Introduction to Optimization

Vincent Leclère (CERMICS, ENPC)[1]

Autumn 2026

[1] Based in part on slides by F. Meunier.

Why should I bother to learn this stuff?

- Engineering is full of decisions: how much to produce, how to design a structure, how to allocate limited resources.
- Optimization provides a systematic way to turn these questions into a model that can be solved numerically.
- The same framework applies to very different fields: mechanics, energy, logistics, economics, machine learning, ...
- Constraints are often what make an engineering problem interesting: capacities, balances, geometry, physical limits, ...

Contents

1 Introduction
- What is optimization?
- Vocabulary
- Practice and tools in optimization

2 First theoretical elements
- Existence of an optimal solution
- Unconstrained optimization

3 Optimization under equality constraints
- Lagrange multipliers
- Tangent cone and constraint qualification

4 Wrap-up

What is optimization?

To optimize $=$ to choose the best decision among feasible alternatives.

This is a natural objective in engineering. Three representative examples:

1. **Production planning.** How much of each product should we manufacture to maximize profit under limited resources?
2. **Traveling salesman.** In which order should we visit a set of cities to minimize total travel cost?
3. **Spring equilibrium.** Which displacement minimizes mechanical energy while satisfying boundary conditions?

These examples lead respectively to linear, discrete, and differentiable constrained optimization models.

A mathematical optimization problem

$$\begin{aligned} \mathrm{Min}_x \quad & f(x) \\ \text{s.t.} \quad & x \in X \end{aligned} \tag{P}$$

An optimization problem is specified by three objects:

- x: the decision variable — what we can choose;
- X: the feasible set — the constraints;
- f: the objective function (or criterion) — how we evaluate a decision.

We mostly use minimization; maximizing f is equivalent to minimizing $-f$.

Course roadmap

The course is organized around four main pillars:

1. **Constrained optimization**: Lagrange multipliers and KKT conditions;
2. **Convex optimization**: global optimality, duality, and first-order algorithms;
3. **Linear optimization**: polyhedra, simplex, and linear-programming duality;
4. **Mixed-integer linear optimization**: discrete modeling and branch-and-bound.

The objective is not only to solve small problems by hand, but to understand how to formulate, analyze, and solve optimization models with modern numerical tools.

Course roadmap

The course is organized around four main pillars:

1. **Constrained optimization**: Lagrange multipliers and KKT conditions;
2. **Convex optimization**: global optimality, duality, and first-order algorithms;
3. **Linear optimization**: polyhedra, simplex, and linear-programming duality;
4. **Mixed-integer linear optimization**: discrete modeling and branch-and-bound.

The objective is not only to solve small problems by hand, but to understand how to formulate, analyze, and solve optimization models with modern numerical tools.

Contents

1 Introduction
- What is optimization?
- Vocabulary
- Practice and tools in optimization

2 First theoretical elements
- Existence of an optimal solution
- Unconstrained optimization

3 Optimization under equality constraints
- Lagrange multipliers
- Tangent cone and constraint qualification

4 Wrap-up

Local and global minima

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $X \subseteq \mathbb{R}^n$.

$x^\sharp$ is a local minimum of f on X if there exists a neighborhood V of $x^\sharp$ such that

$$f(x^\sharp) \leq f(x) \qquad \text{for all } x \in V \cap X.$$

$x^\sharp$ is a global minimum of f on X if

$$f(x^\sharp) \leq f(x) \qquad \text{for all } x \in X.$$

$$\text{global minimum} \quad \Longrightarrow \quad \text{local minimum.}$$

Domain and sublevel sets

In optimization it is convenient to allow an objective to take the value $+\infty$. We denote

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}.$$

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$.

- The domain of f is

$$\operatorname{dom} f := \{x \in \mathbb{R}^n \mid f(x) < +\infty\}.$$

- The sublevel set at level $M \in \mathbb{R}$ is

$$\operatorname{lev}_M(f) := \{x \in \mathbb{R}^n \mid f(x) \leq M\}.$$

Assigning $+\infty$ outside a set is a useful way to encode constraints inside an objective.

Domain and sublevel sets

In optimization it is convenient to allow an objective to take the value $+\infty$. We denote

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}.$$

Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$.

- The domain of f is
$$\operatorname{dom} f := \{x \in \mathbb{R}^n \mid f(x) < +\infty\}.$$
- The sublevel set at level $M \in \mathbb{R}$ is
$$\operatorname{lev}_M(f) := \{x \in \mathbb{R}^n \mid f(x) \leq M\}.$$

Assigning $+\infty$ outside a set is a useful way to encode constraints inside an objective.

Optimal solution and optimal value

$$\boxed{\begin{array}{ll} \mathrm{Min}_{x} & f(x) \\ \text{s.t.} & x \in X \end{array}}$$

$x^\sharp$ is an optimal solution if it is a global minimum of f on X.

The optimal value is

$$p^\star := \inf\{f(x) \mid x \in X\}.$$

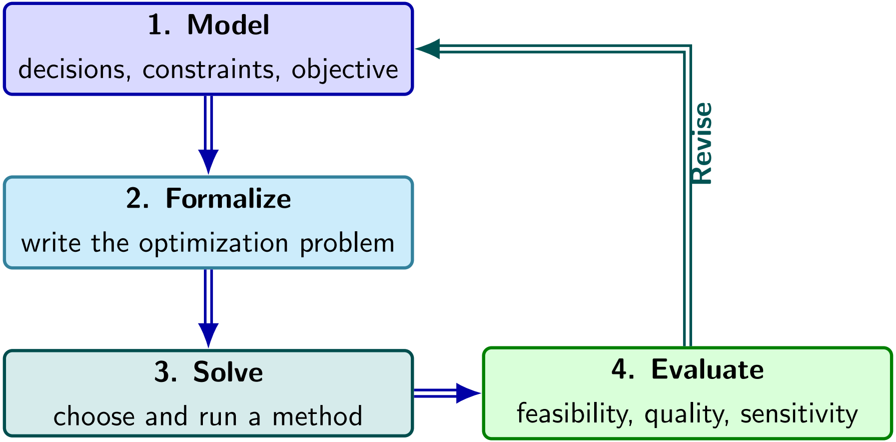
The optimal value need not be attained by an optimal solution.

By convention, $p^\star = +\infty$ when $X = \varnothing$.

Contents

1. Introduction
   - What is optimization?
   - Vocabulary
   - Practice and tools in optimization
2. First theoretical elements
   - Existence of an optimal solution
   - Unconstrained optimization
3. Optimization under equality constraints
   - Lagrange multipliers
   - Tangent cone and constraint qualification
4. Wrap-up

Optimization in practice: an iterative loop



An unsatisfactory solution may indicate a numerical issue or a modeling issue.

Optimization tools

- **Modeler** (e.g. AMPL, GAMS, Pyomo, JuMP, CVXPY): expresses variables, constraints, objectives, and indexed models in a readable form.

➥ First write the mathematical optimization problem clearly, then translate it into code.

- **Solver** (e.g. CPLEX, Gurobi, MOSEK, IPOPT): implements numerical algorithms for one or several classes of optimization problems.

➥ In general, use established solvers rather than reimplementing numerical algorithms from scratch.

- **Post-processing**: check residuals and status, interpret the solution, perform sensitivity analyses, and visualize results.
- **Programming language** (e.g. Julia, Python, C++): handles data and orchestrates the full workflow.

Optimization tools

- **Modeler** (e.g. AMPL, GAMS, Pyomo, JuMP, CVXPY): expresses variables, constraints, objectives, and indexed models in a readable form.

➥ First write the mathematical optimization problem clearly, then translate it into code.

- **Solver** (e.g. CPLEX, Gurobi, MOSEK, IPOPT): implements numerical algorithms for one or several classes of optimization problems.

➥ In general, use established solvers rather than reimplementing numerical algorithms from scratch.

- **Post-processing**: check residuals and status, interpret the solution, perform sensitivity analyses, and visualize results.
- **Programming language** (e.g. Julia, Python, C++): handles data and orchestrates the full workflow.

Optimization tools

- **Modeler** (e.g. AMPL, GAMS, Pyomo, JuMP, CVXPY): expresses variables, constraints, objectives, and indexed models in a readable form.

➥ First write the mathematical optimization problem clearly, then translate it into code.

- **Solver** (e.g. CPLEX, Gurobi, MOSEK, IPOPT): implements numerical algorithms for one or several classes of optimization problems.

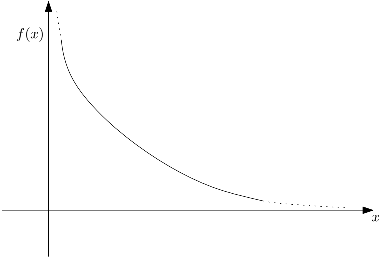
➥ In general, use established solvers rather than reimplementing numerical algorithms from scratch.

- **Post-processing**: check residuals and status, interpret the solution, perform sensitivity analyses, and visualize results.
- **Programming language** (e.g. Julia, Python, C++): handles data and orchestrates the full workflow.

Contents

1. Introduction
   - What is optimization?
   - Vocabulary
   - Practice and tools in optimization
2. First theoretical elements
   - Existence of an optimal solution
   - Unconstrained optimization
3. Optimization under equality constraints
   - Lagrange multipliers
   - Tangent cone and constraint qualification
4. Wrap-up

Optimal value versus optimal solution



A problem can have a finite optimal value but no optimal solution.

For example,

$$\begin{aligned} \mathrm{Min}_x \quad & \frac{1}{x} \\ \text{s.t.} \quad & x > 0. \end{aligned}$$

Its optimal value is 0, but no feasible x attains it.

Some topology recalls

- $C \subseteq \mathbb{R}^n$ is bounded if there exists $M > 0$ such that $\|x\| \leq M$ for all $x \in C$.
- $C \subseteq \mathbb{R}^n$ is closed if every convergent sequence in C has its limit in C.
- $C \subseteq \mathbb{R}^n$ is compact if it is closed and bounded.
- $f : \mathbb{R}^n \to \mathbb{R}$ is continuous if $x_k \to x$ implies $f(x_k) \to f(x)$.
- $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is coercive if

$$f(x) \to +\infty \qquad \text{when } \|x\| \to +\infty.$$

Theorem (Extreme value theorem)

A continuous real-valued function on a nonempty compact set attains its minimum and its maximum.

Some topology recalls

- $C \subseteq \mathbb{R}^n$ is bounded if there exists $M > 0$ such that $\|x\| \leq M$ for all $x \in C$.
- $C \subseteq \mathbb{R}^n$ is closed if every convergent sequence in C has its limit in C.
- $C \subseteq \mathbb{R}^n$ is compact if it is closed and bounded.
- $f : \mathbb{R}^n \to \mathbb{R}$ is continuous if $x_k \to x$ implies $f(x_k) \to f(x)$.
- $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is coercive if

$$f(x) \to +\infty \qquad \text{when } \|x\| \to +\infty.$$

Theorem (Extreme value theorem)

A continuous real-valued function on a nonempty compact set attains its minimum and its maximum.

Some existence results

Consider

$$\begin{aligned} \mathrm{Min}_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & x \in C. \end{aligned}$$

Theorem

If C is nonempty and compact and f is continuous, then the problem admits an optimal solution.

Theorem

If C is nonempty and closed and f is continuous and coercive, then the problem admits an optimal solution.

These are sufficient existence conditions, not necessary ones.

Some existence results

Consider

$$\begin{aligned} \mathrm{Min}_{x\in\mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & x \in C. \end{aligned}$$

Theorem

If C is nonempty and compact and f is continuous, then the problem admits an optimal solution.

Theorem

If C is nonempty and closed and f is continuous and coercive, then the problem admits an optimal solution.

These are sufficient existence conditions, not necessary ones.

Some existence results

♡

Consider

$$\begin{aligned} \mathrm{Min}_{x\in\mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & x \in C. \end{aligned}$$

Theorem

If C is nonempty and compact and f is continuous, then the problem admits an optimal solution.

Theorem

If C is nonempty and closed and f is continuous and coercive, then the problem admits an optimal solution.

These are sufficient existence conditions, not necessary ones.

Contents

1 Introduction
- What is optimization?
- Vocabulary
- Practice and tools in optimization

2 First theoretical elements
- Existence of an optimal solution
- Unconstrained optimization

3 Optimization under equality constraints
- Lagrange multipliers
- Tangent cone and constraint qualification

4 Wrap-up

Unconstrained optimization

♡

The first-order necessary condition is also known as Fermat's rule.

Theorem (FONC)

Let $x^\sharp$ be a local minimum of $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. If f is differentiable at $x^\sharp$, then

$$\nabla f(x^\sharp) = \mathbf{0}.$$

A second-order sufficient condition is:

Theorem (SOSC)

Assume $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is twice differentiable on its domain. If $x^\sharp \in \operatorname{dom} f$, $\nabla f(x^\sharp) = \mathbf{0}$, and $\nabla^2 f(x^\sharp)$ is positive definite, then $x^\sharp$ is a local minimum.

Unconstrained optimization

The first-order necessary condition is also known as Fermat's rule.

Theorem (FONC)

Let $x^\sharp$ be a local minimum of $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. If f is differentiable at $x^\sharp$, then

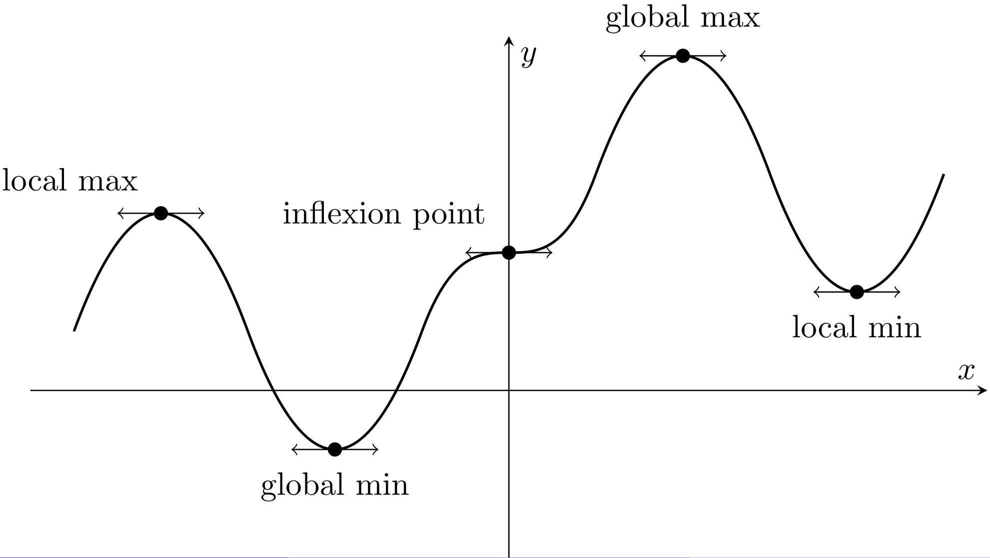
$$\nabla f(x^\sharp) = \mathbf{0}.$$

A second-order sufficient condition is:

Theorem (SOSC)

Assume $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is twice differentiable on its domain. If $x^\sharp \in \operatorname{dom} f$, $\nabla f(x^\sharp) = \mathbf{0}$, and $\nabla^2 f(x^\sharp)$ is positive definite, then $x^\sharp$ is a local minimum.

Critical points



Contents

1 Introduction
- What is optimization?
- Vocabulary
- Practice and tools in optimization

2 First theoretical elements
- Existence of an optimal solution
- Unconstrained optimization

3 Optimization under equality constraints
- Lagrange multipliers
- Tangent cone and constraint qualification

4 Wrap-up

Equality constraints: Lagrange multipliers

♡

Consider

$$\begin{array}{lll} \mathrm{Min}_{x \in \mathbb{R}^n} & f(x) & \\ \text{s.t.} & g_i(x) = 0, & i \in [n_E], \end{array}$$

where f and the g_i are differentiable, and let

$$X := \{x \in \mathbb{R}^n \mid g_i(x) = 0, \ i \in [n_E]\}.$$

Theorem (Lagrange multipliers)

Let $x^\sharp$ be a local minimum of f on X. Under a technical regularity condition on the constraints at $x^\sharp$, there exists $\boldsymbol{\lambda} \in \mathbb{R}^{n_E}$ such that

$$\nabla f(x^\sharp) + \sum_{i=1}^{n_E} \lambda_i \nabla g_i(x^\sharp) = \mathbf{0}.$$

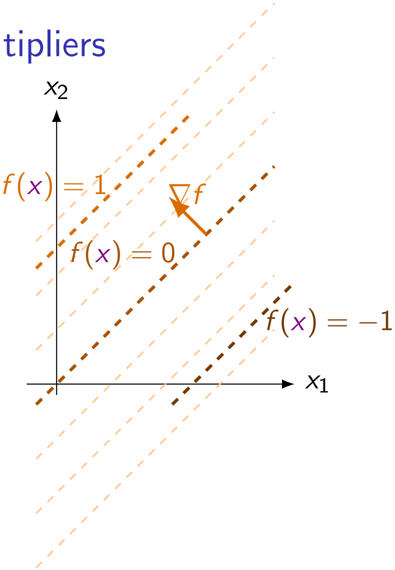
We will make the regularity condition precise later.

Geometric interpretation of Lagrange multipliers

$$f(x) = x_2 - x_1$$
$$g(x) = x_1^2 - x_2$$
$$X = \{x \in \mathbb{R}^2 : g(x) = 0\}$$

$$\nabla f(x) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Level lines of f are parallel, and ∇f is orthogonal to them.



Geometric interpretation of Lagrange multipliers

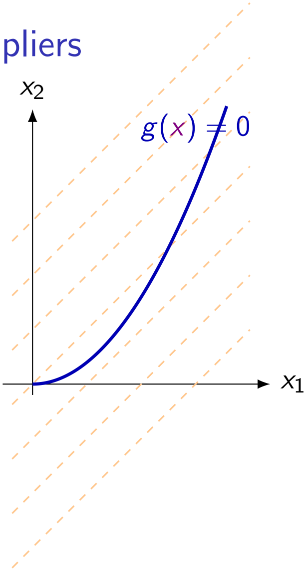
$$f(x) = x_2 - x_1$$

$$g(x) = x_1^2 - x_2$$

$$X = \{x \in \mathbb{R}^2 : g(x) = 0\}$$

$$X = \{(x_1, x_2) : x_2 = x_1^2\}$$

The feasible set is the parabola $g(x) = 0$.



Geometric interpretation of Lagrange multipliers

$$f(x) = x_2 - x_1$$

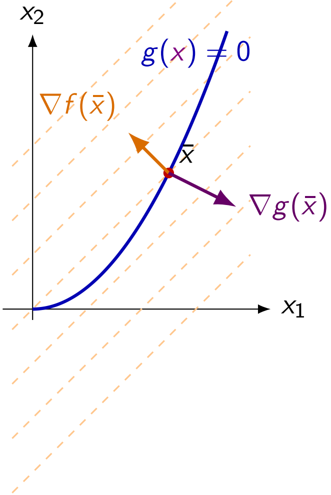
$$g(x) = x_1^2 - x_2$$

$$X = \{x \in \mathbb{R}^2 : g(x) = 0\}$$

$$\bar{x} = (1,1)$$

$$\nabla f(\bar{x}) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \qquad \nabla g(\bar{x}) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

At $\bar{x}$, the gradients are not collinear:
$\bar{x}$ is not a constrained optimum.



Geometric interpretation of Lagrange multipliers

$$f(x) = x_2 - x_1$$

$$g(x) = x_1^2 - x_2$$

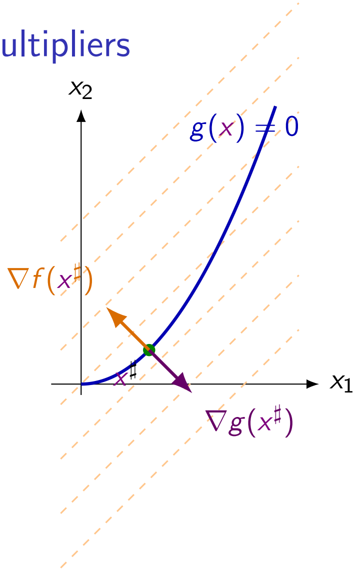
$$X = \{x \in \mathbb{R}^2 : g(x) = 0\}$$

$$x^\sharp = \left(\tfrac{1}{2}, \tfrac{1}{4}\right)$$

$$\nabla f(x^\sharp) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \qquad \nabla g(x^\sharp) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

At $x^\sharp$, the gradients are collinear:

$$\nabla f(x^\sharp) + \nabla g(x^\sharp) = \mathbf{0}.$$

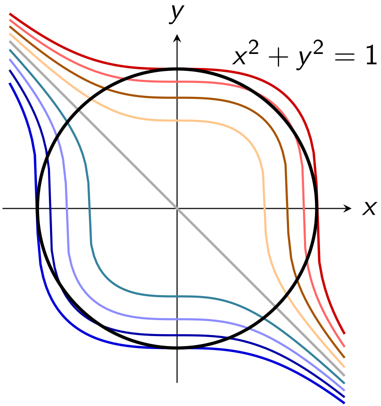


Lagrange multipliers: example

Consider

$$\begin{array}{ll} \text{Min} & x^3 + y^3 \\ \text{s.t.} & x^2 + y^2 = 1. \end{array}$$

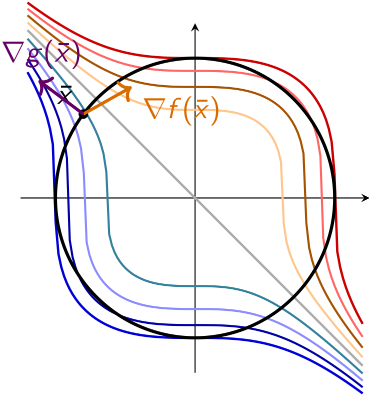
The feasible set is the unit circle. The objective is represented on the right by its level lines.



Lagrange multipliers: example

Consider

$$\begin{array}{ll} \text{Min} & x^3 + y^3 \\ \text{s.t.} & x^2 + y^2 = 1. \end{array}$$



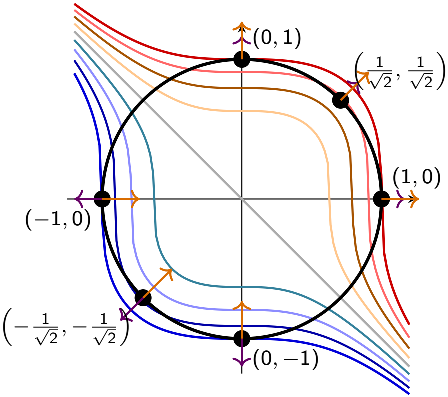
$$\nabla f(x,y) = \begin{pmatrix} 3x^2 \\ 3y^2 \end{pmatrix}, \qquad \nabla g(x,y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}.$$

At an optimum, the two gradients must be collinear.

Lagrange multipliers: example

Consider

$$\begin{array}{ll} \text{Min} & x^3 + y^3 \\ \text{s.t.} & x^2 + y^2 = 1. \end{array}$$



At any optimal solution $(x^\sharp, y^\sharp)$, there exists $\lambda \in \mathbb{R}$ such that

$$\begin{cases} 3x^{\sharp 2} + 2\lambda x^\sharp = 0, \\ 3y^{\sharp 2} + 2\lambda y^\sharp = 0, \\ x^{\sharp 2} + y^{\sharp 2} = 1. \end{cases}$$

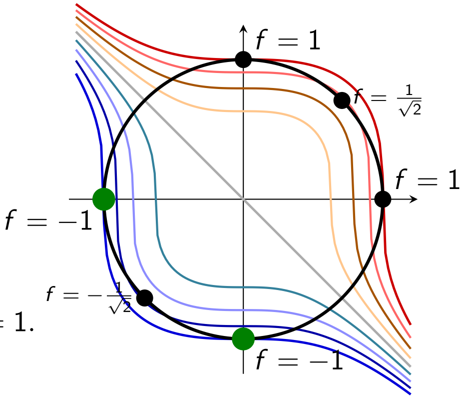
Hence every optimum is among

$$(\pm 1, 0),\ (0, \pm 1),\ \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right),\ \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right).$$

Lagrange multipliers: example

Consider

$$\begin{array}{ll} \text{Min} & x^3 + y^3 \\ \text{s.t.} & x^2 + y^2 = 1. \end{array}$$



Comparing objective values gives

$$f(-1,0) = f(0,-1) = -1,$$

$$f\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = -\frac{1}{\sqrt{2}}, \qquad f(1,0) = f(0,1) = 1.$$

Therefore the two global minimizers are

$$\boxed{(-1,0) \qquad (0,-1)}.$$

Contents

1. Introduction
   - What is optimization?
   - Vocabulary
   - Practice and tools in optimization
2. First theoretical elements
   - Existence of an optimal solution
   - Unconstrained optimization
3. Optimization under equality constraints
   - Lagrange multipliers
   - Tangent cone and constraint qualification
4. Wrap-up

Unconstrained first-order optimality

Let $x^\sharp$ be a local minimum of a differentiable function f.
For any direction $d \in \mathbb{R}^n$ and $t > 0$ small enough,

$$f(x^\sharp + td) = f(x^\sharp) + t\nabla f(x^\sharp) \cdot d + o(t).$$

Since $x^\sharp$ is a local minimum,

$$f(x^\sharp + td) - f(x^\sharp) \geq 0,$$

hence, dividing by t and letting $t \downarrow 0$,

$$\nabla f(x^\sharp) \cdot d \geq 0 \qquad \forall d \in \mathbb{R}^n.$$

Choose $d = -\nabla f(x^\sharp)$:

$$\nabla f(x^\sharp) \cdot d = -\|\nabla f(x^\sharp)\|^2 \leq 0.$$

Therefore

$$\boxed{\nabla f(x^\sharp) = \mathbf{0}.}$$

Unconstrained first-order optimality

Let $x^\sharp$ be a local minimum of a differentiable function f.
For any direction $d \in \mathbb{R}^n$ and $t > 0$ small enough,

$$f(x^\sharp + td) = f(x^\sharp) + t\nabla f(x^\sharp) \cdot d + o(t).$$

Since $x^\sharp$ is a local minimum,

$$f(x^\sharp + td) - f(x^\sharp) \geq 0,$$

hence, dividing by t and letting $t \downarrow 0$,

$$\nabla f(x^\sharp) \cdot d \geq 0 \qquad \forall d \in \mathbb{R}^n.$$

Choose $d = -\nabla f(x^\sharp)$:

$$\nabla f(x^\sharp) \cdot d = -\|\nabla f(x^\sharp)\|^2 \leq 0.$$

Therefore

$$\boxed{\nabla f(x^\sharp) = \mathbf{0}.}$$

Unconstrained first-order optimality

Let $x^\sharp$ be a local minimum of a differentiable function f.
For any direction $d \in \mathbb{R}^n$ and $t > 0$ small enough,

$$f(x^\sharp + td) = f(x^\sharp) + t\nabla f(x^\sharp) \cdot d + o(t).$$

Since $x^\sharp$ is a local minimum,

$$f(x^\sharp + td) - f(x^\sharp) \geq 0,$$

hence, dividing by t and letting $t \downarrow 0$,

$$\nabla f(x^\sharp) \cdot d \geq 0 \qquad \forall d \in \mathbb{R}^n.$$

Choose $d = -\nabla f(x^\sharp)$:

$$\nabla f(x^\sharp) \cdot d = -\|\nabla f(x^\sharp)\|^2 \leq 0.$$

Therefore

$$\boxed{\nabla f(x^\sharp) = \mathbf{0}.}$$

What does local optimality imply under constraints? $\diamondsuit$

Let $x^\sharp$ be a local minimum of f on X. Consider

$$x^k \in X, \qquad x^k \to x^\sharp.$$

By differentiability, for k large enough,

$$f(x^k) = f(x^\sharp) + \nabla f(x^\sharp)^\top (x^k - x^\sharp) + o(\|x^k - x^\sharp\|) \geq f(x^\sharp)$$

Hence

$$\nabla f(x^\sharp)^\top (x^k - x^\sharp) + o(\|x^k - x^\sharp\|) \geq 0.$$

Set

$$t_k := \|x^k - x^\sharp\|, \qquad \boldsymbol{d}^k := \frac{x^k - x^\sharp}{t_k}.$$

Dividing by t_k gives

$$\nabla f(x^\sharp)^\top \boldsymbol{d}^k + o(1) \geq 0.$$

What does local optimality imply under constraints? ♢

Let $x^\sharp$ be a local minimum of f on X. Consider

$$x^k \in X, \qquad x^k \to x^\sharp.$$

By differentiability, for k large enough,

$$f(x^k) = f(x^\sharp) + \nabla f(x^\sharp)^\top (x^k - x^\sharp) + o(\|x^k - x^\sharp\|) \geq f(x^\sharp)$$

Hence

$$\nabla f(x^\sharp)^\top (x^k - x^\sharp) + o(\|x^k - x^\sharp\|) \geq 0.$$

Set

$$t_k := \|x^k - x^\sharp\|, \qquad \boldsymbol{d}^k := \frac{x^k - x^\sharp}{t_k}.$$

Dividing by t_k gives

$$\nabla f(x^\sharp)^\top \boldsymbol{d}^k + o(1) \geq 0.$$

What does local optimality imply under constraints? ◊

Let $x^\sharp$ be a local minimum of f on X. Consider

$$x^k \in X, \qquad x^k \to x^\sharp.$$

By differentiability, for k large enough,

$$f(x^k) = f(x^\sharp) + \nabla f(x^\sharp)^\top (x^k - x^\sharp) + o(\|x^k - x^\sharp\|) \geq f(x^\sharp)$$

Hence

$$\nabla f(x^\sharp)^\top (x^k - x^\sharp) + o(\|x^k - x^\sharp\|) \geq 0.$$

Set

$$t_k := \|x^k - x^\sharp\|, \qquad \boldsymbol{d}^k := \frac{x^k - x^\sharp}{t_k}.$$

Dividing by t_k gives

$$\nabla f(x^\sharp)^\top \boldsymbol{d}^k + o(1) \geq 0.$$

The directions allowed by the constraints

Since $\|\boldsymbol{d}^k\| = 1$, we can extract a subsequence such that

$$\boldsymbol{d}^k \to d.$$

Passing to the limit,

$$\nabla f(x^\sharp)^\top d \geq 0.$$

This motivates collecting all directions that can arise from feasible points approaching $x^\sharp$.

Definition (Tangent cone)

$$T_X(x^\sharp) := \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \exists\, t_k \searrow 0,\ \exists\, \boldsymbol{d}^k \to d, \\ x^\sharp + t_k \boldsymbol{d}^k \in X \end{array} \right\}.$$

Therefore, at a constrained local minimum,

$$\boxed{\nabla f(x^\sharp)^\top d \geq 0 \qquad \forall d \in T_X(x^\sharp).}$$

The directions allowed by the constraints

Since $\|\boldsymbol{d}^k\| = 1$, we can extract a subsequence such that

$$\boldsymbol{d}^k \to d.$$

Passing to the limit,

$$\nabla f(x^\sharp)^\top d \geq 0.$$

This motivates collecting all directions that can arise from feasible points approaching $x^\sharp$.

Definition (Tangent cone)

$$T_X(x^\sharp) := \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \exists\, t_k \searrow 0,\ \exists\, \boldsymbol{d}^k \to d, \\ x^\sharp + t_k \boldsymbol{d}^k \in X \end{array} \right\}.$$

Therefore, at a constrained local minimum,

$$\boxed{\nabla f(x^\sharp)^\top d \geq 0 \qquad \forall d \in T_X(x^\sharp).}$$

The directions allowed by the constraints

Since $\|\boldsymbol{d}^k\| = 1$, we can extract a subsequence such that

$$\boldsymbol{d}^k \to d.$$

Passing to the limit,

$$\nabla f(x^\sharp)^\top d \geq 0.$$

This motivates collecting all directions that can arise from feasible points approaching $x^\sharp$.

Definition (Tangent cone)

$$T_X(x^\sharp) := \left\{ d \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \exists\, t_k \searrow 0,\ \exists\, \boldsymbol{d}^k \to d, \\ x^\sharp + t_k \boldsymbol{d}^k \in X \end{array} \right\}.$$

Therefore, at a constrained local minimum,

$$\boxed{\nabla f(x^\sharp)^\top d \geq 0 \qquad \forall d \in T_X(x^\sharp).}$$

Linearizing the constraints

Suppose now

$$X = \{x \in \mathbb{R}^n \mid g_i(x) = 0, \; i \in [n_E]\}.$$

Take $d \in T_X(x^\sharp)$ and corresponding feasible points x^k.
Since $g_i(x^k) = g_i(x^\sharp) = 0$,

$$0 = g_i(x^k) - g_i(x^\sharp) = t_k \nabla g_i(x^\sharp) \cdot d + o(t_k).$$

Therefore every tangent direction satisfies

$$\nabla g_i(x^\sharp) \cdot d = 0 \qquad \forall i \in [n_E].$$

Define the linearized tangent cone

$$T_X^\ell(x^\sharp) := \left\{ d \in \mathbb{R}^n \mid \nabla g_i(x^\sharp) \cdot d = 0, \; \forall i \right\}.$$

Hence

$$\boxed{T_X(x^\sharp) \subseteq T_X^\ell(x^\sharp).}$$

Linearizing the constraints

Suppose now

$$X = \{x \in \mathbb{R}^n \mid g_i(x) = 0,\ i \in [n_E]\}.$$

Take $d \in T_X(x^\sharp)$ and corresponding feasible points x^k.
Since $g_i(x^k) = g_i(x^\sharp) = 0$,

$$0 = g_i(x^k) - g_i(x^\sharp) = t_k \nabla g_i(x^\sharp) \cdot d + o(t_k).$$

Therefore every tangent direction satisfies

$$\nabla g_i(x^\sharp) \cdot d = 0 \qquad \forall i \in [n_E].$$

Define the linearized tangent cone

$$T_X^\ell(x^\sharp) := \left\{ d \in \mathbb{R}^n \mid \nabla g_i(x^\sharp) \cdot d = 0,\ \forall i \right\}.$$

Hence

$$\boxed{T_X(x^\sharp) \subseteq T_X^\ell(x^\sharp).}$$

Constraint qualification and Lagrange multipliers

Linearization may introduce directions that are not truly feasible.

We say that the constraints are qualified at $x^\sharp$ if

$$\boxed{T_X(x^\sharp) = T_X^\ell(x^\sharp).}$$

Let G be the Jacobian of $g = (g_1, \dots, g_{n_E})$. Under qualification,

$$\nabla f(x^\sharp) \cdot d \geq 0 \qquad \forall d \in \ker G.$$

But $\ker G$ is a vector space: if $d \in \ker G$, then $-d \in \ker G$. Thus

$$\nabla f(x^\sharp) \cdot d = 0 \qquad \forall d \in \ker G.$$

Therefore

$$\nabla f(x^\sharp) \in (\ker G)^\perp = \mathrm{Im}(G^\top),$$

so there exists λ such that

$$\boxed{\nabla f(x^\sharp) + G^\top \lambda = 0.}$$

Constraint qualification and Lagrange multipliers

Linearization may introduce directions that are not truly feasible.
We say that the constraints are qualified at $x^\sharp$ if

$$\boxed{T_X(x^\sharp) = T_X^\ell(x^\sharp).}$$

Let G be the Jacobian of $g = (g_1, \dots, g_{n_E})$. Under qualification,

$$\nabla f(x^\sharp) \cdot d \geq 0 \qquad \forall d \in \ker G.$$

But $\ker G$ is a vector space: if $d \in \ker G$, then $-d \in \ker G$. Thus

$$\nabla f(x^\sharp) \cdot d = 0 \qquad \forall d \in \ker G.$$

Therefore

$$\nabla f(x^\sharp) \in (\ker G)^\perp = \mathrm{Im}(G^\top),$$

so there exists λ such that

$$\boxed{\nabla f(x^\sharp) + G^\top \lambda = 0.}$$

Constraint qualification and Lagrange multipliers

Linearization may introduce directions that are not truly feasible.
We say that the constraints are qualified at $x^\sharp$ if

$$\boxed{T_X(x^\sharp) = T_X^\ell(x^\sharp).}$$

Let G be the Jacobian of $g = (g_1, \dots, g_{n_E})$. Under qualification,

$$\nabla f(x^\sharp) \cdot d \geq 0 \qquad \forall d \in \ker G.$$

But $\ker G$ is a vector space: if $d \in \ker G$, then $-d \in \ker G$. Thus

$$\nabla f(x^\sharp) \cdot d = 0 \qquad \forall d \in \ker G.$$

Therefore

$$\nabla f(x^\sharp) \in (\ker G)^\perp = \mathrm{Im}(G^\top),$$

so there exists $\boldsymbol{\lambda}$ such that

$$\boxed{\nabla f(x^\sharp) + G^\top \boldsymbol{\lambda} = 0.}$$

Simple sufficient qualification conditions

How can we check qualification in practice?

Proposition

If all equality constraints g_i are affine, then they are qualified at every feasible point.

Corollary (LICQ)

Assume the g_i are continuously differentiable near x. If

$$\{\nabla g_i(x) : i \in [n_E]\}$$

is linearly independent, then the constraints are qualified at x.

These conditions are sufficient, not necessary.

Simple sufficient qualification conditions

How can we check qualification in practice?

Proposition

If all equality constraints g_i are affine, then they are qualified at every feasible point.

Corollary (LICQ)

Assume the g_i are continuously differentiable near x. If

$$\{\nabla g_i(x) : i \in [n_E]\}$$

is linearly independent, then the constraints are qualified at x.

These conditions are sufficient, not necessary.

Simple sufficient qualification conditions

How can we check qualification in practice?

Proposition

If all equality constraints g_i are affine, then they are qualified at every feasible point.

Corollary (LICQ)

Assume the g_i are continuously differentiable near x. If

$$\{\nabla g_i(x) : i \in [n_E]\}$$

is linearly independent, then the constraints are qualified at x.

These conditions are sufficient, not necessary.

What you should know

- The basic vocabulary of an optimization problem: variables, objective, constraints, feasible set, local/global optimum, optimal solution/value.
- The main sufficient conditions ensuring existence of an optimal solution.
- The first- and second-order optimality conditions for unconstrained differentiable problems.
- The Lagrange multiplier condition for equality-constrained problems.
- The practical workflow: model $\rightarrow$ formalize $\rightarrow$ solve $\rightarrow$ evaluate/revise.

What you should be able to do

- Formulate a simple optimization problem from an engineering question.
- Check standard sufficient conditions for existence of an optimum.
- Use first- and second-order conditions to study a small unconstrained problem.
- Solve a small equality-constrained problem using Lagrange multipliers.