Topology and calculus background

Mauro Passacantando

Department of Computer Science, University of Pisa
mauro.passacantando@unipi.it

Numerical Methods and Optimization
Master in Computer Science – University of Pisa

Euclidean space $\mathbb{R}^n$

$$\mathbb{R}^n := \{(x_1, x_2, \dots, x_n) : \ x_i \in \mathbb{R} \text{ for any } i = 1, \dots, n\}$$

Given $x, y \in \mathbb{R}^n$,

$$x + y := (x_1 + y_1, \dots, x_n + y_n).$$

Given $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$,

$$\alpha\, x := (\alpha x_1, \dots, \alpha x_n).$$

Euclidean distance

The Euclidean distance between x and y is

$$d(x,y) := \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$$

Properties:

- $d(x,y) \geq 0$ for all $x, y \in \mathbb{R}^n$
- $d(x,y) = 0$ if and only if $x = y$
- $d(\alpha x, 0) = |\alpha| d(x,0)$ for all $x \in \mathbb{R}^n, \alpha \in \mathbb{R}$
- $d(x,y) \leq d(x,z) + d(z,y)$ for all $x, y, z \in \mathbb{R}^n$ (triangle inequality)

Euclidean norm

The Euclidean norm of a vector $x \in \mathbb{R}^n$ is

$$\|x\| := \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + \cdots + x_n^2}$$

Properties:

- $\|x\| \geq 0$ for all $x \in \mathbb{R}^n$
- $\|x\| = 0$ if and only if $x = 0$
- $\|\alpha x\| = |\alpha| \|x\|$ for all $x \in \mathbb{R}^n, \alpha \in \mathbb{R}$
- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathbb{R}^n$

Scalar product

The scalar product between x and y is

$$\langle x, y \rangle := \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \cdots + x_n y_n$$

Properties:

- $d(x, y) = \|x - y\|$ for all $x, y \in \mathbb{R}^n$
- $\|x\| = \sqrt{\langle x, x \rangle}$ for all $x \in \mathbb{R}^n$
- $\langle x, y \rangle = \|x\| \|y\| \cos\theta$, where θ is the angle between x and y.
  (x and y are said orthogonal if $\langle x, y \rangle = 0$)
- $|\langle x, y \rangle| \leq \|x\| \|y\|$ for all $x, y \in \mathbb{R}^n$ (Cauchy-Schwarz inequality)
- $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle$

Interior and boundary points

Given $x \in \mathbb{R}^n$ and $r > 0$ the open ball with center x and radius r is

$$B(x, r) := \{y \in \mathbb{R}^n : \ \|y - x\| < r\}.$$

Let $D \subseteq \mathbb{R}^n$.
A point $\bar{x} \in D$ is an interior point of D if there exists $r > 0$ such that $B(\bar{x}, r) \subseteq D$.
A point $\bar{x} \in \mathbb{R}^n$ is an boundary point of D if for any $r > 0$ there exist $x, y \in B(\bar{x}, r)$ such that $x \in D$ and $y \notin D$.

The interior of D, denoted by $\text{int}(D)$, is the set of interior points of D.
The boundary of D, denoted by ∂D, is the set of boundary points of D.
The closure of D, denoted by $\text{cl}(D)$, is the union of $\text{int}(D)$ and ∂D.

Open sets

A set $D \subseteq \mathbb{R}^n$ is open if any point of D is an interior point, i.e., $D = \text{int}(D)$.

Properties:

- If $\{D_i\}_{i \in I}$ is a family of open sets (possibily infinite), then $\bigcup_{i \in I} D_i$ is open.
- If D_1 and D_2 are open, then $D_1 \cap D_2$ is open.

Exercise. If $\{D_i\}_{i \in I}$ is a family of open sets (possibly infinite), then is $\bigcap_{i \in I} D_i$ open?

Closed sets

A set $D \subseteq \mathbb{R}^n$ is closed if its complement

$$\mathbb{R}^n \setminus D := \{x \in \mathbb{R}^n : \ x \notin D\}$$

is open or, equivalently, if $D = \mathrm{cl}(D)$.

Properties:

- If $\{D_i\}_{i \in I}$ is a family of closed sets (possibly infinite), then $\bigcap_{i \in I} D_i$ is closed.
- If D_1 and D_2 are closed, then $D_1 \cup D_2$ is closed.

Exercise. If $\{D_i\}_{i \in I}$ is a family of closed sets (possibly infinite), then is $\bigcup_{i \in I} D_i$ closed?

Bounded sets

Bounded set

A set $D \subseteq \mathbb{R}^n$ is bounded if there exists $r > 0$ such that $D \subseteq B(0, r)$.

A sequence $\{x^k\} \subseteq \mathbb{R}^n$ converges to a point $\bar{x}$ if $\lim_{k \to \infty} \|x^k - \bar{x}\| = 0$.

Bolzano-Weierstrass Theorem

If $D \subseteq \mathbb{R}^n$ is closed and bounded, then any sequence $\{x^k\} \subseteq D$ has a subsequence $\{x^{k_j}\}$ which converges to a point $\bar{x} \in D$.

Continuous functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is continuous in $\bar{x} \in \mathbb{R}^n$ if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(\bar{x})| < \varepsilon$ for any $x \in B(\bar{x}, \delta)$.
f is continuous on a set $D \subseteq \mathbb{R}^n$ if it is continuous in any point $\bar{x} \in D$.

Theorem

If $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are continuous in $\bar{x}$, then

- αf is continuous in $\bar{x}$ for any $\alpha \in \mathbb{R}$
- $|f|$ is continuous in $\bar{x}$
- $f + g$ are continuous in $\bar{x}$
- fg is continuous in $\bar{x}$

Partial derivatives

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $\bar{x} \in \mathbb{R}^n$.
If

$$\lim_{t\to 0} \frac{f(\bar{x}_1, \ldots, \bar{x}_{i-1}, \bar{x}_i + t, \bar{x}_{i+1}, \ldots, \bar{x}_n) - f(\bar{x})}{t}$$

exists and it is finite, then it is called the partial derivative of f with respect to the variable x_i at $\bar{x}$ and it is denoted by $\frac{\partial f}{\partial x_i}(\bar{x})$.

The vector of all partial derivatives of f at $\bar{x}$ is the gradient of f at $\bar{x}$:

$$\nabla f(\bar{x}) := \left(\frac{\partial f}{\partial x_1}(\bar{x}), \ldots, \frac{\partial f}{\partial x_n}(\bar{x}) \right)$$

Exercise. Compute the gradient of $f(x_1, x_2) = x_1^2 x_2$ at $\bar{x} = (2, 3)$.

Directional derivatives

Let $f : \mathbb{R}^n \to \mathbb{R}$, $\bar{x} \in \mathbb{R}^n$, direction $d \in \mathbb{R}^n$.
If

$$\lim_{t \to 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t}$$

exists and it is finite, then it is called the directional derivative of f along d at $\bar{x}$ and it is denoted by $\dfrac{\partial f}{\partial d}(\bar{x})$.

Exercise. Compute the directional derivative of $f(x_1, x_2) = x_1^2 x_2$ along the direction $d = (d_1, d_2)$ at $\bar{x} = (2, 3)$.

Derivability and continuity

Case $n = 1$: if f has derivative at $\bar{x}$ then it is continuous at $\bar{x}$.
Case $n > 1$: if f has directional derivatives at $\bar{x}$, then it is not necessarily continuous, as the following example shows.

Example.

$$f(x_1, x_2) = \begin{cases} \left(\frac{x_1^2 x_2}{x_1^4 + x_2^2}\right)^2 & \text{if } (x_1, x_2) \neq (0,0) \\ 0 & \text{if } (x_1, x_2) = (0,0) \end{cases}$$

f has directional derivatives along any direction d at $(0,0)$:

$$\frac{\partial f}{\partial d}(0,0) = \lim_{t \to 0} \frac{f(td_1, td_2)}{t} = \lim_{t \to 0} \frac{t d_1^4 d_2^2}{(t^2 d_1^4 + d_2^2)^2} = 0,$$

but it is not continuous at $(0,0)$ since

$$\lim_{k \to +\infty} f\left(\frac{1}{k}, \frac{1}{k^2}\right) = \frac{1}{4} \neq 0 = f(0,0).$$

Differentiability

f is differentiable at $\bar{x}$ if there exists a linear function $\phi : \mathbb{R}^n \to \mathbb{R}$ such that

$$\lim_{h \to 0} \frac{f(\bar{x}+h) - f(\bar{x}) - \phi(h)}{\|h\|} = 0.$$

ϕ is called the differential of f at $\bar{x}$ and is denoted by $df_{\bar{x}}$.

Theorem

If f is differentiable at $\bar{x}$, then f is continuous at $\bar{x}$ and has directional derivatives at $\bar{x}$ along any direction d with

$$\frac{\partial f}{\partial d}(\bar{x}) = df_{\bar{x}}(d) = \langle \nabla f(\bar{x}), d \rangle.$$

Theorem

If f has partial derivatives at any point of $B(\bar{x}, \delta)$ for some $\delta > 0$ and such derivatives are continuous, then f is differentiable at $\bar{x}$.

Second order partial derivatives

If there exist the partial derivative with respect to x_j of $\frac{\partial f}{\partial x_i}$, we have the second order partial derivative

$$\frac{\partial^2 f}{\partial x_j \partial x_i}.$$

The matrix

$$\nabla^2 f(\bar{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\bar{x}) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\bar{x}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\bar{x}) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\bar{x}) & \frac{\partial^2 f}{\partial x_2^2}(\bar{x}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_2}(\bar{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\bar{x}) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\bar{x}) & \dots & \frac{\partial^2 f}{\partial x_n^2}(\bar{x}) \end{pmatrix}$$

is called Hessian matrix of f at $\bar{x}$.

Exercise. Compute the Hessian matrix of $f(x_1, x_2) = x_1^2 x_2$.

Second order partial derivatives

Schwartz Theorem

If the mixed second order partial derivatives $\frac{\partial^2 f}{\partial x_j \partial x_i}(x)$ and $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ exist for any $x \in B(\bar{x}, \delta)$ for some $\delta > 0$ and are continuous at $\bar{x}$, then

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{x}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}). \tag{1}$$

Remark. The assumption of continuity of the mixed derivatives is necessary to get (1). In fact, if we consider

$$f(x_1, x_2) = \begin{cases} \dfrac{x_2^3}{x_1^2 + x_2^2} & \text{if } x_2 \neq 0, \\ 0 & \text{if } x_2 = 0, \end{cases}$$

we obtain $\frac{\partial^2 f}{\partial x_2 \partial x_1}(0,0) = 1$ and $\frac{\partial^2 f}{\partial x_1 \partial x_2}(0,0) = 0$.

Taylor's formulas

Assume that first and second derivatives of $f : \mathbb{R}^n \to \mathbb{R}$ exist and are continuous in $B(\bar{x}, \delta)$ for some $\delta > 0$.

First order Taylor's formula

$$f(\bar{x} + h) = f(\bar{x}) + \langle \nabla f(\bar{x}), h \rangle + R_1(h),$$

where the remainder $R_1(h)$ is such that $\lim_{h \to 0} \frac{R_1(h)}{\|h\|} = 0$.

Second order Taylor's formula

$$f(\bar{x} + h) = f(\bar{x}) + \langle \nabla f(\bar{x}), h \rangle + \frac{1}{2} \langle h, \nabla^2 f(\bar{x}) h \rangle + R_2(h)$$

where the remainder $R_2(h)$ is such that $\lim_{h \to 0} \frac{R_2(h)}{\|h\|^2} = 0$.