Math Background

Linear 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 $\quad \|x\|$ satisfies properties :

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

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

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

(d) $\|x+y\| \leq \|x\| + \|y\|$ ← Triangle Ineq.

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

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

Definition (Convergence) $x_k \to x$, $\lim_{k\to\infty} x_k = x$

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 $\{x_k\}$.

$\{x_k : k \in \mathcal{K}\}$ or $\{x_k\}_{\mathcal{K}}$, $\mathcal{K}$: 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 (below) and non-decreasing (increasing) it converges
- The largest (smallest) limit point of $\{x_k\}$ is $\limsup_{k\to\infty} x_k$ (inf)
- $\{x_k\}$ converges $\iff -\infty < \liminf_{k\to\infty} x_k = \limsup_{k\to\infty} x_k < \infty$

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

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

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

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

f is continuous if it is continuous at all points x

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

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) $x \in \mathbb{R}^2$, $f(x) = x_1^2 + x_2^2 - 2x_1x_2$

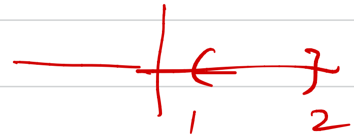
$= (x_1 - x_2)^2$ — not coercive

Closed and Open sets

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

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

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



$\mathcal{S}$ is closed if $\mathcal{S}^c$ is open

$\equiv$ if $\mathcal{S}$ contains all limit points of all sequences in $\mathcal{S}$.

Examples

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

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

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

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

$\mathbb{R}$ is both open and closed

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

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 $A \subset \mathbb{R}$.

- 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 $S \subseteq \mathbb{R}^n$, $f : S \to \mathbb{R}$

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

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

$\sup f(x) = 2$
$\inf f(x) = -1$) neither achieved

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

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

$\sup f(x) = \infty$, 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 $\{z_k\} \subseteq S$ be s.t.

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

Then $$\lim_{k \to \infty} f(z_k) = \inf_{x \in S} f(x)$$

S is bounded $\Rightarrow$ $\{z_k\}$ has at least one limit point x_1

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

f is continuous $\Rightarrow$ $f(x_1) = \lim_{k\to\infty} f(z_k) = \inf_{x \in S} f(x)$ ∎

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

Proof Consider $\{z_k\}$ as in proof of WT.

f is coercive on S $\Rightarrow$ $\{z_k\}$ 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:

$$S_1 = \{x \in \mathbb{R}^3 : x_1^2 + 2x_2^2 + 3x_3^2 \leq 6\} ?$$

S_1 is compact and f is continuous

$\Rightarrow$ both min and max achieved on S_1 by WT

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

- $f \to \infty$ whenever $\|x\| \to \infty$ $\Rightarrow$ 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 $f \to \infty$ as $\|x\| \to \infty$.

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

$$S_2 = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 3\}.$$

- S_2 is closed, but not bounded.
  For example, $x_1 = 0$, $x_2 = 3 - x_3$ and $x_3 \to \infty$
- Since f is coercive, min achieved
- max does not exist since setting $x_1 = 0$, $x_2 = 3 - x_3$ and letting $x_3 \to \infty$ makes $f \to \infty$.

Matrices

$$A_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix} \qquad \begin{aligned} &a_{ij} \\ &\equiv A_{ij} \\ &\equiv [A]_{ij} \end{aligned}$$

- 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 $= \dfrac{1}{\det(A)} \operatorname{adj}(A)$
- If A, B are $n\times n$ matrices,

$$\det(AB) = \det(A)\det(B) = \det(BA)$$

Eigenvalues and Eigenvectors

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

$$Au = \lambda u$$

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

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

Solution to characteristic polynomial gives eigenvalues and corresponding eigenvectors

<u>Example</u> $\quad A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$

$$\det(A - \lambda I) = (1-\lambda)\left((2-\lambda)^2 - 1\right)$$
$$= (1-\lambda)(3 - 4\lambda + \lambda^2)$$
$$= (1-\lambda)^2(3-\lambda)$$

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

<u>Eigenvectors</u> $\quad (A - 3I)\, u_1 = 0$

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

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

$\Rightarrow \quad u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \leftarrow$ unit norm eigenvector for λ_1

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

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

u_{j1} arbitrary

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

Then $\quad u_1, u_2, u_3$ are <u>orthonormal</u>

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

$$Au = \lambda u \Rightarrow A u^* = \lambda^* u^*$$

$$\Rightarrow u^T A u^* = \lambda^* u^T u^* \Rightarrow (Au)^T u^* = \lambda^* u^T u^*$$

$$\Rightarrow (\lambda u)^T u^* = \lambda^* u^T u^* \qquad \Rightarrow \qquad \lambda = \lambda^*$$

Result Eigenvectors corresponding to distinct eigenvalues are orthogonal

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

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

$$\lambda_1 u_1^T u_2 = (A u_1)^T u_2 = u_1^T A u_2$$

$$= u_1^T \lambda_2 u_2 = \lambda_2 u_1^T u_2$$

$$\Rightarrow u_1^T u_2 = 0 \quad \text{since } \lambda_1 \neq \lambda_2$$

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 <u>orthonormal</u> eigenvectors, $u_1, \dots, u_n$

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