~~Week 10~~

Unit 11

Nonlinear Optimization

11.1 ~~10.1~~ Motivating Examples

In the following sections of this chapter, we concern ourselves with three broad classes of successively more difficult nonlinear optimization problems.

Moreover, when analyzing these problems it is always helpful to put them into a standard form.

By doing so, we are able to apply the same generic tools to solve these classes of NLPPs. These 'standard forms' are akin to the standard forms of LPPs

We also use these examples to demonstrate how problems can be converted to these three 'standard forms'.

10.2 Types of Nonlinear Optimization Problems

10.2.1 Unconstrained NLPPs

An **unconstrained NLPP** has an objective function that is being minimized, but does not have any constraints on what values the decision variables can take. This problem can be generically written as:

$$\min_{x \in \mathbb{R}^n} \quad f(x),$$

$f: \mathbb{R}^n \to \mathbb{R}$

C.R. es $\mathbb{R}^n$

where $f : \mathbb{R}^n \to \mathbb{R}$ is the objective function. The Facility-Location Problem in ... is an example of an unconstrained problem.

10.2.2 Equality-Constrained NLPPs

An **equality-constrained NLPP** has an objective function that is being minimized and a finite set of equality constraints that have zeros on their right-hand sides. This problem can be generically written as:

PPL

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in C, \end{aligned}$$

PPNL

f lineal $\wedge$ Todas las funciones de las restric. son lineales

f No lineal

where $C := \{x \in \mathbb{R}^n \ ; \ h_1(x) = 0, \ldots, h_p(x) = 0\}$ is the feasible set, $h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ are the equality-constraint functions, $p \in \{1, 2, \ldots\}$ and $f : \mathbb{R}^n \to \mathbb{R}$ is the objective function. The Hanging-Chain Problem in ... is an example of an equality-constrained problem.

✓ alguna de las f.... ... No es lineal

10.2.3 Equality- and Inequality-Constrained NLPPs

An **equality- and inequality-constrained NLPP** has an objective function that is being minimized and a finite set of equality and less-than-or-equal-to constraints that have zeros on their right-hand sides. This problem can be generically written as:

max x
s.a. x<1

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in C, \end{aligned}$$

$\boxed{x<1}$

where $C := \{x \in \mathbb{R}^n \ ; \ h_1(x) = 0, \ldots, h_p(x) = 0, g_1(x) \leq 0, \ldots, g_q(x) \leq 0\}$ is the feasible set, $h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ are the equality-constraint functions, $g_1, \ldots, g_q : \mathbb{R}^n \to \mathbb{R}$ are the less-than-or-equal-to-constraint functions, $p, q \in \{1, 2, \ldots\}$ and $f : \mathbb{R}^n \to \mathbb{R}$ is the objective function. To demonstrate how a problem can converted to the generic form, take as an example the Return-Maximization Problem ...

10.3 Convexity

10.3.1 Convex sets

Definition 10.1. A set $C \subseteq \mathbb{R}^n$ is **convex** if for every $a, b \in C$ and $t \in (0,1)$: $a + t(b-a)$ is in C. $(1-t)a + t \cdot b$ ♢

Example 10.2. Let $a \in \mathbb{R}^n$, let $d \in \mathbb{R}^n \setminus \{0\}$, let $c \in \mathbb{R}$ and let $r > 0$. The following sets are convex:

1. $\mathcal{L}(a,d) := \{a + td \; ; \; t \in \mathbb{R}\}$; La recta
2. $\mathcal{H}^-(d,c) := \{v \in \mathbb{R}^n \; ; \; \langle v, d\rangle \le c\}$; Semi-espacio
3. $\mathcal{H}(d,c) := \{v \in \mathbb{R}^n \; ; \; \langle v, d\rangle = c\}$; Hiperplano.
4. $B[a,r] := \{v \in \mathbb{R}^n \; ; \; |v-a| \le r\}$. Bola cerrada
5. $B(a,r) := \{v \in \mathbb{R}^n \; ; \; |v-a| < r\}$. " abierta
6. $\{(v_1, \dots, v_n) \in \mathbb{R}^n \; ; \; v_1 \ge 0, \dots, v_n \ge 0\}$.

♢

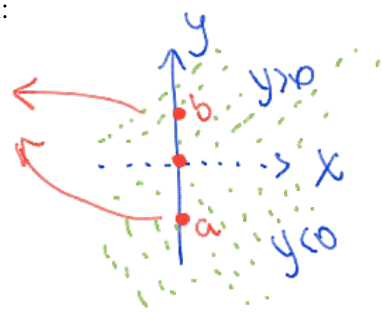
Indeed, let $x, y \in \mathbb{R}^n$ and $t \in (0,1)$.

1. $x, y \in \mathcal{L}(a,d)$. $x = a + t_1 \cdot d$, $y = a + t_2 \cdot d$ $\{$ $(1-t)\cdot x + t \cdot y$
   $(1-t)\cdot(a + t_1 \cdot d) + t \cdot (a + t_2 \cdot d)$
   $a + [(1-t)\cdot t_1 + t \cdot t_2]\cdot d$
2. $x, y \in \mathcal{H}^-(d,c)$ $\langle (1-t)x + ty, d\rangle = (1-t)\underbrace{\langle x, d\rangle}_{\le c} + t\underbrace{\langle y, d\rangle}_{\le c} \le (1-t)c + tc = c.$
3.
4. Luego $(1-t)\cdot\langle x, d\rangle \le (1-t)c$
5. $t \cdot \langle y, d\rangle \le t \cdot c$
6. $\le c$

$\therefore [(1-t)\cdot x + t \cdot y] \in \mathcal{H}^-(d,c)$ □

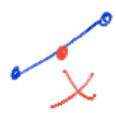
Example 10.3. Let $a \in \mathbb{R}^n$, let $d \in \mathbb{R}^n \setminus \{0\}$, let $c \in \mathbb{R}$ and let $r > 0$. The following sets are not convex:

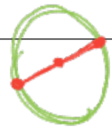
1. $\{v \in \mathbb{R}^n \; ; \; \langle v, d\rangle \ne c\}$;
   ||
   X



∀ a,b

∃ a,b





idea intuitive

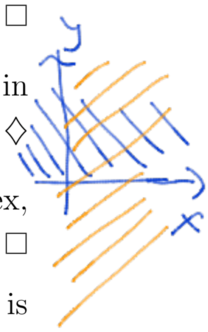
2. $S(a, r) := \{v \in \mathbb{R}^n \ ; \ |v - a| = r\}$;
3. $\{v \in \mathbb{R}^n \ ; \ |v - a| \geq r\}$.

$\diamondsuit$

Indeed:

1.
2. Let $b := (c, 0, 0, \dots, 0)$, let $\boxed{x} := a - b$, let $\boxed{y} := a + b$ and let $t := 1/2$. Then, $x, y \in S(a, r)$, $t \in (0, 1)$, but $\boxed{(1 - t)x + ty} = a$ is not in $S(a, r)$.
3.

$\square$



Property 10.4. Let $(C_j)_{j \in J}$ be a collection of convex sets of $\mathbb{R}^n$ indexed in an arbitrary set J. Then $C := \cap_{j \in J} C_j$ is a convex set. $\diamondsuit$

Proof. Let $a, b \in C$ and let $t \in (0, 1)$. Let $j \in J$. Then, as C_j is convex, $a + t(b - a)$ is in C_j. And, by the arbitrariness of j, $a + t(b - a)$ is in C. $\square$

Remark 10.5. By virtue of Property 10.4, we can easily prove that a set is convex by expressing this as a union of convex set, e.g., the first quadrant is convex since it is a finite intersection of semi-spaces, which by virtue of item 2 of Example 10.2 are convex sets. $\diamondsuit$

Property 10.6. Let $A, B \subseteq \mathbb{R}^p$ and $C \subseteq \mathbb{R}^q$ be convex sets and let $\alpha \in \mathbb{R}$. Then

1. $A + B$ is convex, $A+B := \{ a+b \ ; \ a \in A, \ b \in B \}$
2. αA is convex and $\alpha \cdot A := \{ \alpha \cdot a \ ; \ a \in A \}$
3. $A \times \cancel{B}\, C$ is convex. $A \times C := \{ (a,c) \ ; \ a \in A, \ c \in C \}$

(en $\mathbb{R}^p \times \mathbb{R}^q$)

$\diamondsuit$

Indeed, let $t \in (0, 1)$.

1.
2.

3. Let $x := (a^1, b^1), y := (a^2, b^2) \in A \times B$ with $a^1, a^2 \in A$ and $b^1, b^2 \in B$. Then, as A and B are convex,

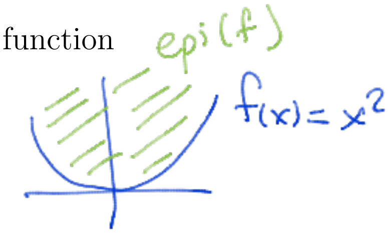
$$(1-t)x + ty = \left(\underbrace{(1-t)a^1 + ta^2}_{\in A}, \underbrace{(1-t)b^1 + tb^2}_{\in B} \right)$$

is in $A \times B$.

□

10.3.2 Convex functions

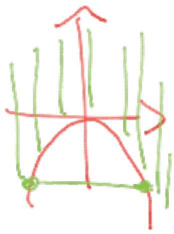
Definition 10.7. Let $C \subseteq \mathbb{R}^n$ be a convex set, let $f : C \to \mathbb{R}$ be a function and $\alpha \in \mathbb{R}$. Then the sets



$$\mathrm{epi}(f) := \{(x, \beta) \in C \times \mathbb{R} \; ; \; f(x) \leq \beta\} \quad \text{and}$$
$$S_\alpha(f) := \{x \in C \; ; \; f(x) \leq \alpha\}$$

epígrafo conjunto de nivel

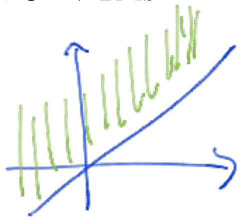
are called **epigraph** and **level sets** of f, respectively. ♢



The notion of convexity is essentially geometric and not analytic.

Definition 10.8. Let $C \subseteq \mathbb{R}^n$ be a convex set. We say that $f : C \to \mathbb{R}$ is **convex** if $\mathrm{epi}(f)$ is convex and **concave** if $-f$ is convex.

Example 10.9. Let $C = \mathbb{R}$ and $f : C \to \mathbb{R}$... son convexas



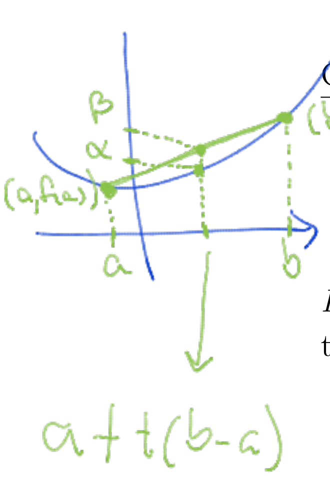
1. $f(x) = x$ es cóncava y convexa.
2. $f(x) = x^2$ No es cóncava, pero es convexa
3. $f(x) = |x|$ "

Example 10.10. Let $C = \mathbb{R}$ and $f : C \to \mathbb{R}$... no son convexas

1. $f(x) = x^3$
2. ...

epi(f) NO es convexo

Property 10.11. Let $C \subseteq \mathbb{R}^n$ be a convex set and $f : C \to \mathbb{R}$. Then are equivalent:



1. epi(f) is convex;

$$\overbrace{(1-t)\cdot f(a) + t\cdot f(b)}^{\beta}$$

2. for every $a, b \in C$ and $t \in (0, 1)$: $\overbrace{f(a+t(b-a))}^{\alpha} \leq f(a)+t[f(b)-f(a)]$.

Proof. First, let us assume that epi(f) is convex. Let $a, b \in C$ and $t \in (0, 1)$; then, as $(a, f(a)), (b, f(b)) \in \text{epi}(f)$,

$$((1-t)a + tb, (1-t)f(a) + tf(b)) = (1-t)\,(a, f(a)) + t\,(b, f(b))$$

is in epi(f) and thus inequality of item 2 holds. Now let us assume that item 2 holds. Let $(a, \alpha), (b, \beta) \in \text{epi}(f)$ and $t \in (0, 1)$. Then, as

$$f\left((1-t)a + tb\right) \leq (1-t)f(a) + tf(b) \leq (1-t)\alpha + t\beta,$$

$$(1-t)\,(a, \alpha) + t\,(b, \beta) = ((1-t)a + tb, (1-t)\alpha + t\beta)$$

is in epi(f). $\square$

10.3.3 Preliminaries from differentiability

Definition 10.12 (differentiability)**.** Let $U \subseteq \mathbb{R}^m$ be a non-empty open set, let $f : U \to \mathbb{R}^n$ and let $a \in U$. We say that f is **differentiable at** a if there exists a linear transformation $f'(a) : \mathbb{R}^m \to \mathbb{R}^n$ such that

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - f'(a) \cdot h}{|h|} = 0,$$

we call $f'(a)$ the **derivative** of f at a. Moreover, we way that f is **differentiable** if it is differentiable at every $x \in U$. $\diamondsuit$

When we have a real-valued function we have some particular terminology.

Definition 10.13. Let $U \subseteq \mathbb{R}^m$ be a non-empty open set, let $f : U \to \mathbb{R}$ and let $a \in U$.

1. If f is differentiable at a, $f'(a)$ is also called the **gradient** of f at a and denoted $\nabla f(a)$.

   $M(m,1)$

2. If f is differentiable, we denote $\nabla f : U \to \mathbb{R}^m$ the function $x \mapsto \nabla f(x)$.

3. If f is differentiable and ∇f is differentiable at a, we say that f is **twice differentiable at** a.

4. If f is twice differentiable at a, the derivative of ∇f at a is called the **Hessian** of f at a and denoted $\nabla^2 f(a)$. $Hf(a)$

5. If f is twice differentiable at every $x \in U$, we say that f is **twice differentiable**.

6. If f is twice differentiable, we denote $\nabla^2 f : U \to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ the function $x \mapsto \nabla^2 f(x)$.

Remark 10.14. With the same notation of Definition 10.13:

1. If f is differentiable at a,

$$\nabla f(a) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_m}(a) \end{bmatrix}.$$

2. If f is twice differentiable at a,

Obs.: Si f es dos veces diferenciable, entonces $D^2 f(a)$ es simétrica.

$$\nabla^2 f(a) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1}(a) & \cdots & \frac{\partial^2 f}{\partial x_m \partial x_m}(a) \end{bmatrix}.$$

◇

In practice when we have to calculate the derivative of a function we do not usually use the Definition 10.12, most of the time we calculate it as a composition of other elementary ones, which are simpler and known. One of the most important tools for this is the following.

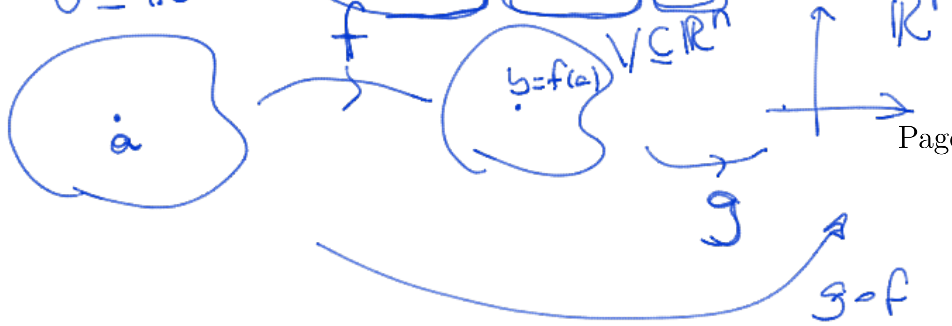
Fórmula de Itô

Property 10.15 (Chain rule)**.** With the same notation of Definition 10.12, let $V \subseteq \mathbb{R}^n$ be an open set such that $\boxed{f(U) \subseteq V}$ and let $g : V \to \mathbb{R}^p$. If f is differentiable at a and g is differentiable at $b := f(a)$, then $\boxed{g \circ f}$ is differentiable at a and

$$(g \circ f)'(a) = g'(f(a))\, f'(a).$$

◇

$U \subseteq \mathbb{R}^m$



$f'(a) : \mathbb{R}^m \to \mathbb{R}^n$

$g'(b) : \mathbb{R}^n \to \mathbb{R}^p$

$n \times m$

$p \times n$

$p \times m$

~~10~~11.4 Exercises

Exercise ~~10~~11.1. Show that the sets in Example 10.2 are convex.

Exercise ~~10~~11.2. Show that the sets in Example 10.3 are not convex.

Exercise ~~10~~11.3. Show that the first quadrant is a convex set.

Exercise ~~10~~11.4. Show that the polytopes are convex sets.

Exercise ~~10~~11.5. Show that the sets in Property 10.6 are convex.

$X := \{ x \in \mathbb{R}^n ; \ Ax \leq b, \ A \in M(m,n), \ b \in M(m,1) \}$