# Lecture 2

- Background mathematics
  - Topology in $\mathbb{R}^n$
  - Analysis in $\mathbb{R}^n$
  - Linear algebra

# Topology in $R^n$

- Open set

A subset $\mathcal{C}$ of $\mathbb{R}^n$ is called open, if for every $\mathbf{x} \in \mathcal{C}$ there exists $\varepsilon > 0$ such that the ball $B_\varepsilon(\mathbf{x}) = \{\mathbf{y} | \|\mathbf{y} - \mathbf{x}\|_2 \leq \varepsilon\}$ is included in $\mathcal{C}$.

Examples: $\{x | a < x < b\}$, $\{\mathbf{x} | \|\mathbf{x} - \mathbf{a}\| < 1\}$, $\{\mathbf{x} | \mathbf{x} > \mathbf{0}\}$, $\mathbb{S}_{++} = \{\mathbf{X} | \mathbf{X} \succ \mathbf{0}\}$

- Closed set

A subset $\mathcal{C}$ of $\mathbb{R}^n$ is called closed, if its complement $\mathcal{C}^c = \mathbb{R}^n \backslash \mathcal{C}$ is open.

Examples: $\{x | a \leq x \leq b\}$, $\{\mathbf{x} | \|\mathbf{x} - \mathbf{a}\| \leq 1\}$, $\{\mathbf{x} | \mathbf{x} \geq \mathbf{0}\}$, $\mathbb{S}_{+} = \{\mathbf{X} | \mathbf{X} \succeq \mathbf{0}\}$

- Bounded set

A subset $\mathcal{C}$ of $\mathbb{R}^n$ is called bounded, if $\exists R > 0$ such that $\|\mathbf{x}\| < R$, $\forall \mathbf{x} \in \mathcal{C}$.

Examples: $\{x | a \leq x < b\}$, $\{\mathbf{x} | \|\mathbf{x} - \mathbf{a}\| < 1\}$, $\{\mathbf{x} | \mathbf{1} > \mathbf{x} \geq \mathbf{0}\}$, $\{\mathbf{X} | \mathbf{I} \succeq \mathbf{X} \succ \mathbf{0}\}$

- Compact set

A subset $\mathcal{C}$ of $\mathbb{R}^n$ is called compact, if it is both bounded and closed.

Examples: $\{x | a \leq x \leq b\}$, $\{\mathbf{x} | \|\mathbf{x} - \mathbf{a}\| \leq 1\}$, $\{\mathbf{x} | \mathbf{1} \geq \mathbf{x} \geq \mathbf{0}\}$, $\{\mathbf{X} | \mathbf{I} \succeq \mathbf{X} \succeq \mathbf{0}\}$

# Topology in R$^n$

- Interior

The interior of $\mathcal{C} \subseteq \mathbb{R}^n$ is defined as $\mathcal{C}^\circ = \{\mathbf{y} | \exists \varepsilon > 0 \text{ such that } B_\varepsilon(\mathbf{y}) \subset \mathcal{C}\}$.

Examples: $(\{x | a \leq x \leq b\})^\circ = \{x | a < x < b\}$, $(\{\mathbf{x} | \|\mathbf{x} - \mathbf{a}\| \leq 1\})^\circ = \{\mathbf{x} | \|\mathbf{x} - \mathbf{a}\| < 1\}$, $(\{\mathbf{x} | \mathbf{1} \geq \mathbf{x} \geq \mathbf{0}\})^\circ = \{\mathbf{x} | \mathbf{1} > \mathbf{x} > \mathbf{0}\}$, $(\{\mathbf{X} | \mathbf{I} \succeq \mathbf{X} \succeq \mathbf{0}\})^\circ = \{\mathbf{X} | \mathbf{I} \succ \mathbf{X} \succ \mathbf{0}\}$

- Closure

The closure of $\mathcal{C} \subset \mathbb{R}^n$ is defined as $\bar{\mathcal{C}} = \mathbb{R}^n \backslash (\mathbb{R}^n \backslash \mathcal{C})^\circ = ((\mathcal{C}^c)^\circ)^c$.

Examples: $\overline{\{x | a \leq x < b\}} = \{x | a \leq x \leq b\}$, $\overline{\{\mathbf{x} | \|\mathbf{x} - \mathbf{a}\| < 1\}} = \{\mathbf{x} | \|\mathbf{x} - \mathbf{a}\| \leq 1\}$, $\overline{\{\mathbf{x} | \mathbf{1} \geq \mathbf{x} > \mathbf{0}\}} = \{\mathbf{x} | \mathbf{1} \geq \mathbf{x} \geq \mathbf{0}\}$, $\overline{\{\mathbf{X} | \mathbf{I} \succ \mathbf{X} \succ \mathbf{0}\}} = \{\mathbf{X} | \mathbf{I} \succeq \mathbf{X} \succeq \mathbf{0}\}$

- Boundary

The boundary of $\mathcal{C} \subseteq \mathbb{R}^n$ is defined as $\partial \mathcal{C} = \bar{\mathcal{C}} \backslash \mathcal{C}^\circ$.

Examples: $\partial(\{x | a \leq x < b\}) = \{a, b\}$, $\partial(\{\mathbf{x} | \|\mathbf{x} - \mathbf{a}\| < 1\}) = \{\mathbf{x} | \|\mathbf{x} - \mathbf{a}\| = 1\}$, $\partial(\{\mathbf{x} | \mathbf{1} \geq \mathbf{x} > \mathbf{0}\}) = \{\mathbf{x} | \exists i \text{ such that } x_i = 0 \text{ or } 1\}$, $\partial(\{\mathbf{X} | \mathbf{I} \succ \mathbf{X} \succ \mathbf{0}\}) = \{\mathbf{X} | 0 \leq \lambda_i(\mathbf{X}) \leq 1, \forall i, \text{ and } \exists j \text{ such that } \lambda_j(\mathbf{X}) = 0 \text{ or } 1\}$

# Analysis in R$^n$

- Sequences

A sequence in $\mathbb{R}^n$ is a set of vectors in $\mathbb{R}^n$ indexed by positive integers, denoted as $\{\mathbf{x}_k\}_{k=1}^{\infty}$ of simply $\{\mathbf{x}_k\}$. Sometimes the index may start from 0 or even negative integers for convenience.

Examples: $x_k = 10^k$, $x_k = 10^{-k}$, $x_k = (-1)^k k$

A sequence $\{x_k\}$ in $\mathbb{R}$ is *increasing* if $x_k < x_{k+1}$ for all $k$. If $x_k \leq x_{k+1}$, then we say that the sequence is *nondecreasing*. Similarly, we can define *decreasing* and *nonincreasing* sequences. Nonincreasing or nondecreasing sequences are called *monotone* sequences.

Examples: $x_k = 10^k$, $x_k = 10^{-k}$, $x_k = \lfloor k/4 \rfloor$, $x_k = (-1)^k k$

A sequence $\{\mathbf{x}_k\}$ is *bounded* if $\exists R > 0$ such that $\|\mathbf{x}_k\| < R$, $\forall k$.

Examples: $x_k = 10^k$, $x_k = 10^{-k}$, $x_k = \lfloor k/4 \rfloor$, $x_k = (-1)^k k$

# Analysis in $R^n$

- Sequences

A sequence $\{x_k\}$ in $\mathbb{R}$ is *upper bounded* (or *bounded above*) if $\exists u$ such that $x_k \leq u$, $\forall k$. Lower boundedness (or *bounded below*) is defined similarly.

Examples: $x_k = 10^k$, $x_k = 10^{-k}$, $x_k = \lfloor k/4 \rfloor$, $x_k = (-1)^k k$

- Supremum and infimum

Suppose $\mathcal{C} \subseteq \mathbb{R}$. A number $a$ is an *upper bound* on $\mathcal{C}$ if for each $x \in \mathcal{C}$, $x \leq a$. The set of upper bounds on a set $C$ is either empty (in which case we say $\mathcal{C}$ is unbounded above), all of $\mathbb{R}$ (only when $\mathcal{C} = \emptyset$), or a closed infinite interval $[b, \infty)$. The number $b$ is called the *least upper bound* or *supremum* of the set $\mathcal{C}$, and is denoted $\sup \mathcal{C}$. We take $\sup \emptyset = -\infty$, and $\sup \mathcal{C} = \infty$ if $\mathcal{C}$ is unbounded above.
We define lower bound, and infimum, in a similar way.

Examples: $x_k = 2 - 10^{-k}$, $x_k = 1 + 10^{-k}$, $x_k = 1 + (-1)^k \times 10^{-k}$

# Analysis in $\mathrm{R}^n$

- Limit

A sequence $\{x_k\}$ in $\mathbb{R}^n$ is *convergent* (or *has a limit*) if $\exists \mathbf{x}^*$, such that for all $\varepsilon > 0$, $\exists K$ such that $\|\mathbf{x}_k - \mathbf{x}^*\| \leq \varepsilon$, $\forall k > K$.

Examples: $x_k = 2 - 10^{-k}$, $x_k = 1 + 10^{-k}$, $x_k = 1 + (-1)^k \times 10^{-k}$

- Accumulation point

Given a sequence $\{\mathbf{x}_k\} \subseteq \mathbb{R}^n$, we call $\mathbf{x}^*$ an accumulation point of $\{\mathbf{x}_k\}$ if for any $\varepsilon > 0$, there exists $k_j$ such that $\|\mathbf{x}_{k_j} - \mathbf{x}^*\| < \varepsilon$.

Examples: $x_k = (-1)^k(1 + 10^{-k})$, $x_k = (k \mod 4) + (-1)^k/k$

# Analysis in R$^n$

- Bolzano-Weierstrass theorem

**Theorem 1.** *Any bounded sequence in* $\mathbb{R}^n$ *contains a convergent subsequence.*

Proof. Suppose that a sequence $\{\mathbf{x}_k\}$ is in a hyper-cube $\mathcal{H} = \{\mathbf{x} | l_i \leq x_i \leq u_i, i = 1, \cdots, n\}$. Each time we divide $\mathcal{H}$ in half, then at least one of the halves has infinite number of points in $\{\mathbf{x}_k\}$, denote it as $\mathcal{H}_j$. Then the diameter of $\mathcal{H}_j$ approaches 0 and $\mathbf{x}^* = \bigcap_{j=1}^{\infty} \mathcal{H}_j$ is an accumulation point. Picking any of the points in $\mathcal{H}_{j-1} \backslash \mathcal{H}_j$ gives a subsequence that converges to $\mathbf{x}^*$.

# Analysis in $R^n$

- Global and local convergence

We say that an iterative algorithm is *globally convergent* if for any arbitrary starting point the algorithm is guaranteed to generate a sequence of points converging to a point that satisfies the (FONC) for a minimizer. When the algorithm is not globally convergent, it may still generate a sequence that converges to a point satisfying the FONC, provided that the initial point is sufficiently close to the point. In this case we say that the algorithm is *locally convergent*.

# Analysis in R$^n$

- Convergence rate

Assume $\mathbf{x}_k \to \mathbf{x}^*$. We define the sequence of errors to be

$$\mathbf{e}_k = \mathbf{x}_k - \mathbf{x}^*.$$

We say that the sequence $\{\mathbf{x}_k\}$ converges to $\mathbf{x}^*$ with rate $r$ and rate constant $C$ if

$$\lim_{k\to\infty} \frac{\|\mathbf{e}_{k+1}\|}{\|\mathbf{e}_k\|^r} = C, \quad (C < \infty).$$

Linear: $r = 1$, $0 < C < 1$; Q-linear

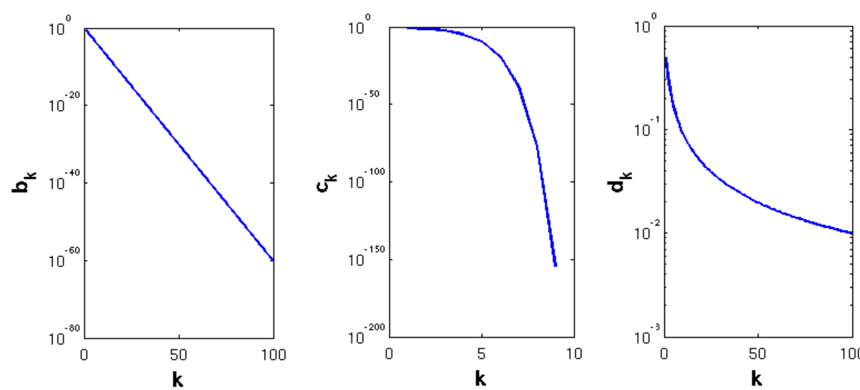
Sublinear: $r = 1$, $C = 1$;

Superlinear: $r = 1$, $C = 0$;

Quadratic: $r = 2$;

Cubic: $r = 3$. $r$ may be non-integers

Can $r = 1$ and $C > 1$ happen?




Examples: $x_k = 10^{-k}$, $x_k = 0.99^k$, $x_k = 10^{-2^k}$, $x_{k+1} = x_k/2 + 2/x_k$ $(x_1 = 4)$,

# Analysis in R$^n$

- Convergence rate

Estimating the order $r$:

$$r \approx \frac{\log \dfrac{x_{k+1} - x_k}{x_k - x_{k-1}}}{\log \dfrac{x_k - x_{k-1}}{x_{k-1} - x_{k-2}}}.$$

Assume $\mathbf{x}_k \to \mathbf{x}^*$. We say that the sequence $\{\mathbf{x}_k\}$ converges to $\mathbf{x}^*$ $R$-linearly if

$$\|\mathbf{x}_k - \mathbf{x}^*\| \leq e_k$$

R-linear

and $\{e_k\}$ converges to 0 $Q$-linearly.

Remedies the issue when $\lim\limits_{k\to\infty} \dfrac{\|\mathbf{e}_{k+1}\|}{\|\mathbf{e}_k\|^r}$ does not exist.

Example: $x_k = \begin{cases} 1 + 2^{-k}, & k \text{ even}, \\ 1, & k \text{ odd}. \end{cases}$

# Analysis in R$^n$

- Continuity

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *continuous* at $\mathbf{x} \in \operatorname{dom} f$ if for all $\varepsilon > 0$ there exists a $\delta$ such that

$$\mathbf{y} \in \operatorname{dom} f, \;\; \|\mathbf{y} - \mathbf{x}\|_2 \le \delta \Rightarrow \|f(\mathbf{y}) - f(\mathbf{x})\|_2 \le \varepsilon.$$

Continuity can be described in terms of limits: whenever the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ in $\operatorname{dom} f$ converges to a point $\mathbf{x} \in \operatorname{dom} f$, the sequence $f(\mathbf{x}_1), f(\mathbf{x}_2), \ldots$ converges to $f(\mathbf{x})$, *i.e.*,

$$\lim_{i\to\infty} f(\mathbf{x}_i) = f(\lim_{i\to\infty} \mathbf{x}_i).$$

A function $f$ is continuous if it is continuous at every point in its domain.

# Analysis in R$^n$

- Minimum and minimal

A point $\mathbf{x}^*$ is called a *minimum point* of a function $f(\mathbf{x})$ if

$$f(\mathbf{x}^*) \leq f(\mathbf{x}), \quad \forall \mathbf{x} \in \operatorname{dom} f.$$

Accordingly, $f(\mathbf{x}^*)$ is called the *minimum value* of $f$.

$\mathbf{x}^*$ is called a *minimal point* of $f$ if for sufficiently small $\varepsilon > 0$

$$f(\mathbf{x}^*) \leq f(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{B}_\varepsilon(\mathbf{x}^*) \cap \operatorname{dom} f.$$

Accordingly, $f(\mathbf{x}^*)$ is called the *minimal value* of $f$.

# Analysis in R$^n$

- Closedness

A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *closed* if, for each $\alpha \in \mathbb{R}$, the sublevel set

$$\{\mathbf{x} \in \operatorname{dom} f | f(\mathbf{x}) \leq \alpha\}$$

is closed. This is equivalent to the condition that the epigraph of $f$,

$$\operatorname{epi} f = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} | \mathbf{x} \in \operatorname{dom} f, f(\mathbf{x}) \leq t\},$$

is closed.

If $f : \mathbb{R}^n \to \mathbb{R}$ is continuous, and $\operatorname{dom} f$ is closed, then $f$ is closed. If $f : \mathbb{R}^n \to \mathbb{R}$ is continuous, with $\operatorname{dom} f$ open, then $f$ is closed iff $f$ converges to $\infty$ along every sequence converging to a boundary point of $\operatorname{dom} f$. In other words, if $\lim_{i\to\infty} \mathbf{x}_i = \mathbf{x} \in \partial(\operatorname{dom} f)$, with $\mathbf{x}_i \in \operatorname{dom} f$, we have $\lim_{i\to\infty} f(\mathbf{x}_i) = \infty$.

Examples: $f(x) = x \log x$ with $\operatorname{dom} f = \mathbb{R}_{++}$; $f(x) = -\log x$ with $\operatorname{dom} f = \mathbb{R}_{++}$; $f(x) = \begin{cases} x \log x, & x > 0 \\ 0, & x = 0, \end{cases}$ with $\operatorname{dom} f = \mathbb{R}_+$

# Analysis in $R^n$

- Derivative

Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$, and $\mathbf{x} \in (\mathrm{dom}\, f)^\circ$. If there exists a matrix $\mathbf{J}$ such that

$$\lim_{\mathbf{z}\in \mathrm{dom}\, f, \mathbf{z}\neq\mathbf{x}, \mathbf{z}\to\mathbf{x}} \frac{\|f(\mathbf{z}) - f(\mathbf{x}) - \mathbf{J}(\mathbf{z}-\mathbf{x})\|_2}{\|\mathbf{z}-\mathbf{x}\|_2} = 0,$$

for all choice of sequence $\{\mathbf{z}\} \subset \mathrm{dom}\, f$, then $f$ is said to be differentiable at $\mathbf{x}$ and denote $Df(\mathbf{x}) = \mathbf{J}$. Let $\mathbf{z} = \mathbf{x} + t\mathbf{e}_i$ and let $t \to 0$. Then

$$\begin{aligned}
\lim_{\mathbf{z}\in \mathrm{dom}\, f, \mathbf{z}\neq\mathbf{x}, \mathbf{z}\to\mathbf{x}} \frac{\|f(\mathbf{z}) - f(\mathbf{x}) - \mathbf{J}(\mathbf{z}-\mathbf{x})\|_2}{\|\mathbf{z}-\mathbf{x}\|_2} &= \lim_{t\to 0} \frac{\|f(\mathbf{x}+t\mathbf{e}_i) - f(\mathbf{x}) - t\mathbf{J}\mathbf{e}_i\|_2}{|t|} \\
&= \lim_{t\to 0} \left\| \frac{f(\mathbf{x}+t\mathbf{e}_i) - f(\mathbf{x})}{t} - \mathbf{J}\mathbf{e}_i \right\|_2 \\
&= \left\| \frac{\partial f(\mathbf{x})}{\partial x_i} - \mathbf{J}\mathbf{e}_i \right\|_2 .
\end{aligned}$$

Jacobian

Therefore, the $i$-th column of $\mathbf{J}$ is $\frac{\partial f(\mathbf{x})}{\partial x_i}$. Thus $\mathbf{J} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}^T} = \left(\frac{\partial f_i(\mathbf{x})}{\partial x_j}\right)$.

# Homework (2)

1. Judge the properties of the following sets (openess, closeness, boundedness, compactness) and give their interiors, closures, boundaries, and accumulation points:

a. $\mathcal{C}_1 = \emptyset$.

b. $\mathcal{C}_2 = \mathbb{R}^n$.

c. $\mathcal{C}_3 = \{x|0 \leq x < 1\} \cup \{x|2 \leq x \leq 3\} \cup \{x|4 < x \leq 5\}$.

d. $\mathcal{C}_4 = \{(x, y)^T|x \geq 0, y > 0\}$.

e. $\mathcal{C}_5 = \{k|k \in \mathbb{Z}\}$.

f. $\mathcal{C}_6 = \{k^{-1}|k \in \mathbb{Z}\}$.

g. $\mathcal{C}_7 = \{(1/k, \sin k)^T|k \in \mathbb{Z}\}$.

2. Prove that a set $\mathcal{C} \subseteq \mathbb{R}^n$ is closed iff (aka. if and only if) it contains the limit point of every convergent sequence in it.

# Homework (2)

3. Prove that a point $\mathbf{x}$ is a boundary point of $\mathcal{C} \subseteq \mathbb{R}^n$ iff for $\forall \epsilon > 0$, there exists $\mathbf{y} \in \mathcal{C}$ and $\mathbf{z} \notin \mathcal{C}$ such that

$$\|\mathbf{y} - \mathbf{x}\|_2 \leq \epsilon, \quad \|\mathbf{z} - \mathbf{x}\|_2 \leq \epsilon.$$

4. Prove that $\mathcal{C} \subseteq \mathbb{R}^n$ is closed iff it contains its boundary, and is open iff it contains no boundary points.

5. Prove the following:

   a. $\overline{\mathcal{A} \cup \mathcal{B}} = \overline{\mathcal{A}} \cup \overline{\mathcal{B}}$; $\overline{\mathcal{A} \cap \mathcal{B}} \subseteq \overline{\mathcal{A}} \cap \overline{\mathcal{B}}$. Give an example showing that $\overline{\mathcal{A} \cap \mathcal{B}} \neq \overline{\mathcal{A}} \cap \overline{\mathcal{B}}$.

   b. $(\overline{\mathcal{A} \cap \mathcal{B}})^\circ = \mathcal{A}^\circ \cap \mathcal{B}^\circ$; $(\overline{\mathcal{A} \cup \mathcal{B}})^\circ \supseteq \mathcal{A}^\circ \cup \mathcal{B}^\circ$. Give an example showing that $(\overline{\mathcal{A} \cup \mathcal{B}})^\circ \neq \mathcal{A}^\circ \cup \mathcal{B}^\circ$.

# Homework (2)

6. For each of the following sequences, determine the rate of convergence and the rate constant.

   a. $x_k = 2^{-k}$, for $k = 1, 2, \cdots$.

   b. $x_k = 1 + 5 \times 10^{-2k}$, for $k = 1, 2, \cdots$.

   c. $x_k = 2^{-2^k}$.

   d. $x_k = 3^{-k^2}$.

   e. $x_k = 1 - 2^{-2^k}$ for $k$ odd, and $x = 1 + 2^{-k}$ for $k$ even.

7. Let $\{x_k\}$ and $\{c_k\}$ be convergent sequences, and assume that

$$\lim_{k\to\infty} c_k = c \neq 0.$$

Consider the sequence $\{y_k\}$ with $y_k = c_k x_k$. Can its convergence rate and rate constant be determined from those of $\{x_k\}$ and $\{c_k\}$?