

Math BackgroundLinear Algebra

Vector $x \in \mathbb{R}^n$ - n-dim Euclidean space

$$x = (x_1, \dots, x_n) \equiv [x_1 \ x_2 \ \dots \ x_n]^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Norm of x $\|x\|$ satisfies properties :

$$(a) \|x\| \geq 0$$

$$(b) \|x\| = 0 \iff x = 0$$

$$(c) \|cx\| = |c| \|x\|, \text{ for } c \in \mathbb{R}$$

$$(d) \|x+y\| \leq \|x\| + \|y\| \leftarrow \text{Triangle Ineq.}$$

Inner Product $x^T y = \sum_{i=1}^n x_i y_i$

$$\text{Euclidean Norm} \quad \|x\| = \sqrt{x^T x} = \sqrt{\sum_{i=1}^n x_i^2}$$

Two important results for Euclidean norm :

1) Pythagorean Theorem : If $x^T y = 0$,
 $\|x+y\|^2 = \|x\|^2 + \|y\|^2$

2) Cauchy-Schwarz Inequality :

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

"=" iff $x = \alpha y$ for some $\alpha \in \mathbb{R}$

Other norms: ℓ_1 -norm : $\|x\|_1 = \sum_{i=1}^n |x_i|$

sup-norm or ℓ_∞ -norm : $\|x\|_\infty = \max_i |x_i|$

Sequences $\{x_k\}_{k=1}^\infty$ or $\{x_k\}$, $x_k \in \mathbb{R}^n$

Definition (convergence) $x_k \rightarrow x$, $\lim_{k \rightarrow \infty} x_k = x$

Given $\varepsilon > 0$, $\exists N_\varepsilon$ s.t. $\|x_k - x\| < \varepsilon \quad \forall k \geq N_\varepsilon$

Definition (Cauchy sequence) $\{x_k\}$ is Cauchy if

given $\varepsilon > 0$, $\exists N_\varepsilon$ s.t. $\|x_k - x_m\| < \varepsilon \quad \forall k, m \geq N_\varepsilon$.

$\{x_k\}$ converges \iff $\{x_k\}$ is Cauchy

Definition (subsequence) infinite subset of $\{x_k\}$.

$\{x_k : k \in K\}$ or $\{x_k\}_{k \in K}$, K : infinite subset of \mathbb{Z}^+

Definition (Limit point) x is a limit point of $\{x_k\}$ if \exists a subsequence of $\{x_k\}$ that converges to x .

Definition (bounded sequence) $\|x_k\| \leq b, \forall k$

Results about Bounded Sequences

1. Every bounded has at least one limit point
2. A bounded sequence converges iff it has a unique limit point

Scalar Sequences $\{x_k\}$, $x_k \in \mathbb{R}$.

- If $\{x_k\}$ is bounded above and non-decreasing it converges (below) (increasing)
(smallest)
- The largest limit point of $\{x_k\}$ is $\limsup_{k \rightarrow \infty} x_k$ (inf)
- $\{x_k\}$ converges $\Leftrightarrow -\infty < \liminf_{k \rightarrow \infty} x_k = \limsup_{k \rightarrow \infty} x_k < \infty$

Definition (continuity) A real-valued function f is continuous at x if for every $\{x_k\}$ converging to x

$$\lim_{k \rightarrow \infty} f(x_k) = f(x)$$

Equivalently, given $\varepsilon > 0$, $\exists \delta > 0$ s.t.

$$|f(x) - f(y)| < \varepsilon \quad \forall \|y - x\| < \delta$$

f is continuous if it is continuous at all points x

Definition (coercive) A real-valued function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is coercive if for every $\{x_k\} \subset \mathbb{R}^n$ s.t. $\|x_k\| \rightarrow \infty$, $f(x_k) \rightarrow \infty$

Examples 1) $x \in \mathbb{R}^2$, $f(x) = x_1^2 + x_2^2$ — coercive

2) $x \in \mathbb{R}$, $f(x) = 1 - e^{-|x|}$ — not coercive

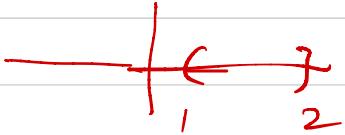
3) $x \in \mathbb{R}^2$, $f(x) = x_1^2 + x_2^2 - 2x_1 x_2$
 $= (x_1 - x_2)^2$ — not coercive

Closed and Open sets

A set $S \subseteq \mathbb{R}^n$ is open if $\forall x \in S$ we can draw a ball around x that is contained in S ,

i.e. $\forall x \in S$, $\exists \varepsilon > 0$ s.t.

$$\{y : \|y - x\| < \varepsilon\} \subseteq S$$



S is closed if S^c is open

\equiv if S contains all limit points of all sequences in S .

Examples $(1, 2) = \{x \in \mathbb{R} : 1 < x < 2\}$ - open

\mathbb{R} is both $(-\infty, 1) = \{x \in \mathbb{R} : x < 1\}$ - open

open and closed $[1, \infty)$ is closed because complement open

$(1, 2]$ is neither open nor closed

Compact Set $S \subseteq \mathbb{R}^n$ is compact if it is closed and bounded. S is bounded if $\exists M$ s.t. $\|x\| \leq M \quad \forall x \in S$

Examples $[1, 2] = \{x \in \mathbb{R} : 1 \leq x \leq 2\}$

$$\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 4\}$$

Extrema of sets of scalars Let $A \subseteq \mathbb{R}$.

- The infimum of A , or $\inf A$, is largest y s.t. $y \leq x$, $\forall x \in A$. If no such y exists, $\inf A = -\infty$
- Similar definition for supremum of A or $\sup A$.
- If $\inf A = x^*$ for some $x^* \in A$, then $x^* = \min A$ or minimum of A .
- similarly $\max A$ or maximum of A .

- Examples
- 1) $A = [1, 2]$, $\inf A = \min A = 1$
 $\sup A = \max A = 2$
 - 2) $A = (1, 2]$, $\inf A = 1$, not achieved
 $\sup A = \max A = 2$
 - 3) $A = (1, \infty)$, $\sup A = \infty$, no maximum

Extrema of Functions Let $S \subseteq \mathbb{R}^n$, $f: S \rightarrow \mathbb{R}$

$$\inf_{x \in S} f(x) = \inf \{f(x) : x \in S\}.$$

If $\exists x^* \in S$ s.t. $\inf_{x \in S} f(x) = f(x^*)$ then

f achieves (attains) its minimum and

$$f(x^*) = \min_{x \in S} f(x)$$

x^* is called a minimizer of f , written as

$$x^* \in \arg \min_{x \in S} f(x)$$

If x^* is unique we write $x^* = \arg \min_{x \in S} f(x)$.

Similarly, supremum and maximum of f .

Example 1 $f(x) = x$, $x \in (-1, 2)$

$$\sup f(x) = 2 \quad) \text{ neither achieved}$$

$$\inf f(x) = -1$$

Example 2 $f(x) = x^2$, $x \in \mathbb{R}$

$$\inf f(x) = \min f(x) = 0, x^* = 0$$

$$\sup f(x) = \infty, \text{ max does not exist.}$$

Weierstrass' Theorem (Extreme Value Theorem)

If f is a continuous function on a compact set, $S \subseteq \mathbb{R}^n$, then f attains its min and max on S , i.e.,

$$\exists x_1 \in S \text{ s.t. } f(x_1) = \inf_{x \in S} f(x)$$

$$\exists x_2 \in S \text{ s.t. } f(x_2) = \sup_{x \in S} f(x)$$

Proof (for existence of min; max is similar)

Let $\{\bar{x}_k\} \subseteq S$ be s.t.

$$\inf_{x \in S} f(x) \leq f(\bar{x}_k) \leq \inf_{x \in S} f(x) + \frac{1}{k}$$

$$\text{Then } \lim_{k \rightarrow \infty} f(\bar{x}_k) = \inf_{x \in S} f(x)$$

S is bounded $\Rightarrow \{\bar{x}_k\}$ has at least one limit point x_1 ,

S is closed $\Rightarrow x_1 \in S$.

$$f \text{ is continuous} \Rightarrow f(x_1) = \lim_{k \rightarrow \infty} f(\bar{x}_k) = \inf_{x \in S} f(x)$$

Corollary Let f be continuous on closed set S , that is not necessarily bounded. If f is coercive on S it attains its min. on S . ($-f$)
(max)

Proof Consider $\{\bar{x}_k\}$ as in proof of WT.

f is coercive on $S \Rightarrow \{\bar{x}_k\}$ is bounded

Rest of proof same as proof of WT.

Example

$$f(x) = f(x_1, x_2, x_3) = x_1^4 + 2x_2^2 + e^{-x_3} + e^{2x_3}$$

(a) Does f achieve its min. and max. on:

$$\mathcal{S}_1 = \{x \in \mathbb{R}^3 : x_1^2 + 2x_2^2 + 3x_3^2 \leq 6\} ?$$

\mathcal{S}_1 is compact and f is continuous

\Rightarrow both min and max achieved on \mathcal{S}_1 by WT

(b) Does f achieve its min and max over \mathbb{R}^3 ?

- $f \rightarrow \infty$ whenever $\|x\| \rightarrow \infty \Rightarrow f$ is coercive

- \mathbb{R}^3 is closed

$\Rightarrow f$ achieves its min. on \mathbb{R}^3 by Corollary to WT.

max. does not exist since $f \rightarrow \infty$ as $\|x\| \rightarrow \infty$.

(c) Does f achieve its min and max over:

$$\mathcal{S}_2 = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 3\} .$$

- \mathcal{S}_2 is closed, but not bounded.

For example, $x_1 = 0$, $x_2 = 3 - x_3$ and $x_3 \rightarrow \infty$

- Since f is coercive, min achieved

- Max does not exist since setting $x_1 = 0$, $x_2 = 3 - x_3$ and letting $x_3 \rightarrow \infty$ makes $f \rightarrow \infty$.

Matrices

$$A_{m \times n} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix}$$

a_{ij}
 $\equiv A_{ij}$
 $\equiv [A]_{ij}$

- Square matrix: $m=n$, $A_{n \times n}$
- Symmetric : $A^T = A$
- Upper Triangular : $a_{ij} = 0$ whenever $i < j$.

Square Matrices $A_{n \times n}$

- A is singular if $\det(A) = 0$, else non-singular
- If $\det(A) \neq 0$, A^{-1} exists and $= \frac{1}{\det(A)} \text{adj}(A)$
- If A, B are $n \times n$ matrices,
 $\det(AB) = \det(A)\det(B) = \det(BA)$

Eigenvalues and Eigen vectors

$A_{n \times n}$ has eigenvector $u \neq 0$ if

$$Au = \lambda u$$

$$\Rightarrow (A - \lambda I)u = 0$$

$$\Rightarrow \det(A - \lambda I) = 0 \leftarrow \text{characteristic poly.}$$

Solution to characteristic polynomial gives eigenvalues and corresponding eigenvectors

Example $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$

$$\begin{aligned}\det(A - \lambda I) &= (1-\lambda)((2-\lambda)^2 - 1) \\ &= (1-\lambda)(3-4\lambda+\lambda^2) \\ &= (1-\lambda)^2(3-\lambda)\end{aligned}$$

$$\lambda_1 = 3, \quad \lambda_2 = \lambda_3 = 1 \quad \leftarrow \text{multiplicity 2}$$

Eigenvectors $(A - 3 I) u_1 = 0$

$$\Rightarrow \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \\ u_{13} \end{bmatrix} = 0$$

$$\Rightarrow u_{11} = 0, \quad u_{12} = -u_{13}$$

$$\Rightarrow u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \quad \leftarrow \text{unit norm eigenvector for } \lambda_1$$

u_2, u_3 found by solving $(A - I) u_j = 0, j=2,3$.

i.e. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} u_{j1} \\ u_{j2} \\ u_{j3} \end{bmatrix} = 0 \Rightarrow u_{j2} = u_{j3}$
 u_{j1} arbitrary

Set $u_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $u_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

Then u_1, u_2, u_3 are orthonormal

Eigen Decomposition of Symmetric Matrices

Let A be a symmetric $n \times n$ matrix, i.e., $A^T = A$

1) Result All eigenvalues of A are real

Proof Let λ be eigenvalue of A and u be the corresponding eigenvector. Then

$$\begin{aligned} Au &= \lambda u \Rightarrow A u^* = \lambda^* u^* \\ \Rightarrow u^T A u^* &= \lambda^* u^T u^* \Rightarrow (Au)^T u^* = \lambda^* u^T u^* \\ \Rightarrow (\lambda u)^T u^* &= \lambda^* u^T u^* \Rightarrow \lambda = \lambda^* \end{aligned}$$

Result Eigenvectors corresponding to distinct eigenvalues are orthogonal

Proof Let λ_1, λ_2 be s.t. $\lambda_1 \neq \lambda_2$

$$A u_1 = \lambda_1 u_1 \quad \text{and} \quad A u_2 = \lambda_2 u_2$$

$$\begin{aligned} \lambda_1 u_1^T u_2 &= (Au_1)^T u_2 = u_1^T A u_2 \\ &= u_1^T \lambda_2 u_2 = \lambda_2 u_1^T u_2 \\ \Rightarrow u_1^T u_2 &= 0 \quad \text{since } \lambda_1 \neq \lambda_2 \end{aligned}$$

3) Result If λ is an eigenvalue with multiplicity k , we can find k orthogonal eigenvectors for λ

Conclusion For real symmetric A , eigenvalues are real, and we can find orthonormal eigenvectors, u_1, \dots, u_n

$$u_i^T u_j = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$$