Optimization and Operations Research

Introduction

Michel Bierlaire

Introduction to optimization and operations research

Modeling

Motivation

- Translating a concrete problem into a mathematical model is difficult.
- It is more of an art, but it requires rigor and systematism.
- We introduce the process and illustrates it on a small example.

Motivation

Mathematical model

- Mathematical representation
- of a phenomenon,
- done in order to better study it.

Model for optimization

Decision variables

$$x \in \mathbb{R}^n$$

Objective function

$$f(x) \in \mathbb{R}$$

Constraints

$$x \in X \subseteq \mathbb{R}^n$$

Example: Château Laupt-Himum

Problem description

Rosé and red wine using local grapes.

- Grape: max. 1 ton of Pinot. Price: 3 €/kilo.
- Vinification as rosé: cost 2 € per kilo of grape.
- Vinification as red (Pinot Noir) : cost 3.50 € per kilo of grape.
- For one liter of wine, one kilo of grapes is needed.

Example: Château Laupt-Himum

Problem description

Marketing strategy: rebate proportional to the production

- Price rosé: 15 €/ℓ. Rebate: 2 € per 100ℓ.
- Example: if 100 ℓ of rosé are produced, the price is 13 €/ℓ
- Price red: 23 €/ℓ. Rebate: 1 € per 100ℓ.
- Example: if 100 ℓ of red are produced, the price is 22 €/ℓ

Château Laupt-Himum: decision variables

Liters of rosé to produce per year

$$x_1$$

Liters of red to produce per year

$$x_2$$

Kilos of grapes to buy

$$x_3$$

Château Laupt-Himum: objective function

Revenue per liter of rosé

$$15 - \frac{2}{100}x_1$$

Revenue per liter of red

$$23 - \frac{1}{100}x_2$$

Total revenues

$$x_1\left(15 - \frac{2}{100}x_1\right) + x_2\left(23 - \frac{1}{100}x_2\right)$$

Château Laupt-Himum: objective function

Cost for producing rosé

$$2x_1$$

Cost for producing red

$$3.5x_2$$

Cost for buying grapes

$$3x_3$$

Château Laupt-Himum: objective function

Total benefits

$$x_1\left(15-\frac{2}{100}x_1\right)+x_2\left(23-\frac{1}{100}x_2\right)-(2x_1+3.5x_2+3x_3)$$

Château Laupt-Himum: constraints

Maximum quantity of grapes

$$x_3 \leq 1000$$

1 kilo of grapes produces 1 liter of wine

$$x_1 + x_2 \leq x_3$$

Non negativity

$$x_1 \geq 0, \; x_2 \geq 0, \; x_3 \geq 0$$

Mathematical model

$$\max_{x\in\mathbb{R}^3} f(x) = x_1\left(15 - \frac{2}{100}x_1\right) + x_2\left(23 - \frac{1}{100}x_2\right) - (2x_1 + 3.5x_2 + 3x_3)$$

subject to

$$\begin{aligned} x_1 + x_2 &\leq x_3 \\ x_3 &\leq 1000 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \\ x_3 &\geq 0 \end{aligned}$$

Notations

$$y = \min_{x \in \mathbb{R}^n} f(x) \qquad y \in \mathbb{R}$$

$$x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} f(x) \qquad x^* \in \mathbb{R}^n$$

$$y = f(x^*)$$

$$f(x) = (x_1 - 1)^2 + (x_2 + 2)^2$$

$$x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} f(x) = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

$$y = \min_{x \in \mathbb{R}^n} f(x) = 0.$$

Problem transformations

Motivation

- There are many ways to write an optimization problem.
- Algorithms usually require a specific type of formulation.
- For instance, most optimization software are designed only for minimization, or only for maximization.

Equivalence

Definition

Problems P_1 and P_2 are equivalent if a feasible point of P_1 can be created from a feasible point of P_2, with the same value of the objective function.

$$P_1 : \min 2 - x \text{ s. t. } x \geq 0 \text{ and } x \leq 1$$

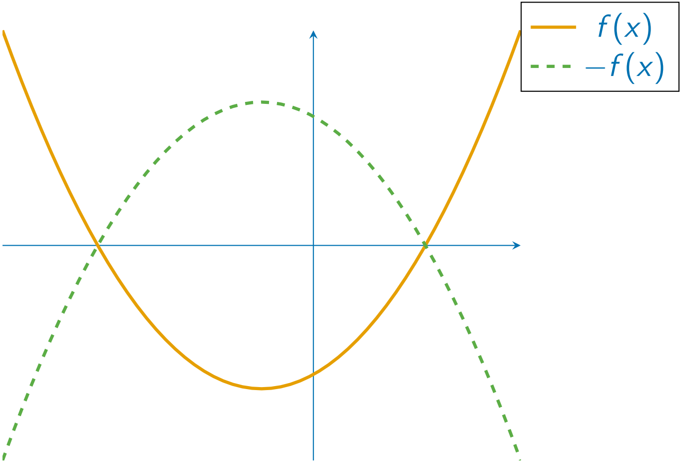
$$P_2 : -\max y \text{ s. t. } y \geq -2 \text{ and } y \leq -1$$

$$x \rightarrow y = x - 2$$

$$x = 0.5, y = -1.5$$

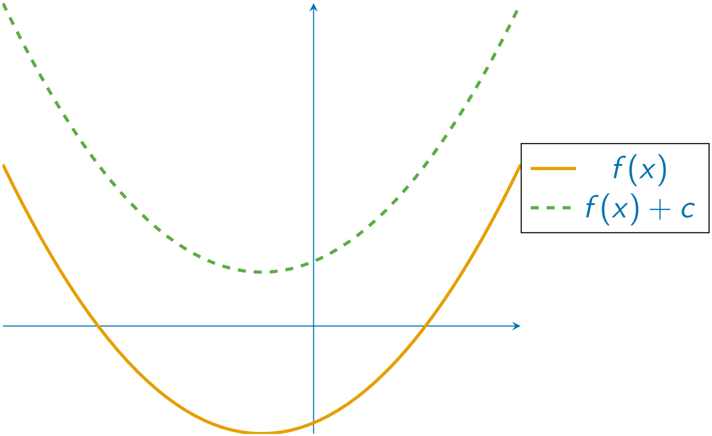
$$\text{Obj. } P_1 = \text{Obj. } P_2 = 1.5$$

Minimization or maximization



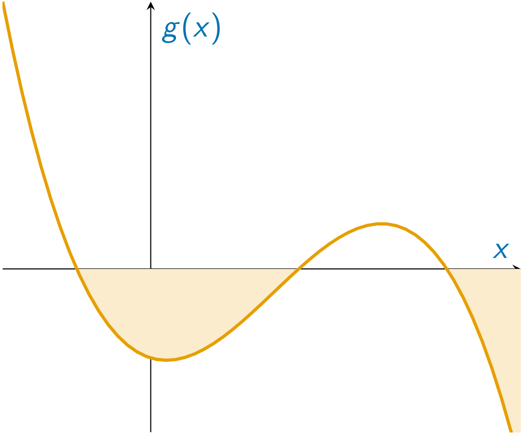
$$\min f(x) = -\max -f(x), \operatorname{argmin} f(x) = \operatorname{argmax} -f(x)$$

Constant term



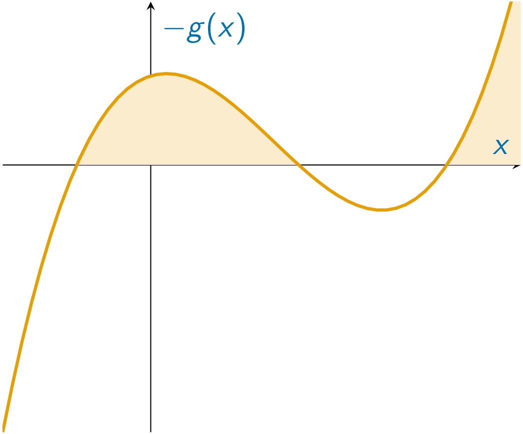
Inequality constraints

$$g(x) \leq 0$$



Inequality constraints

$$-g(x) \geq 0$$



Equality and inequality constraints

$$g(x) = 0 \Longleftrightarrow \begin{cases} g(x) \leq 0 \\ g(x) \geq 0 \,. \end{cases}$$

Signed variables

Some software impose only non negative variables
$x \in \mathbb{R}$, $x = x^+ - x^-$ with $x^+ \geq 0$, $x^- \geq 0$

Translation

Constraint $x \geq a$. Change of variable: $x = \tilde{x} + a$. Constraint become $\tilde{x} \geq 0$

Slack variables

Linear

$$g(x) \leq 0 \iff \begin{cases} g(x) + y = 0 \\ y \geq 0\,. \end{cases}$$

Non linear

$$g(x) \leq 0 \iff g(x) + z^2 = 0.$$

Problem definition

Motivation

- We provide a generic definition of an optimization problem.
- We discuss the types of solutions that can be found.

Problem definition

$$\min_{x \in \mathbb{R}^n} f(x) \quad f : \mathbb{R}^n \to \mathbb{R}, n > 0,$$

subject to

$$\begin{aligned} h(x) &= 0, \quad h : \mathbb{R}^n \to \mathbb{R}^m, m \geq 0, \\ g(x) &\leq 0, \quad g : \mathbb{R}^n \to \mathbb{R}^p, p \geq 0, \end{aligned}$$

and

$$x \in X, \quad X \text{convex.}$$

Feasible set

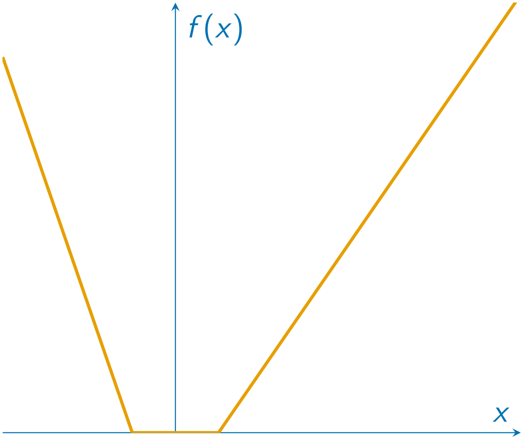
$$Y = \left\{x \in \mathbb{R}^n \mid h(x) = 0\,,\ g(x) \leq 0 \text{ and } x \in X\right\}$$

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

subject to

$$x \in Y$$

Global optimum

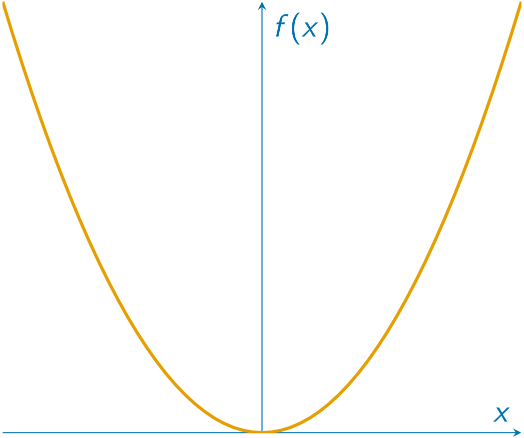


Global minimum

x^* is a global minimum if

$$f(x^*) \leq f(x)\,, \qquad \forall x \in Y\,.$$

Strict global optimum

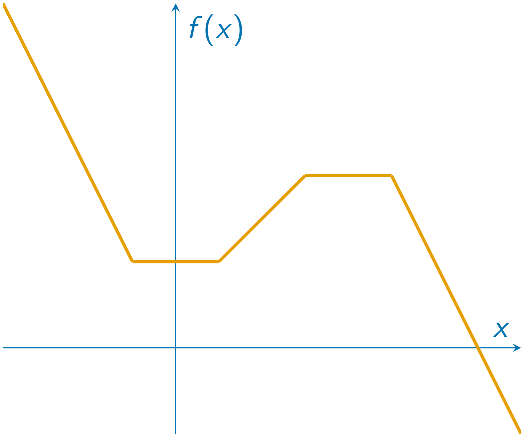


Strict global minimum

x^* is a strict global minimum if

$$f(x^*) < f(x)\,, \qquad \forall x \in Y,\, x \neq x^*.$$

Local optimum

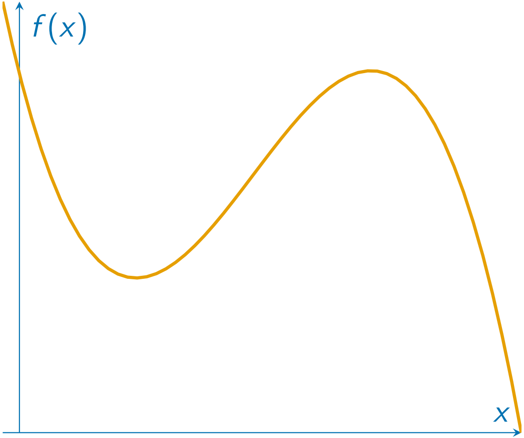


Local minimum

x^* is a local minimum if there exists $\varepsilon > 0$ such that

$$f(x^*) \leq f(x), \quad \forall x \in Y \text{ such that } \left\| x - x^* \right\| < \varepsilon .$$

Strict local optimum



Strict local minimum

x^* is a strict local minimum if there exists $\varepsilon > 0$ such that

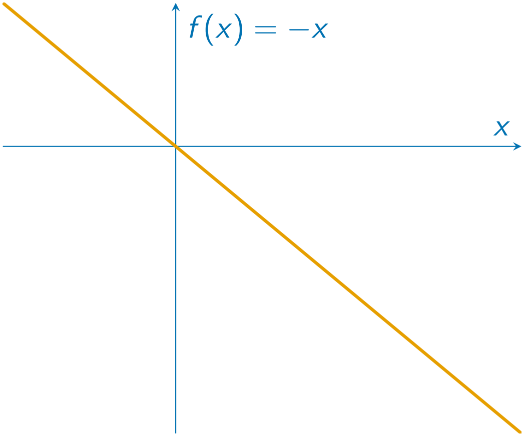
$$f(x^*) < f(x)\,, \quad \forall x \in Y \text{ such that } \left\|x - x^*\right\| < \varepsilon, x \neq x^*.$$

Properties

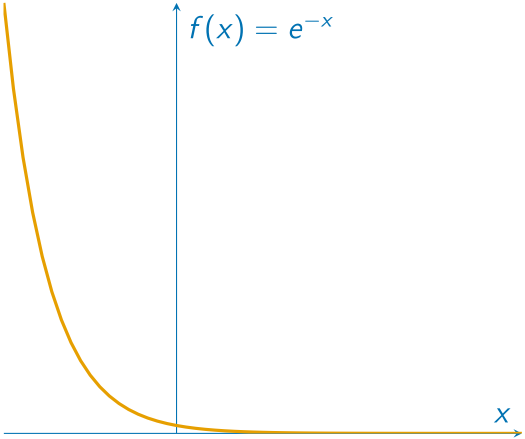
Motivation

- We introduce some important properties of optimization problems.
- We consider from now on only minimization.
- The lecture provides the intuitions. We refer to Section 1.4 of the book for a more rigorous discussion.

Function bounded from below



Infimum vs optimum



Bounded from below, but no optimum

- Give me any x.
- $x + 1$ is always better.
- Indeed, $\exp(-(x+1)) < \exp(-x)$.

Infimum vs optimum

Infimum

Largest lower bound

$$\inf_{y \in Y} f(y) \leq f(x), \ \forall x \in Y$$

For each $M > \inf_{y \in Y} f(y)$,

$$\exists x \in Y \text{ such that } f(x) < M.$$

Always exists is f is bounded.

Optimum

$$x^* \text{ such that } f(x^*) = \inf_{y \in Y} f(y)$$

Does not always exist. See Weierstrass theorem 1.14 for sufficient conditions.

Optimality conditions

Motivation

- Characterization of optimal solutions.
- Central role in optimization.
- Key ingredients for the algorithms.
- Unconstrained optimization: Fermat's theorem.
- Constrained optimization: Karush-Kuhn-Tucker conditions.

Unconstrained optimization

Problem definition

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

where f is twice differentiable.

Fermat's theorem

- x^* is a local minimum of $f : \mathbb{R}^n \to \mathbb{R}$.
- If f is differentiable around x^*, then
$$\nabla f(x^*) = 0.$$
- If f is twice differentiable around x^*, then
$$\nabla^2 f(x^*) \geq 0 \quad \text{[positive semidefinite]}.$$

Theorem 5.1

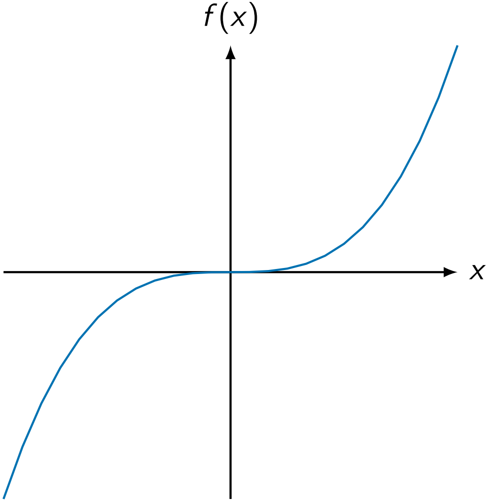
Necessary, but not sufficient

Example

$$f(x) = x^3$$

$$f'(x) = 3x^2, f''(x) = 6x$$

$$f'(0) = 0, f''(0) \geq 0$$



Sufficient optimality conditions

Theorem 5.7

- Consider x^* and $f : \mathbb{R}^n \to \mathbb{R}$ twice differentiable.
- If

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

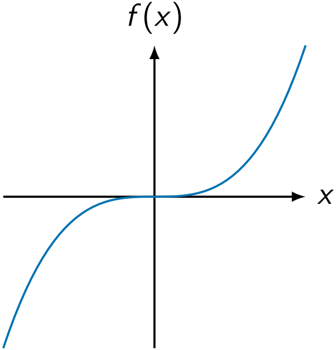
- and

$$\nabla^2 f(x^*) > 0 \quad \text{[positive definite]}$$

- then x^* is a local minimum of f.

Examples

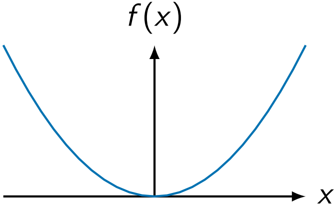
$$f(x) = x^3, x^* = 0$$



$$f'(x) = 3x^2, f''(x) = 6x$$

$$f'(0) = 0, f''(0) \not> 0$$

$$f(x) = x^2, x^* = 0$$

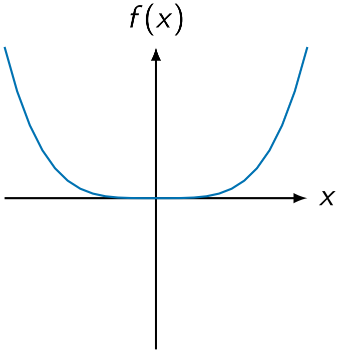


$$f'(x) = 2x, f''(x) = 2$$

$$f'(0) = 0, f''(0) > 0$$

Examples

$$f(x) = x^4, x^* = 0$$



$$f'(x) = 4x^3, f''(x) = 12x^2$$

$$f'(0) = 0, f''(0) \not> 0$$

Convexity

Set

A set X is convex if, $\forall x, y \in X$, $\forall 0 \leq \alpha \leq 1$,

$$\alpha x + (1 - \alpha) y \in X.$$

Function

A function f is convex $\forall x, y$, $\forall 0 \leq \alpha \leq 1$,

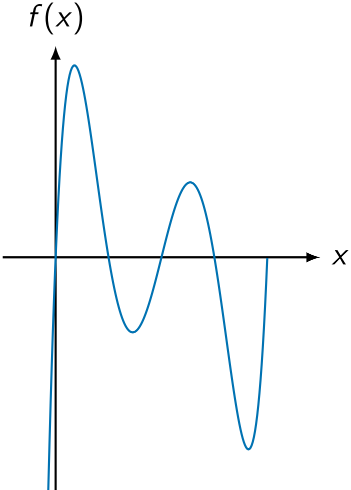
$$f(\alpha x + (1 - \alpha) y) \leq \alpha f(x) + (1 - \alpha) f(y).$$

Convexity and second derivatives

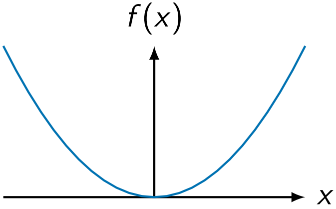
- Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $\nabla^2 f(x^*)$
- $\nabla^2 f(x^*) \geq 0 \iff \lambda_i \geq 0$, $i = 1, \ldots, n$
- Eigenvalue $=$ curvature along the eigenvector.

Global and local optimum

Sufficient conditions are verified both for local and global minima.



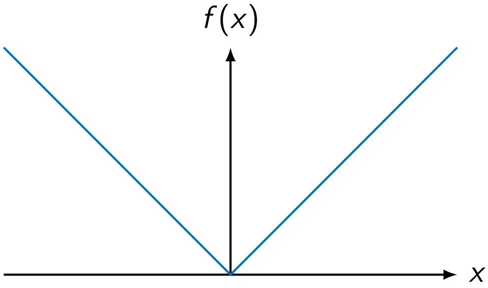
For convex functions, local optimum = global optimum.



Global optimum

Theorem 5.9

- $f : \mathbb{R}^n \to \mathbb{R}$ continuous.
- $x^* \in \mathbb{R}^n$ a local minimum of f.
- If f is convex, then x^* is a global minimum of f.
- If f is strictly convex, x^* is the unique global minimum of f.



Constrained optimization: the convex case

$$\min_{x \in \mathbb{R}^n} f(x) \quad f : \mathbb{R}^n \to \mathbb{R}, n > 0, \text{convex},$$

subject to

$$x \in X, \quad X \text{convex}.$$

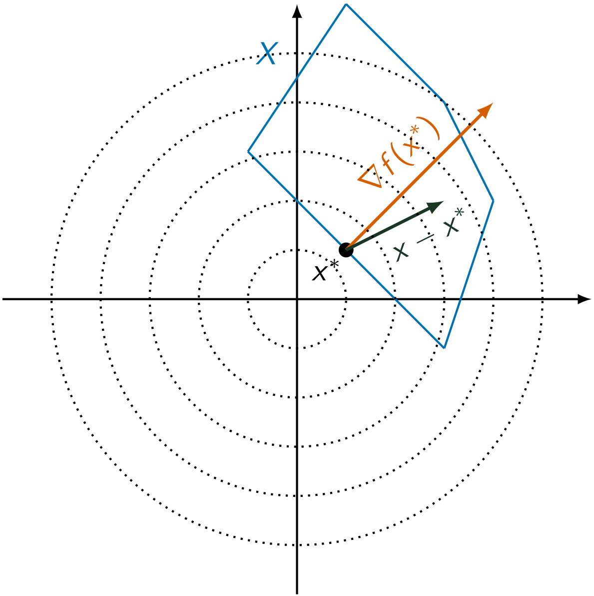
Necessary and sufficient condition

$$x^* \text{ global optimum}$$

$$\Longleftrightarrow$$

$$\forall x \in X, \nabla f(x^*)^T (x - x^*) \geq 0.$$

Necessary and sufficient condition



Constrained optimization: the general case

$$\min_{x\in\mathbb{R}^n} f(x) \quad f:\mathbb{R}^n \to \mathbb{R}, n>0,$$

subject to

$$\begin{aligned} h(x) &= 0, \quad h:\mathbb{R}^n \to \mathbb{R}^m, m \geq 0, \\ g(x) &\leq 0, \quad g:\mathbb{R}^n \to \mathbb{R}^p, p \geq 0. \end{aligned}$$

Definition: Lagrangian

$$L(x,\lambda,\mu) = f(x) + \lambda^T h(x) + \mu^T g(x).$$

$$\lambda \in \mathbb{R}^m, \quad \mu \in \mathbb{R}^p.$$

Karush-Kuhn-Tucker: first-order necessary conditions

Theorem 6.13

- $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^p$, $h : \mathbb{R}^n \to \mathbb{R}^m$ continuously differentiable.
- x^* local optimal of the problem.
- If the constraints are “qualified” at x^*, there exists a unique $\lambda^* \in \mathbb{R}^m$, a unique $\mu^* \in \mathbb{R}^p$, $\mu^* \geq 0$, such that

  $$\nabla_x L(x^*, \lambda^*, \mu^*) = 0,$$

  and

  $$\mu_i^* g_i(x^*) = 0, \ i = 1, \ldots, p.$$

Karush-Kuhn-Tucker: second-order necessary conditions

Theorem 6.13

- $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^p$, $h : \mathbb{R}^n \to \mathbb{R}^m$ twice differentiable.
- x^* local optimal of the problem.
- If the constraints are "qualified" at x^*, there exists a unique $\lambda^* \in \mathbb{R}^m$, a unique $\mu^* \in \mathbb{R}^p$, $\mu^* \geq 0$, such that

$$y^T \nabla^2_{xx} L(x, \lambda^*, \mu^*) y \geq 0,$$

for each $y \in \mathbb{R}^n$ such that

$$\begin{aligned} y^T \nabla h_i(x^*) &= 0, \; i = 1, \ldots, m, \\ y^T \nabla g_j(x^*) &= 0, \; j = 1, \ldots, p, \; \text{if } g_i(x^*) = 0. \end{aligned}$$

Notes

- Linear constraints are always “qualified”.
- Lagrangian, and Lagrange multipliers λ^* and μ^* are related to duality.
- The fact that $g_j(x^*) = 0$ plays a role. We say that the constraint is “active”.

Active constraints

Motivation

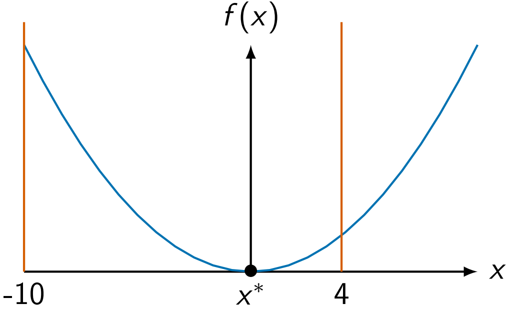
- Active constraints = constraints that matter.
- In linear optimization, finding the optimum solution amounts to finding the constraints that are active at the solution.

First example

$$\min_{x\in\mathbb{R}} x^2$$

subject to

$$-10 \le x \le 4.$$

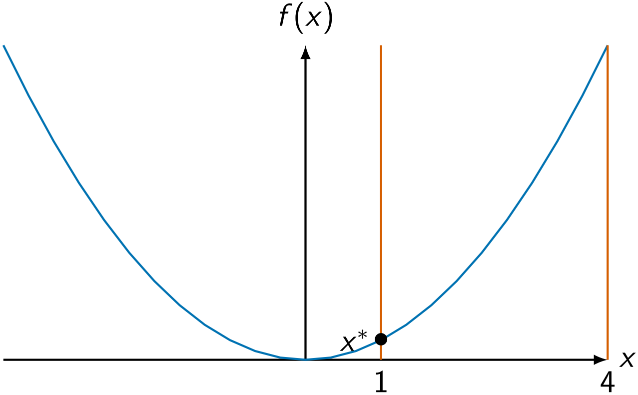


Second example

$$\min_{x \in \mathbb{R}} x^2$$

subject to

$$1 \leq x \leq 4.$$



Definition

Inequality constraints

Consider $g : \mathbb{R}^n \to \mathbb{R}$ and the constraint

$$g(x) \leq 0.$$

It is *active* at x^* if

$$g(x^*) = 0.$$

Equality constraints

Consider $h : \mathbb{R}^n \to \mathbb{R}$ and the constraint

$$h(x) = 0.$$

It is *active* at x^* if

$$h(x^*) = 0.$$

Optimization

- Consider the problem

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

  subject to

  $$g(x) \leq 0,$$

  where $g : \mathbb{R}^n \to \mathbb{R}^p$.
- x^* is a local optimum and

  $$\mathcal{A}(x^*) = \{i = 1, \ldots, p \,|\, g_i(x^*) = 0\}.$$

- x^* is also a local optimum of

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

  subject to

  $$g_i(x) = 0, \forall i \in \mathcal{A}(x^*).$$

Summary

- Modeling.
- Problem transformations.
- Definitions of optimum: global/local, strict.
- In the following, we work only with minimization.
- We assume that there is at least one optimum.
- It means that the objective function must be bounded from below on the feasible set.
- But it is not sufficient. See Weierstrass theorem for sufficient conditions.
- Optimality conditions: Fermat, Karush-Kuhn-Tucker.
- Constraints that do not play a role at the solution can be ignored.