Basic concepts of Optimization

A. Brilli[1]

These notes are a summary of the material used for the Mathematical Programming course for the Bachelor program in Computer Engineering[2]

[1] Dip. Ingegneria Informatica Automatica e Gestionale, "Sapienza" Univ. di Roma

[2] Acknowledgement: the material was developed by revisiting Prof. Liuzzi's lecture notes.

The basics

Let $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$; consider the set

$$Y = \{t \in \mathbb{R} : t = f(x), \forall\ x \in X\} \subseteq \mathbb{R}$$

It makes sense to ask what the infimum (supremum) of Y is, and therefore to write

$$f^* = \inf Y \equiv \inf\{t \in \mathbb{R} : t = f(x), \forall\ x \in X\} \equiv \inf\{f(x) : x \in X\} \equiv \inf_{x \in X} f(x)$$

N.B. f^* is always (well) defined because $f^* \in \mathbb{R} \cup \{-\infty\}$

It may happen that $f^* = f(x^*)$ for some $x^* \in X$. When this is the case, we say that f^* is the minimum value of f on X and we write

$$f^* = \min_{x \in X} f(x)$$

First definitions

Definition (Optimization problem)

Given $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$, *the problem*

$$\begin{aligned} &\min\ f(x) \\ &\quad x \in S \subseteq X \end{aligned}$$

is called an **optimization problem**

- **unconstrained** *when* $S = \mathbb{R}^n$
- **constrained** *when* $S \subset \mathbb{R}^n$

Infeasible and unbounded problems

The optimization problem

$$\begin{aligned} \min\ & f(x) \\ & x \in S \subseteq X \end{aligned}$$

is called

- **infeasible** when $S = \emptyset$
- **unbounded below** when for every $M > 0$ there exists $x \in S$ such that $f(x) < -M$

Minimum points

Definition (Global minimum)

A point $x^* \in S$ is a **global minimum** *of f on S when*

$$f(x^*) \leq f(x) \qquad \forall\ x \in S$$

a **strict** *global minimum when* $f(x^*) < f(x) \qquad \forall\ x \in S,\ x \neq x^*$

Definition (Local minimum)

A point $x^* \in S$ is a **local minimum** *of f on S when there exists $\epsilon > 0$ such that*

$$f(x^*) \leq f(x) \qquad \forall\ x \in S \cap B(x^*, \epsilon)$$

a **strict** *local minimum when* $f(x^*) < f(x) \qquad \forall\ x \in S \cap B(x^*, \epsilon),\ x \neq x^*$

What can happen? (Examples)

$$\begin{array}{c} \min \; f(x) \\ x \in S \end{array}$$

1. Let $f(x) = x^2$ and $S = \{x \in \mathbb{R} : \; x > 1, x < -1\}$. (P) is infeasible
2. Let $f(x) = -x^2$ and $S = [0, +\infty)$. (P) is unbounded below
3. Let $f(x) = x^2$ and $S = [1, +\infty)$. (P) has a minimum with optimal solution $f^* = 1 = f(1)$, $x^* = 1 \in S$
4. Let $f(x) = x^2$ and $S = (1, +\infty)$. (P) has no optimal solution but $f^* = 1 = \inf\{x^2 : \; x \in S\}$

Summary

Given $f \colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ and $S \subseteq X$, and considering (assuming it is well-posed)

$$\begin{aligned} \min \ & f(x) \\ & x \in S \subseteq X \end{aligned} \tag{P}$$

we have the following cases

1. (P) is **infeasible** when $S = \emptyset$
2. (P) is **unbounded below** when for every $M > 0$ there exists $x \in S$ such that $f(x) < -M$
3. (P) **admits an optimal solution** when there exists $x^* \in S$ such that $f(x^*) \leq f(x)$ for every $x \in S$
4. (P) admits an infimum but **no optimal solution**, i.e. $\inf\{f(x) : \ x \in S\} = f^* > -\infty$ but there exists no $x^* \in S$ such that $f(x^*) = f^*$

Mathematical programming

Definition (Mathematical programming problem)

Given $f: X \subseteq \mathbb{R}^n \to \mathbb{R}$, the problem

$$\begin{aligned} \min \ & f(x) \\ & x \in S \subseteq X \end{aligned}$$

is called a **mathematical programming problem** *when S is described by a finite number of equations and inequalities, i.e.*

$$\begin{aligned} S = \{x \in \mathbb{R}^n : & g_i(x) \leq 0, i = 1, \ldots, m \\ & h_j(x) = 0, j = 1, \ldots, p\} \end{aligned}$$

Introduction

A **mathematical programming** problem

$$\begin{aligned} \min\ & f(x) \\ & g_i(x) \leq 0, \quad i = 1, \dots, m \\ & h_j(x) = 0, \quad j = 1, \dots, p \end{aligned}$$

such that:

1. $f(x) = c^\top x$, **LINEAR** objective function
2. equality and inequality constraints that are **AFFINE**

is called a **linear programming** problem, or LP problem

$$\begin{aligned} \min\ & c^\top x \\ & Ax \leq b \end{aligned}$$

Optimization problems

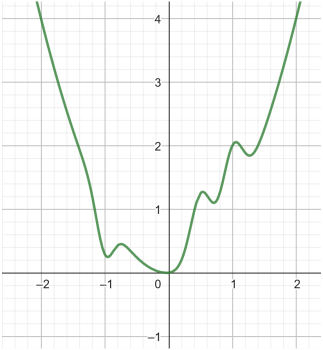
optimization problem		mathematical programming problem		linear programming problem
$\min f(x)$ $x \in S$ $f : \mathbb{R}^n \to \mathbb{R},\ S \subseteq \mathbb{R}^n$	$\rightsquigarrow$	$\min f(x)$ $h(x) = 0$ $g(x) \leq 0$ $h : \mathbb{R}^n \to \mathbb{R}^p,$ $g : \mathbb{R}^n \to \mathbb{R}^m$	$\rightsquigarrow$	$\min c^\top x$ $Ax = b$ $x \geq 0$ $c \in \mathbb{R}^n, b \in \mathbb{R}^m,$ $A \in \mathbb{R}^{m \times n}$

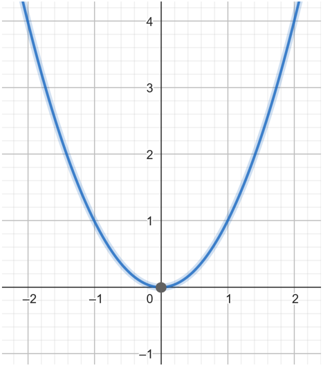
Fundamental theorem of LP: exactly one of the following holds

- infeasible problem
- unbounded below problem
- optimal solution exists at a vertex

Global and local minima

$\min\{f(x):\ x \in S\}$: we have seen that besides the global minimum, local minima may also exist. Example of a function with local minima and one without





Existence of the minimum

Given $X \subseteq \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$, how can we be sure that it makes sense to write $\min\{f(x), x \in X\}$, i.e., that at least one global minimum of f on X exists?

We distinguish two cases:

- $X \neq \mathbb{R}^n$
- $X = \mathbb{R}^n$

Constrained optimization ($X \neq \mathbb{R}^n$)

Theorem (Weierstrass)

Let $X \subset \mathbb{R}^n$ be non-empty and compact, and let f be continuous on X. Then f attains both its global minimum and maximum on X.

Unconstrained optimization ($X = \mathbb{R}^n$)

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $\alpha \in \mathbb{R}$. Define

$$L_\alpha = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$$

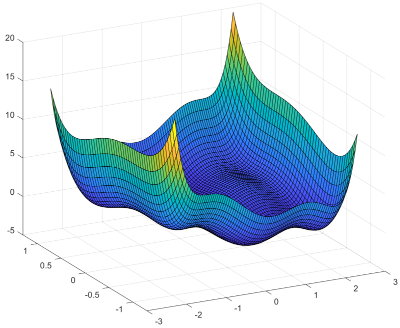
Definition (Level set)

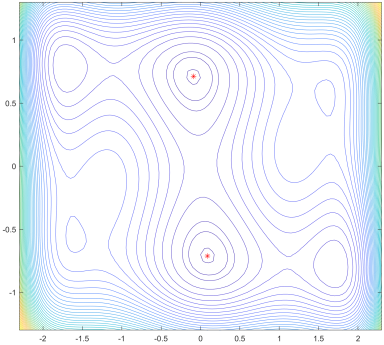
If L_α is non-empty for some $\alpha \in \mathbb{R}$, it is a lower level set of f

Theorem (Weierstrass)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous on $\mathbb{R}^n$. If there exists a non-empty and compact lower level set of f, then f attains a global minimum on $\mathbb{R}^n$.

Level sets





The tools we need

Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined on an open neighborhood of x

Definition (Descent direction)

A direction $d \in \mathbb{R}^n, d \neq 0$ is a descent direction for f at x when there exists $\bar{t} > 0$ such that

$$f(x + td) < f(x), \quad \forall t \in (0, \bar{t}]$$

1. If x^* is a global/local minimum of f on $\mathbb{R}^n$, can a descent direction for f at x^* exist?
2. If d is a descent direction for f at x, can x be a global/local minimum of f?

Descent direction

Let $f \colon \mathbb{R}^n \to \mathbb{R}$ and $\bar{x} \in \mathbb{R}^n$. If there exists a direction $d \neq 0$ such that moving from $\bar{x}$ along the half-line $\bar{x} + td$ yields points strictly better than $\bar{x}$, then $\bar{x}$ is not a minimum of f

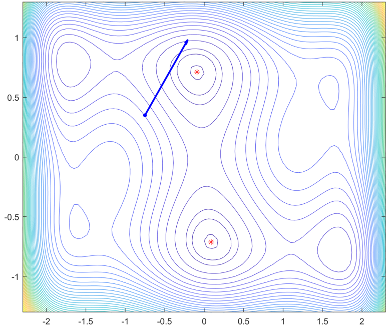
Formally

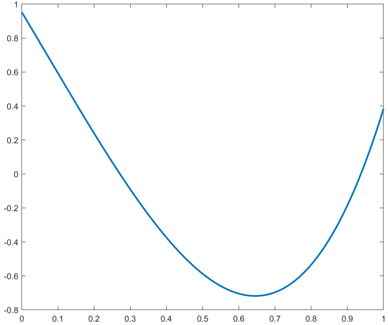
Definition

A vector $d \in \mathbb{R}^n$, $d \neq 0$, is a descent direction for f at $\bar{x}$ if there exists $\bar{\alpha} > 0$ such that

$$f(\bar{x} + \alpha d) < f(\bar{x}), \quad \forall\, \alpha \in (0, \bar{\alpha}]$$

Example


1
0.5
0
-0.5
-1
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2


1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

First-order descent condition

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable in a neighborhood of a point x, and let $d \in \mathbb{R}^n$, $d \neq 0$ be such that $\nabla f(x)^\top d < 0$. Then d is a descent direction for f at x.

It suffices to consider the definition of directional derivative of f at x along d

$$\lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t} = \nabla f(x)^\top d < 0$$

then for small t we must have $f(x + td) < f(x)$

First-order necessary condition

Proposition (First-order necessary condition)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable in a neighborhood of a point $x^* \in \mathbb{R}^n$. If x^* is a local minimum of f, then

$$\nabla f(x^*) = 0.$$

$$f(x, y) = x^4 + y^4 - 36xy$$

Does $(0, 0)$ satisfy the first-order necessary condition? That is, is it a **stationary point**?

Geometry in $\mathbb{R}^n$

Definition (Segment between two points)

Let $x, y \in \mathbb{R}^n$. The segment joining x and y is the set

$$[x, y] = \{z \in \mathbb{R}^n : \; z = ty + (1-t)x, t \in [0,1]\}$$
$$(x, y) = \{z \in \mathbb{R}^n : \; z = ty + (1-t)x, t \in (0,1)\}$$

But why are segments useful?

Definition (Convex set)

$C \subseteq \mathbb{R}^n$ is convex when, for any $x, y \in C$,

$$[x, y] \subseteq C$$

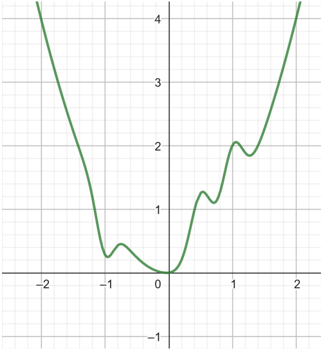
Properties

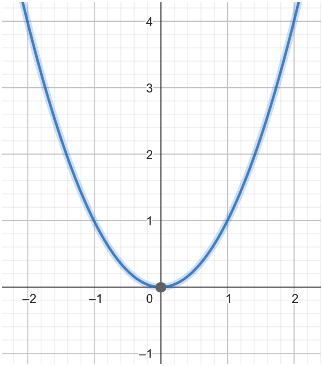
1. $\mathbb{R}^n$ is a convex set (obvious)
2. $\emptyset$ is convex (for lack of counterexample)
3. if $C_1, C_2 \subseteq \mathbb{R}^n$ are convex $\Rightarrow$ $C_1 \cap C_2$ is convex
4. if $C_1, C_2, \dots, C_p$ are p convex sets $\Rightarrow$

$$\bigcap_{i=1}^{p} C_i = C \text{ convex}$$

5. (in general) the intersection of a finite or infinite number of convex sets is convex
6. $\{x \in \mathbb{R}^n : \; a^\top x \leq b\}$ (halfspace) is convex
7. $\{x \in \mathbb{R}^n : \; a^\top x = b\}$ (hyperplane) is convex (because it is the intersection of two halfspaces)
8. a polyhedron $P = \{x \in \mathbb{R}^n : \; Ax \leq b\}$ is convex

Why are they so different?





- for many reasons!
- the **BLUE** curve always lies **below every secant segment**
- the **GREEN** curve does not

Convex functions

Intuitively: a function is "convex" when its graph always lies below every secant segment

Definition (Convex function)

A function $f : \mathbb{R}^n \to \mathbb{R}$ *defined on a convex set* $S \subseteq \mathbb{R}^n$ *is convex when, for every* $x, y \in S$

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y), \quad \forall\ \lambda \in [0,1]$$

Definition (Strictly convex function)

A function $f : \mathbb{R}^n \to \mathbb{R}$ *defined on a convex set* $S \subseteq \mathbb{R}^n$ *is strictly convex when, for every* $x, y \in S$, $x \neq y$

$$f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y), \quad \forall\ \lambda \in (0,1)$$

Linear functions

A **linear** or **affine** function

$$f(x) = c^\top x$$

is it convex?

Local and global minima

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex on a convex set $S \subseteq \mathbb{R}^n$. Then every local minimum of f on S is global.

Proof. Let x^* be a local minimum of f; then there exists $\epsilon > 0$ such that

$$f(x^*) \leq f(x), \quad \forall\ x \in B(x^*, \epsilon) \cap S$$

Let $x \in S \setminus B(x^*, \epsilon)$; consider the segment $[x^*, x]$. By convexity there must exist $\lambda \in (0, 1)$ such that $(1 - \lambda)x^* + \lambda x \in B(x^*, \epsilon) \cap S$, therefore

$$f(x^*) \leq f((1 - \lambda)x^* + \lambda x) \leq (1 - \lambda)f(x^*) + \lambda f(x)$$

that is

$$f(x^*) \leq f(x)$$

Set of global minima

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex on a convex set $S \subseteq \mathbb{R}^n$. The set of global minimum points is convex.

Proof. Let X^* be the set of global minima. The theorem is trivially true if $X^* = \emptyset$ or $X^* = \{x^*\}$. Suppose therefore that f admits two global minima $\bar{x} \in X^*$ and $\bar{y} \in X^*$. Then $f(\bar{x}) = f(\bar{y})$. Consider a point $z \in [\bar{x}, \bar{y}]$, i.e. $z = (1 - \lambda)\bar{x} + \lambda\bar{y}$, $\lambda \in [0, 1]$. By convexity of f we have

$$f((1 - \lambda)\bar{x} + \lambda\bar{y}) \leq (1 - \lambda)f(\bar{x}) + \lambda f(\bar{y}) = f(\bar{x})$$

Linear programming

Since $c^\top x$ is convex and $P = \{x : Ax = b, x \geq 0\}$ is a convex region, then

$$\begin{aligned} &\min \; c^\top x \\ &\quad Ax = b, \; x \geq 0 \end{aligned}$$

1. can it have minima that are not global?
2. What does the set of global minima of the problem look like?

First-order convexity condition

$f : \mathbb{R}^n \to \mathbb{R}$ of class 1 on a convex open set $S \subseteq \mathbb{R}^n$.

Theorem (Necessary and sufficient condition for convexity)

f is convex on S **if and only if** *for every* $y \in S$

$$f(y) + \nabla f(y)^\top (x - y) \leq f(x) \quad \forall\, x \in S$$

Theorem (Necessary and sufficient condition for convexity)

f is strictly convex on S **if and only if** *for every* $y \in S$

$$f(y) + \nabla f(y)^\top (x - y) < f(x) \quad \forall\, x \in S$$

Convexity of quadratic functions

Let Q be symmetric and

$$f(x) = \frac{1}{2}x^\top Q x + c^\top x$$

then

1. f is **convex** on $\mathbb{R}^n$ **if and only if** Q is **positive semidefinite**
2. f is **strictly convex** on $\mathbb{R}^n$ **if and only if** Q is **positive definite**

Global minimum condition

Proposition (Necessary and sufficient condition for global minimum)

Let $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(\mathbb{R}^n)$ and convex. $x^* \in \mathbb{R}^n$ is a global minimum of f if and only if $\nabla f(x^*) = 0$.

Proof.

$\boxed{\Rightarrow}$ Necessity follows from the fact that if x^* is a global minimum, it is also a local minimum of f.

$\boxed{\Leftarrow}$ If f is convex on $\mathbb{R}^n$, for every $x \in \mathbb{R}^n$,

$$f(x) \geq f(x^*) + \nabla f(x^*)^\top (x - x^*)$$

Therefore the result follows from the hypothesis $\nabla f(x^*) = 0$. $\square$