is **dense** in $\mathbb{R}$.

*Taking any real number, we can have a sequence of rational numbers that can get arbitrarily close to the real number, however the sequence will never reach the real number.

3.14 Definition - Bounded

A set $A \subset X$ is **bounded** if there exists $D > 0$ such that

$$\forall u, v \in A, \quad d(u, v) \leq D$$

* Where D is some scalar

* Taking any 2 pairs of points from set A,then the largest possible distance is smaller than D

* D $>$ 0

4 Norms

In general, there are no algebraic operations defined on a metric space, only a distance function. Most of the spaces that arise in analysis are vector, or linear, spaces, and the metrics on them are usually derived from a norm, which gives the "length" of a vector

4.1 Definition

Let V be a vector space. A **norm** on V is a **function** $\|\cdot\| : V \to \mathbb{R}$ such that for all $x, y \in V$ and $\lambda \in \mathbb{F}$, the following conditions hold:

(P1) $\|\lambda x\| = |\lambda| \|x\|$ (*homogeneous*)

(P2) $\|x + y\| \leq \|x\| + \|y\|$ (*triangle inequality*)

(P3) $x = 0 \implies \|x\| = 0$ (*norm property*)

(P4) $\|x\| = 0 \implies x = 0$ (*vector property*)

$\|\cdot\|$ is a **semi-norm** if (P1)–(P3) are satisfied.

Intuition: The norm of x represents the **"length of x"** or the **distance** between $(x, 0)$, (where 0 is the origin).

Example:

- **Euclidean norm** on $\mathbb{R}^d$:

$$\|x\| = \left(\sum_{i=1}^{d} x_i^2 \right)^{1/2}$$

  Each x_i is one of the coordinates of x

- **Manhattan distance**:

$$\|x\| = \sum_{i=1}^{d} |x_i|$$

  Absolute value of the coordinates of x

4.2 P-Norm

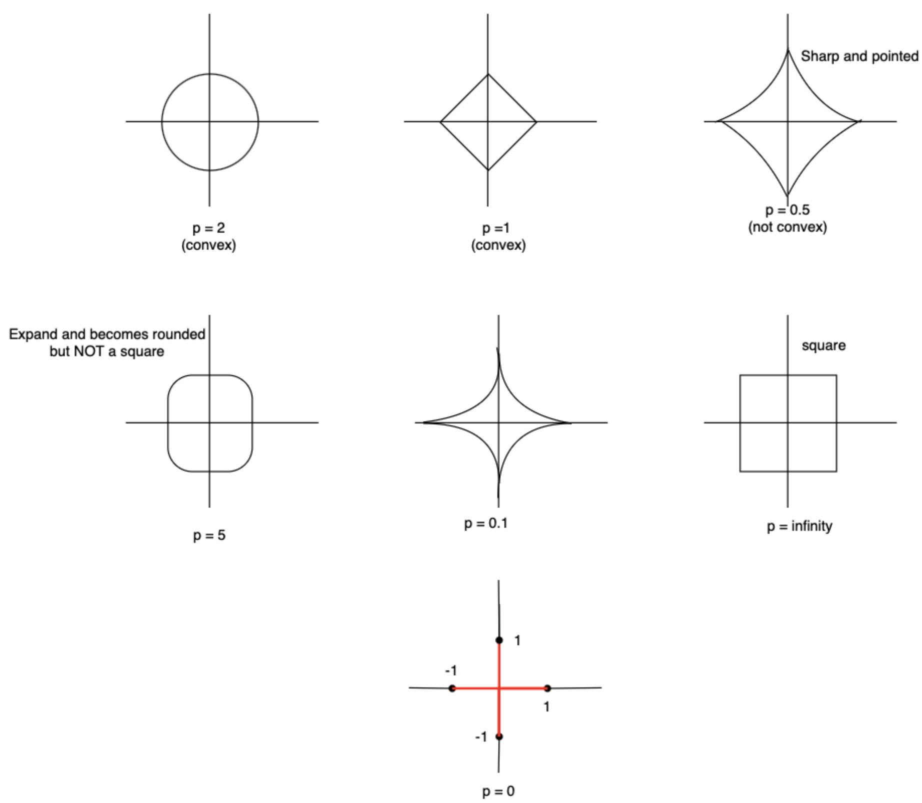
Consider $V = \mathbb{R}^d$. Define:

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

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

$$B_p := \{ x \in \mathbb{R}^d \mid \|x\|_p \leq 1 \}$$

Examples:



Def:

$$\|x\|_\infty := \max |x_i| \quad \text{(is a norm)}$$

This represents the maximum value over the coordinates of the vector.

$$\|x\|_0 := \text{number of non-zero coordinates}$$

This can be written as:

$$\|x\|_0 = \sum_{i=1}^{d} \mathbf{1}_{\{x_i \neq 0\}}$$

where **1** is the indicator function, which equals 1 when the condition $x_i \neq 0$ is satisfied and 0 otherwise.

Note: $\|x\|_0$ is not a norm.

Example:

$$x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \|x\|_0 = 1$$

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

$$\|\lambda x\|_0 \neq \lambda \|x\|_0$$

5 References

Diagonalization. (n.d.). https://textbooks.math.gatech.edu/ila/diagonalization.html

Triangular Matrix - Lower and upper triangular matrix, examples. (n.d.). Cuemath. https://www.cuemath.com/algebra/ matrix/

Bert, M. (n.d.). Metric spaces. In Metric Spaces (pp. 93–96). https://www.math.ucdavis.edu/ hunter/m125a/intro$_{a}$*nalys*