Let us relate the structures from the last session to $\mathbb{R}^n$:

We denote $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$. Then:

(or dot product)

- The inner product (or scalar product) is $\langle x, y \rangle \equiv x \cdot y := \sum_{i=1}^{n} x_i y_i$.

  (Recall that $\langle x, y \rangle = \|x\| \, \|y\| \underbrace{\cos(\sphericalangle x, y)}_{\text{angle between } x \text{ and } y}$.)

- It induces the norm $\|x\|_2 = (\langle x, x \rangle)^{\frac{1}{2}} = \left( \sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}$.
- Also $\|x\|_p := \left( \sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$ define norms $\forall\, 1 \le p < \infty$ and the following Hölder inequality holds:
  $|\langle x, y \rangle| \le \|x\|_p \, \|y\|_q$.
- Also $\|x\|_\infty := \max_{1 \le i \le k} |x_i|$ defines a norm. (In fact, $\lim_{p \to \infty} \|x\|_p = \|x\|_\infty$.)

  $\|\cdot\|_A$ and $\|\cdot\|_B$ are any two norms here
- In fact, all norms on $\mathbb{R}^n$ are equivalent, i.e., $\forall x \; \exists C_1, C_2 > 0$ s.t. $C_1 \|x\|_A \le \|x\|_B \le C_2 \|x\|_A$

Finally, we briefly discuss the notion of compactness.

Recall that we call a subset of $\mathbb{R}^n$ open if it can be written as the union of open balls

$$B_r(x) := \{ y \in \mathbb{R}^n : \underbrace{\|x - y\|}_{\text{unless specified otherwise, we always mean } \|x\|^2 := \sum_{i=1}^{n} x_i^2 \text{ for } x \in \mathbb{R}^n} < r \}.$$

A set $E \subset \mathbb{R}^n$ is called compact if every open cover of $E$ has a finite subcover.

$\hookrightarrow$ A family of open sets $(V_\alpha)$ such that $\bigcup_\alpha V_\alpha \supset E$.

$\hookrightarrow$ A subfamily of the open cover with finitely many elements.

Important result: As in $\mathbb{R}$, the Heine-Borel theorem also holds in $\mathbb{R}^n$:

$$E \subset \mathbb{R}^n \text{ is compact} \iff E \underbrace{\text{closed}}_{\substack{\text{it contains all} \\ \text{its limit points}}} \text{ and } \underbrace{\text{bounded}}_{E \subset B_r(x) \text{ for some } r>0,\ x \in \mathbb{R}^n}.$$

This implies, e.g., that continuous functions $E \to \mathbb{R}$, $\mathbb{R}^n \supset E$ compact, attain their maximum and minimum.

<u>Ex.:</u>
- $\mathbb{R}^n$ is not compact, since it is not bounded ($\mathbb{R}^n$ is closed (and open)).
- $B_r(x)$ is not compact, since it is not closed ($B_r(x)$ is bounded).
- $\overline{B_r(x)} := \{ y \in \mathbb{R}^n : \|x-y\| \leq r \}$ is closed and bounded, and thus compact.

# 2. Derivatives

## 2.1 Total and Partial Derivatives

Some notation:

- We write vectors $x \in \mathbb{R}^n$ as $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.
- Special vectors are the basis vectors $e_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow$ j-th component, i.e., $x = \sum_{j=1}^{n} x_j e_j$

Recall from linear Algebra:

- A map $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear if $L(\lambda x + y) = \lambda L(x) + L(y) \quad \forall x, y \in \mathbb{R}^n, \lambda \in \mathbb{R}$.

  For linear maps we usually write $L(x) = Lx$.

- linear maps $L : \mathbb{R}^n \to \mathbb{R}^m$ are in one-to-one correspondence to $m \times n$ matrices

$$A_L = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$ by choosing a basis.

Choosing a basis $(e_j)$ of $\mathbb{R}^n$ and $(\tilde{e}_i)$ of $\mathbb{R}^m$, we have $(Lx)_i = \langle \tilde{e}_i, Lx \rangle = \langle \tilde{e}_i, L(\sum_j x_j e_j) \rangle$
$= \sum_j x_j \underbrace{\langle \tilde{e}_i, L e_j \rangle}_{=: a_{ij}}$.

Recall $(Ax)_i = \sum_{j=1}^{n} a_{ij} x_j$, $(AB)_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$ for $A$ $m \times n$, and $B$ $n \times p$ matrix.

(then $AB$ is an $m \times p$ matrix)

- For linear maps $\mathbb{R}^n \to \mathbb{R}^m$ we define the operator norm $\|L\| := \sup_{\substack{u \in \mathbb{R}^n \\ \|u\| = 1}} \|Lu\| < \infty$

since unit sphere is compact and $L$ linear (thus continuous), the maximum is attained (for at least one $u \in \mathbb{R}^n$)

Since $\|L \frac{u}{\|u\|}\| \leq \|L\|$, we have $\|Lu\| \leq \|L\| \, \|u\|$.

Recall that for functions $f: \mathbb{R} \to \mathbb{R}$ we defined differentiability at $\tilde{x}$ as:

$\exists m \in \mathbb{R}$ s.t. for small enough $h$: $f(\tilde{x}+h) = f(\tilde{x}) + mh + r_x(h)$, with $\lim_{h\to 0} \left|\frac{r_x(h)}{h}\right| = 0$.

Clearly, $L_m: \mathbb{R} \to \mathbb{R}$, $h \mapsto mh$ is a linear map.

The idea "derivatives are the best linear approximation" can be generalized:

<u>Definition:</u> Let $U \subset \mathbb{R}^n$ be open and $f: U \to \mathbb{R}^m$. Then $f$ is called differentiable at $\tilde{x} \in U$ if there is a linear map $A: \mathbb{R}^n \to \mathbb{R}^m$ s.t.

$$\underbrace{\boxed{f(\tilde{x}+h) = f(\tilde{x}) + Ah + r_{\tilde{x}}(h) \qquad \text{with} \quad \lim_{h\to 0} \frac{\|r_{\tilde{x}}(h)\|}{\|h\|} = 0.}}$$

In other words: $\lim_{h\to 0} \dfrac{\|f(\tilde{x}+h) - f(\tilde{x}) - Ah\|}{\|h\|} = 0.$

We call $A \equiv Df|_{\tilde{x}} \equiv f'(\tilde{x})$ the $\boxed{\text{total derivative}}$ of $f$ at $\tilde{x}$.

If $f$ is differentiable for all $\tilde{x} \in U$, we say $f$ is differentiable in $U$.