1 21st of September 2018 — A. Frangioni

1.1 Mathematical background for optimization problems

Definition 1.1 (Minimum problem)**.** *Let X be a set, called **feasible region** and let $f : X \to \mathbb{R}$ be* any *function, called **objective function** we call **problem** the following*

$$(P) \qquad f_* = \min\{ f(x) \ : x \in X\}$$

Definition 1.2 (Feasible solution)**.** *Let $x \in F$ be a solution of the minumim problem in which the domain is a superset of $X \subset F$. We say that x is a **feasible solution** if $x \in X$. On the other hand, x is **unfeasible** if $x \in F \setminus X$.*

Definition 1.3 (Optimal solution)**.** *Under the same hypothesis of the above definition, we define x_* such that $f(x_*) = f_*$ an **optimal solution**, where $f_* \leq f(x) \, \forall x \in X$, $\forall v > f_* \, \exists \, x \in X$ s.t. $f(x) < v$.*

It is possible to find problems where there is no optimal solution at all.

Example 1.1. *There are two cases in which it is not possible to find an optimal solution:*

1. *The domain is empty, which may be not trivial to prove, since it is an NP-hard problem sometimes;*

2. *We want to find the minimum of the objective function but it is unbounded below ($\forall M \, \exists x_M \in X$ s.t. $f(x_M) \leq M$). On the other hand, we need to maximize the function, but it is unbounded above.*

We can now rewrite the problem of solving an optimization problem as:

1. Finding x_* and proving it is optimal

2. Or proving $X = \emptyset$

3. Or constructively prove $\forall M \, \exists x_M \in X$ s.t. $f(x_M) \leq M$.

Most of the times we consider optimal a solution which is close to the true optimal value, modulo some error.

Definition 1.4 (Absolute error)**.** *We call **absolute error** the gap between the real value and the one we obtained. Formally,*

$$f(\bar{x}) - f_* \leq \varepsilon$$

Definition 1.5 (Relative error)**.** *We define as **relative error** the absolute error, normalized by the true value of the function*

$$(\, f(\bar{x}) - f_* \,)/|\, f_* \,| \leq \varepsilon$$

Let us consider an iterative algorithm that moves towards the optimum. It may happen that the function decreases and decreases along a certain direction but its non-continuity leads to the impossibility of reaching the optimum. As an example, let us take the following

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 1 & \text{if } x = 0 \end{cases}$$

Definition 1.6 (Totally ordered set)**.** *We say that set X is* ***totally ordered*** *if $\forall\, x, y \in X$, either $f(x) \leq f(y)$ or $f(y) \leq f(x)$.*

Definition 1.7 (Infima and suprema)**.** *Given a totally ordered set R and one of its subsets (say $S \subseteq R$)*

$$s \textit{ is the } \textbf{infimum} \textit{ of } S \Leftrightarrow \underline{s} = \inf S \quad \Leftrightarrow \quad \underline{s} \leq s \;\; \forall s \in S \;\; \wedge \;\; \forall t > \underline{s}\; \exists\, s \in S \textit{ s.t. } s \leq t$$

$$s \textit{ is the } \textbf{supremum} \textit{ of } S \Leftrightarrow \bar{s} = \sup S \quad \Leftrightarrow \quad \bar{s} \geq s \;\; \forall s \in S \;\; \wedge \;\; \forall t < \bar{s}\; \exists\, s \in S \textit{ s.t. } s \geq t$$

What happens if we have more than one objective function? We are provided with two tools in order to reduce them into one:

Scalarization: using a linear combination of the two functions: $f(x) = \alpha f_1(x) + \beta f_2(x)$;

Budgeting: $f(x) = f_1(x)$, $X := X \cup \{\, f_2(x) \leq b \,\}$, which intuitively corresponds to taking into account only one objective function, provided that the values of the othr functions are not too high.

Definition 1.8 (Extended real)**.** *In the case of unbounded functions the value of infima or suprema are ∞, and we call* ***extended reals*** *$\overline{\mathbb{R}} = -\infty \cup \mathbb{R} \cup +\infty$.*

We are interested in studying sequences, because iterative methods start from a certain point and move towards the optimal, hopefully.

Definition 1.9 (Limit)**.** *Given a sequence $\{x_i\}$ the* ***limit*** *for $i \to \infty$ is defined as*

$$\lim_{i \to \infty} v_i = v \quad \Longleftrightarrow \quad \forall \varepsilon > 0 \; \exists \; h \textit{ s.t. } |v_i - v| \leq \varepsilon \; \forall i \geq h$$

It may happen that a sequence has or does not have a limit. For example $\{\frac{1}{n}\}$ has limit 0 for $n \to +\infty$, while $\{(-1)^n\}$ does not.

Fact 1.1. *Let us be given a monotone sequence, then the sequence* ***does*** *have a limit.*

Notice that given a sequence either it is monotone or it can be "split" into two monotone sequences (for example $\{(-1)^n\}$ can be transformed into $\{(-1)^{2n}\}$ and $\{(-1)^{2n+1}\}$ and these two sequences are both monotone).

Definition 1.10 (Euclidean vector space)**.** *We call* ***Euclidean space***

$$\mathbb{R}^n := \{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \ : \ x_i \in \mathbb{R}, \ i = 1, \dots, n\}$$

Equivalently, we can characterize the Euclinean space as Cartesian product of $\mathbb{R}$ n *times:* $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \dots \mathbb{R}$.

The main operations on elements of the Euclidean space (vectors) are:

Sum: $x + y := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$

Scalar multiplication: $\alpha x = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}$

In order to be able to compute limits in a vector space we need to use norms (see **??**).

Fact 1.2. *The norms on a vector space have the following properties:*

1. $\|x\| \geq 0$ *and* $\forall x \in \mathbb{R}^n$, $\|x\| = 0 \iff x = 0$;
2. $\|\alpha x\| = |\alpha| \, \|x\|, \ \forall x \in \mathbb{R}^n, \ \alpha \in \mathbb{R}$;
3. $\|x + y\| \leq \|x\| + \|y\|, \ \forall x, y \in \mathbb{R}^n$ *(triangle inequality).*

Definition 1.11 (Ball)**.** *We term* ***ball*** *centered in* $\bar{x}$ *and having* ε *as radius as the set of points that are close enough to* $x \in \mathbb{R}^n$*:* $B(\bar{x}, \varepsilon) = \{x \in \mathbb{R}^n \ : \ \|x - \bar{x}\| \leq \varepsilon\}$.

In Figure 1.1 we may observe the different shapes of the same ball varying the value of p in the p-norm.

Definition 1.12 (Scalar product)**.** *Let* $x, y \in \mathbb{R}^n$ *we define the* ***scalar product*** *between these two vectors*

$$< x, y > := y^T x = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \dots + x_n y_n$$

Fact 1.3. *A scalar product has the following properties:*

1. $< x, y > = < y, x > \quad \forall x, y \in \mathbb{R}^n$ *(symmetry)*

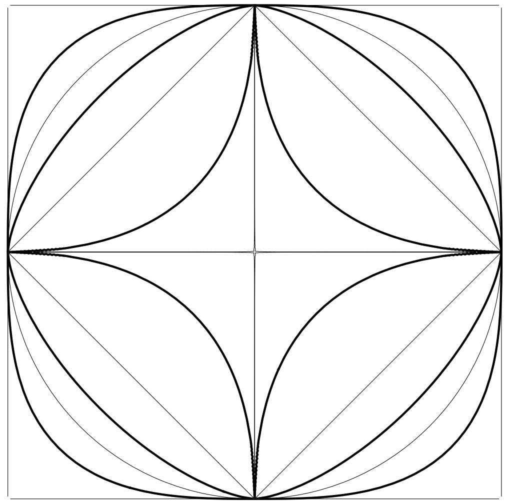

FIGURE 1.1: The shapes of balls centered in the origin of radius 1 varying the value of p.

2. $< x, x > \geq 0, \ \forall x \in \mathbb{R}^n, \ < x, x >= 0 \iff x = 0$;

3. $< \alpha x, y >= \alpha < x, y >, \ \forall x \in \mathbb{R}^n, \alpha \in \mathbb{R}$;

4. $< x + y, z >=< x, z > + < y, z >, \ \forall x, y, z \in \mathbb{R}^n$.

Fact 1.4 (Cauchy-Schwartz inequality)**.** *Let $x, y \in \mathbb{R}^n$. The following holds:*

$$< x, y >^2 \leq < x, x >< y, y > \equiv |< x, y >| \leq \|x\| \, \|y\| \,, \ \forall x, y \in \mathbb{R}^n$$

An important characterization of the scalar product is the one that uses angles:
$< x, y >= \|x\| \, \|y\| \cos\theta$: $x \perp y \iff < x, y >= 0$ and $< x, y >> 0 \iff$ "x and y point in the same direction"

We have now all the tools to define the notion of limit of a sequence in $\mathbb{R}^n$.

Definition 1.13 (Limit of a sequence in the Euclidean space)**.** *Let $\{x_i\} \subset \mathbb{R}^n$ be a sequence in $\mathbb{R}^n$. The **limit** of $\{x_i\}$ for $i \to +\infty$ is the following:*

$$\lim_{i \to \infty} x_i = x \equiv \{x_i\} \to x$$

$$\Updownarrow$$

$$\forall \varepsilon > 0 \ \exists h \ s.t. \ d(x_i, x) \leq \varepsilon \ \forall i \geq h$$

$$\Updownarrow$$

$$\forall \varepsilon > 0 \ \exists h \ s.t. \ x_i \in \mathcal{B}(x, \varepsilon) \ \forall i \geq h$$

$$\Updownarrow$$

$$\lim_{i \to \infty} d(x_i, x) = 0$$