PEOPELE’S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

Faculty of Sciences

Setif 1 University - Ferhat ABBAS

Optimization 1

Lecture Notes

Prepared at Department of Mathematics

By

Dr. RAHAL Mohamed

Academic year 2024-2025

Contents

General Introduction

What is Optimization?

Optimization involves searching for the "best" element from a given set. The study of the properties of optimal solutions precisely forms the objectives of optimization. It is a branch of applied mathematics and numerical analysis, which has been developing for several years and shows relationships with many other fields of mathematics. This topic examines whether local and global extrema exist for a function of one or more variables, with or without constraints.

This document is particularly intended for undergraduate students (L3) in mathematics, in accordance with the curriculum of this program. It serves as a course support rich in exercises and numerical examples on unconstrained optimization. It consists of three chapters.

In the first chapter, we recall some concepts of differential calculus and notions of convexity that are useful for the rest of the document. At the end of this chapter, a series of exercises is provided, along with a sample exam question with detailed solutions.

The second chapter presents the conditions for existence and uniqueness for a nonlinear optimization problem without constraints. We will then present the necessary and sufficient conditions for optimality in the case of a general unconstrained optimization problem and in the convex case.

The third chapter is dedicated to algorithms for solving a nonlinear optimization problem without constraints. This chapter concludes with a series of exercises and an exam question without solutions.

Chapter 1

A brief review of differential calculus and convexity

Contents

1.1 Introduction

In this chapter, we review some basic concepts of differential calculus and the notion of convexity for sets and functions, which will be essential for the following chapters, particularly the second and third chapters.

1.2 Differentiability

Let n be a natural integer and $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} ... \times \mathbb{R}$ be the usual Euclidean space equipped with the inner product denoted by $\langle .;. \rangle$. We denote by $\|.\|$ the associated norm.
For any $x \in \mathbb{R}^n$, we denote by $x = (x_1, x_2, ..., x_n)^\intercal$ as the column vector.

$$\begin{aligned}\forall x &\in \mathbb{R}^n,\ \forall y \in \mathbb{R}^n \quad \langle x; y\rangle = \sum_{i=1}^{n} x_i y_i \\ \text{and } \|x\|_2 &= \sqrt{\langle x; x\rangle} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}.\end{aligned}$$

The canonical basis of $\mathbb{R}^n$ is denoted by $e_1, e_2, ..., e_n$, where e_k is the vector whose k-th component is 1 and 0 elsewhere.

$$\begin{aligned} e_i &\in \mathbb{R}^n, i = 1, ..., n \\ \text{and } (e_i)_j &= \gamma_{ij} = \begin{cases} 1 & si \quad i = j \\ 0 & \text{else} \end{cases}, j = 1, ..., n \end{aligned}$$

$e_1 = (1, 0, ..., 0)^\intercal, e_2 = (0, 1, ..., 0)^\intercal, ...,\ e_n = (0, 0, ..., 1)^\intercal.$

Definition 1.2.1 *Let Ω be an open subset of $\mathbb{R}^n$ and $a \in \Omega$. Let $f : \Omega \longrightarrow \mathbb{R}$. We say that f is continuous at a if $\lim\limits_{x \longrightarrow a} f(x) = f(a)$. In other words, f is continuous at $a \in \Omega$ if and only if*

$$\forall\ \varepsilon > 0 \ \ \exists\ \tau > 0 \quad \textit{such that } \|x - a\| < \tau \Longrightarrow |f(x) - f(a)| < \varepsilon.$$

Remark 1.2.1

1. *A function f is said to be continuous on Ω if f is continuous at every point $a \in \Omega$.*
2. *If*

$$\begin{aligned} f : \Omega \subseteq \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x &\longmapsto f(x) = (f_1(x), f_2(x), ..., f_m(x))^\intercal \end{aligned}$$

   then f is continuous at $a \in \Omega$ if and only if each component function f_j is continuous at a for every $j = 1, ..., m$.
3. *f is continuous at $a \in \Omega$ if and only if for every sequence $(x^{(k)})_k$ in Ω such that $x^{(k)} \longrightarrow a$ it follows that $f(x^{(k)}) \longrightarrow f(a)$.*

Definition 1.2.2 *Let $f : \Omega \subseteq \mathbb{R}^n \longrightarrow R$ and $v \in \mathbb{R}^n \diagdown \{0_{\mathbb{R}^n}\}$. The directional derivative of f at $a \in \Omega$ in the direction of the vector v, if it exists, is given by:*

$$\frac{\partial f}{\partial v}(a) = \lim_{t \longrightarrow 0} \frac{f(a + tv) - f(a)}{t}.$$

..

Remark 1.2.2

1. *Let*

$$f:\Omega\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^m$$
$$x\longrightarrow f(x)=(f_j(x))_{1\leq j\leq m}$$

then $\dfrac{\partial f}{\partial v}(a)=\left(\dfrac{\partial f_1}{\partial v}(a),......,\dfrac{\partial f_m}{\partial v}(a)\right).$

2. $\forall x=(x_i)_{1\leq i\leq n}\in\mathbb{R}^n$

$$\begin{aligned}\frac{\partial f}{\partial x_i}(a)=\frac{\partial f}{\partial e_i}(a)&=\lim_{t\longrightarrow 0}\frac{f(a+te_i)-f(a)}{t}\\&=\lim_{t\longrightarrow 0}\frac{f(a_1,...,a_i+t,a_{i+1},...,a_n)-f(a_1,..,a_i,a_{i+1},...,a_n)}{t}.\end{aligned}$$

Definition 1.2.3 *Let* $f:\Omega\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^m$. *We say that* f *is differentiable at* $a\in\Omega$ *if there exists a linear map* L *from* $\mathbb{R}^n$ *to* $\mathbb{R}^m$ *such that, for all* $h\in\mathbb{R}^n$, *with* $a+h\in\Omega$, *we have:*

$$f(a+h)=f(a)+L(h)+\|h\|\varepsilon(h),\ \textit{where } \varepsilon(h)\longrightarrow 0 \textit{ as } h\longrightarrow 0.$$

The linear map L *is denoted by* $L(a)$, $df(a)$ *or* $Df(a)$

Lemma 1.2.1

1. *If* f *is differentiable at* $a\in\Omega\Longrightarrow f$ *is continuous at* a.
2. *If* f *is est differentiable at* a *then the first-order partial derivatives of* f *exist at* a.

Special cases

1. When $f:\Omega\subseteq\mathbb{R}\longrightarrow\mathbb{R}$ $(m=1,n=1)$ then the linear map $L(a)=f\prime(a)$, ($L(a)$ is simply the derivative of f at a).
2. When $f:\Omega\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}$ $(m=1)$ then
$$L(a)=\left(\frac{\partial f}{\partial x_1}(a),\frac{\partial f}{\partial x_2}(a),...,\frac{\partial f}{\partial x_n}(a)\right).$$
In this case (where$L(a)$ represents a linear transformation), the vector $L(a)$ is the transpose of a vector called the gradient of f at point a, denoted as $\nabla f(a)$.
3. When $f:\Omega\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^m$, the linear map $L(a)$ is a matrix of order $m\times n$ $(L=L(a)\in\mathcal{M}_{m\times n}(\mathbb{R}))$. This matrix is called the Jacobian matrix of f at a, and it's denoted as $J_f(a)$.
$$J_f(a)=\left(\frac{\partial f_i}{\partial x_j}(a)\right)_{\substack{1\leq i\leq m\\1\leq j\leq n}}=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(a)&\cdots&\frac{\partial f_1}{\partial x_j}(a)&\cdots&\frac{\partial f_1}{\partial x_n}(a)\\\vdots&\ddots&\vdots&\ddots&\vdots\\\frac{\partial f_m}{\partial x_1}(a)&\cdots&\frac{\partial f_m}{\partial x_j}(a)&\cdots&\frac{\partial f_m}{\partial x_n}(a)\end{pmatrix}$$

Examples.

1. $$\begin{array}{rcl} f : \mathbb{R} & \longrightarrow & \mathbb{R} \\ t & \longmapsto & f(t) = e^t - \cos(t) + 3t^2. \end{array}$$

   Let $a \in \mathbb{R}$. Then $df(a) = e^a + \sin(a) + 6a$.

2. $$\begin{array}{rcl} f : \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} & \longmapsto & f(x) = e^{x_1+x_2} - 3x_2x_3. \end{array}$$

   Let $a = (a_1, a_2, a_3)^t \in \mathbb{R}^3$. Then, $L(a) = df(a) = \nabla f(a)^t = \left(e^{a_1+a_2}, e^{a_1+a_2} - 3a_3, -3a_2\right)$.

3. $$\begin{array}{rcl} f : \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \longmapsto & f(x) = \begin{pmatrix} 2x_1x_2 \\ 4x_1 + e^{x_2} \\ x_1 \end{pmatrix} \end{array}$$

   Let $a = (a_1, a_2)^t \in \mathbb{R}^2$. Then, $L(a) = df(a_1, a_2) = J_f(a) = \begin{pmatrix} 2a_2 & 2a_1 \\ 4 & e^{a_2} \\ 1 & 0 \end{pmatrix}$.

Definition 1.2.4 *Let $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ and $a \in \Omega$. We say that f is twice differentiable at a if there exists a symmetric matrix A (a linear map from $\mathbb{R}^n$ to $\mathbb{R}^n$) of order n such that: $\forall h \in \mathbb{R}, a + h \in \Omega$, we have:*

$$f(a+h) = f(a) + (\nabla f(a))^\intercal h + h^\intercal A h + \|h\|^2 \varepsilon(h),$$

where $\lim\limits_{h \longrightarrow 0} \varepsilon(h) = 0$.

Lemma 1.2.2

1. *f is twice differentiable $\Longrightarrow$ f is differentiable.*
2. *If f is twice differentiable, then the partial derivatives up to order 2 exist.*
3. *Let $k \in \mathbb{N}^*$. We say that f is of class C^k on Ω and write $f \in C^k(\Omega)$ if the partial derivatives up to order k exist and are continuous.*

Definition 1.2.5 *Let f be a twice differentiable function at the point $a \in \Omega$. The Hessian matrix of f at a is the matrix denoted $H_f(a)$ or $\nabla^2 f(a)$, defined by:*

$$H_f(a) = \nabla^2 f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{\substack{1 \le i \le n \\ 1 \le j \le n}}.$$

1. *If the function f is twice differentiable at the point a, then by Schwarz's theorem,*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a), \quad \forall i = 1, ..., n \ \text{ and } \ j = 1, ..., n$$

2. *The Hessian matrix of f at the point $a \in \Omega$ is always symmetric.*

3. *If $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ is of class C^2 on Ω, then*

- $\dfrac{\partial f}{\partial \upsilon}(a) \ = \ \langle \nabla f(a), \upsilon \rangle, \quad \forall \upsilon \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}.$
- $H_f(a) = J_{\nabla f}(a) = \nabla_{J_f}(a).$
- $H_f(a) \cdot \upsilon = \nabla \langle \nabla f(a), \upsilon \rangle, \quad \forall a \in \Omega, \forall \upsilon \in \mathbb{R}^n.$

Examples.

1. *Let*

$$\begin{array}{rccl} f & : \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & x & \longmapsto & Ax \end{array}$$

where $A \in M_{m \times n}(\mathbb{R})$, then $J_f(x) = A$, $\nabla f(x) = (J_f(x))^{\intercal} = A^{\intercal}$.

2.

$$\begin{array}{rccl} f & : \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \langle a, x \rangle + b \end{array}$$

$\nabla f(x) = a$ and $H_f(x) = \nabla^2 f(x) = 0.$

3.

$$\begin{array}{rccl} f & : \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \langle Ax, x \rangle = x^{\intercal} A x = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_j x_i \end{array}$$

where $A \in M_n(\mathbb{R})$, the function f in this case is called a quadratic function.

Show that:

- $\nabla f(x) = (A + A^{\intercal})x, \quad \forall x \in \mathbb{R}^n.$
- $H_f(x) = \nabla^2 f(x) = A + A^{\intercal}.$
- *When A is symmetric, then:* $\begin{cases} \nabla f(x) & = 2Ax \\ \nabla^2 f(x) & = 2A \end{cases}.$

Theorem 1.2.1 *(Taylor Expansion).*
Let U be an open subset of $\mathbb{R}^n$, $a \in U$, and $f : U \longrightarrow \mathbb{R}$. Let $h \in \mathbb{R}^n$ such that the line segment $[a, a + h]$ is contained in U.

1. *If $f \in C^1(U)$, then*

   (a) *The first-order Taylor formula with Maclaurin remainder is given by:*

   $$f(a+h) = f(a) + \langle \nabla f(a+\theta h), h\rangle, \quad 0 < \theta < 1.$$

   (b) *The first-order Taylor formula with Young remainder is given by:*

   $$f(a+h) = f(a) + \langle \nabla f(a), h\rangle + o(\|h\|).$$

2. *If $f \in C^2(U)$, then*

   (a) *The second-order Maclaurin Taylor formula is given by:*

   $$f(a+h) = f(a) + \nabla f(a)^\intercal h + \frac{1}{2} h^\intercal \nabla^2 f(a+\theta h) h, \quad 0 < \theta < 1.$$

   (b) *The second-order Young Taylor formula is given by:*

   $$f(a+h) = f(a) + \langle \nabla f(a), h\rangle + \frac{1}{2} h^\intercal \nabla^2 f(a) h + o(\|h\|^2).$$

Examples.

1.
$$\begin{array}{rcccl} f & : & I \subset \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & f(x) = x^4 \end{array}$$

We have $f \in C^\infty(I)$

$$\begin{aligned} f(0+h) &= f(0) + f'(0)h + \frac{1}{2}h^2 f''(0) + o(h^2) \\ f(h) &= 0 + 0 + 0 + o(h^2), \end{aligned}$$

2.
$$\begin{array}{rcccl} f & : & U \subset \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & & \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & f(x,y) = \ln(1+xy) \end{array}$$

The domain is $D_f = \{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2;\ 1+xy > 0\}$.
We have $f \in C^2(U)$, so the second-order Young Taylor formula at the point $a = (0,0)^\intercal$ is:

$$\begin{aligned} f(a+h) &= f(a) + \langle \nabla f(a), h\rangle + \frac{1}{2} h^\intercal \nabla^2 f(a) h + o(\|h\|^2) \\ f(a) &= f(0,0) = \ln(1+0) = \ln(1) = 0. \end{aligned}$$

$$\nabla f(x,y) = \begin{pmatrix} \frac{\partial f}{\partial x}(x,y) \\ \frac{\partial f}{\partial y}(x,y) \end{pmatrix} = \begin{pmatrix} \frac{y}{1+xy} \\ \frac{x}{1+xy} \end{pmatrix} \Rightarrow \nabla f(0,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = (0,0)^\intercal.$$

$$\nabla^2 f(x,y) = \begin{pmatrix} \dfrac{-y^2}{(1+xy)^2} & \dfrac{1}{(1+xy)^2} \\ \dfrac{1}{(1+xy)^2} & \dfrac{-x^2}{(1+xy)^2} \end{pmatrix} \Rightarrow \nabla^2 f(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Therefore

$$f(h) = 0 + \langle 0, h\rangle + \frac{1}{2}(h_1,h_2)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + o(h_1^2+h_2^2)$$

$$\begin{aligned} f(h_1,h_2) &= \frac{1}{2}(h_1,h_2)\begin{pmatrix} h_2 \\ h_1 \end{pmatrix} + o(h_1^2+h_2^2) \\ &= h_1h_2 + o(h_1^2+h_2^2). \end{aligned}$$

1.3 Convexity

1.3.1 Convex Set

Definition 1.3.1 *Let C be a non-empty subset of $\mathbb{R}^n$. We say that C is convex if for all $a,b \in C$ the line segment $[a,b] \subseteq C$,*
where

$$[a,b] = \{ta + (1-t)b;\ 0 \le t \le 1\}.$$

In other words:

$$C \text{ convex} \iff \left.\begin{array}{l} \forall\ \ t \in [0,1] \\ \forall\ \ a,b \in C \end{array}\right\} \text{ then } ta + (1-t)b \in C.$$

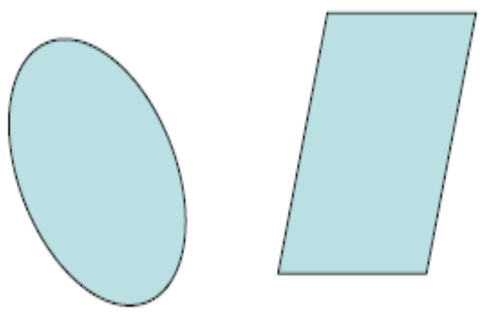

Figure 1.1: Example of convex sets

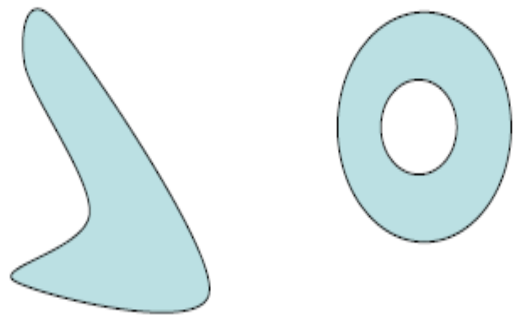

Figure 1.2: Example of non-convex sets

Examples.

1. If $\|\cdot\|$ is a norm on $\mathbb{R}^n$, then the open balls $(B(x_0, r))$ and closed balls $(B_f(x_0, r))$ are convex in $\mathbb{R}^n$ (see topology).
2. The convex sets in $\mathbb{R}$ are exactly the intervals.
3. In $\mathbb{R}^2$, the set $C = \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$ is not convex because: $(1, 0)^\intercal \in C$ and $(0, 1)^\intercal \in C$, but for $t = \frac{1}{2} \in [0, 1]$ we have
$$\frac{1}{2}(1,0)^\intercal + \left(1 - \frac{1}{2}\right)(0,1)^\intercal = \left(\frac{1}{2}, \frac{1}{2}\right)^\intercal \notin C \quad \left(\text{since } \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \neq 0\right).$$
4. In $\mathbb{R}^n$ $(n \geq 2)$, the set $S_n = \{(\alpha_1, \alpha_2, ..., \alpha_n)^\intercal \in \mathbb{R}^n_+ \mid \sum_{i=1}^n \alpha_i = 1\}$ is a convex set of $\mathbb{R}^n$ called the simplex of $\mathbb{R}^n$.

Proposition 1.3.1

1. *The intersection of convex sets is convex.*
2. *The union of two convex sets is not generally convex. For example:*
$$C_1 = \{(x, y)^\intercal \in \mathbb{R}^2 : x = 0\} \text{ and } C_2 = \{(x, y)^\intercal \in \mathbb{R}^2 : y = 0\}$$
are convex sets in $\mathbb{R}^2$*, but* $C_1 \cup C_2$ *is not convex in* $\mathbb{R}^2$*.*
3. *The Cartesian product of convex sets is convex.*

Proposition 1.3.2 *Let* $C \subseteq \mathbb{R}^n$ *be a convex set, then*

1. $C + \alpha = \{x + a \mid x \in C\}$ *is convex.*
2. $\alpha C = \{\alpha x \mid x \in C\}$ *is convex.*
3. *If* $f : \mathbb{R}^n \to \mathbb{R}^m$ *is a linear function, then* $f(C)$ *is convex in* $\mathbb{R}^m$*.*
4. *If* $f : \mathbb{R}^m \to \mathbb{R}^n$ *is a linear mapping, then* $f^{-1}(C)$ *is convex in* $\mathbb{R}^n$*.*

1.3.2 Convex function

Definition 1.3.2 *Let C be a convex subset of $\mathbb{R}^n$ and $f : C \to \mathbb{R}$ be a function defined on C.*

1. *We say that f is convex on C if and only if: $\forall x, y \in C$ and $\forall \theta \in [0, 1]$*
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y).$$
2. *We say that f is strictly convex on C if: $\forall x, y \in C$ $(x \neq y)$ and $\forall \theta \in (0, 1)$*
$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y).$$
3. *We say that f is strongly convex on C if there exists $\alpha > 0$ such that: $\forall x, y \in C$ and $\forall \theta \in [0, 1]$*
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) - \alpha\theta(1-\theta)\|x-y\|^2.$$
4. *We say that f is concave (resp. strictly concave, resp. strongly concave) if $(-f)$ is convex (resp. strictly convex, resp. strongly convex).*

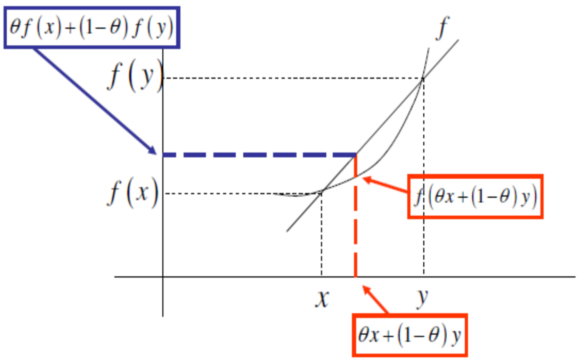

Figure 1.3: Convex function

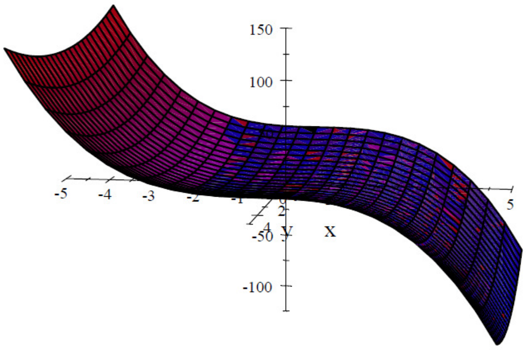

Figure 1.4: Example of non-convex function of two variables

Examples.

1.

$$\begin{array}{cccc} \|\cdot\| & : & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & \|x\| \end{array}$$

2.

$$\begin{array}{cccc} f & : & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & f(x) = \langle x, a\rangle + b \end{array}$$

f is both convex and concave. Let $x, y \in \mathbb{R}^n$ and $t \in [0,1]$:

$$\begin{aligned} f(tx + (1-t)y) &= \langle tx + (1-t)y, a\rangle + b \\ &= \langle tx, a\rangle + \langle (1-t)y, a\rangle + tb + (1-t)b \\ &= t\left(\langle x, a\rangle + b\right) + (1-t)\left(\langle y, a\rangle + b\right) \\ &= tf(x) + (1-t)f(y). \end{aligned}$$

3.

$$\begin{array}{cccc} f & : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & x^2 \end{array}$$

f is a convex function (moreover, f is strictly convex). Indeed, $\forall x, y \in \mathbb{R}$ and $\forall t \in [0,1]$:

$$\begin{aligned} f(tx + (1-t)y) &= (tx + (1-t)y)^2 \\ &= t^2x^2 + (1-t)^2y^2 + 2t(1-t)xy \\ &\le t^2x^2 + (1-t)^2y^2 + t(1-t)(x^2 + y^2) \\ &= tx^2 + (1-t)y^2 \\ &= tf(x) + (1-t)f(y). \end{aligned}$$

Theorem 1.3.1

1. *Let $\{f_i\}_{i\in I}$ be a family of convex functions on U (a convex subset of $\mathbb{R}^n$) mapping to $\mathbb{R}$. Then $\sup_{i\in I} f_i$ is a convex function,*

   where $\left(\sup_{i\in I} f_i\right)(x) = \sup_{i\in I}\left(f_i(x)\right)$.

2. *If $U \subset \mathbb{R}^n$ is convex and $f_i : U \to \mathbb{R}$ $(i = 1, ..., n)$ are convex, then for all $\lambda_i \geq 0$ $(i = 1, ..., n)$, the function $\sum_{i=1}^n \lambda_i f_i$ is convex.*

3. *Let $f : U \to \mathbb{R}$ be convex (U is a convex subset of $\mathbb{R}^n$) and $\varphi : \mathbb{R} \to \mathbb{R}$ be a function. Then:*

   (a) If f is convex and φ is convex and increasing $\Rightarrow \varphi \circ f$ is convex.

   (b) If f is concave and φ is convex and decreasing $\Rightarrow \varphi \circ f$ is convex.

   (c) If f is concave and φ is concave and increasing $\Rightarrow \varphi \circ f$ is concave.

Proof.

1. For each $f_i : U \to \mathbb{R}$ convex $(\forall i \in I)$, and for all $x, y \in U$ and $t \in [0, 1]$:

$$\begin{aligned} f_i(tx + (1-t)y) &\leq t f_i(x) + (1-t) f_i(y) \quad \forall i \in I \\ &\leq t\left(\sup_{i\in I} f_i(x)\right) + (1-t)\left(\sup_{i\in I} f_i(y)\right) \\ &= t\left(\sup_{i\in I} f_i\right)(x) + (1-t)\left(\sup_{i\in I} f_i\right)(y) \end{aligned}$$

   Therefore, $\sup_{i\in I} f_i$ is a convex function.

2. For all $t \in [0, 1]$:

$$\begin{aligned} f(tx + (1-t)y) &\leq t f(x) + (1-t) f(y) && \text{(since } f \text{ is convex)} \\ \varphi(f(tx + (1-t)y)) &\leq \varphi(t f(x) + (1-t) f(y)) && \text{(since } \varphi \text{ is increasing)} \\ &\leq t\varphi(f(x)) + (1-t)\varphi(f(y)) && \text{(since } \varphi \text{ is convex)} \\ \Rightarrow (\varphi \circ f)(tx + (1-t)y) &\leq t(\varphi \circ f)(x) + (1-t)(\varphi \circ f)(y) && \text{(proving } \varphi \circ f \text{ is convex)} \end{aligned}$$

3.

$$\begin{aligned} f \text{ concave} &\Leftrightarrow (-f) \text{ convex} \\ &\Leftrightarrow -f(tx + (1-t)y) \leq t(-f(x)) + (1-t)(-f(y)) \\ &\Leftrightarrow f(tx + (1-t)y) \geq t f(x) + (1-t) f(y). \end{aligned}$$

Examples of Convex Functions

1. The function $(x, y) \mapsto x^2 + y^2$ is convex because:
   - The norm $(x, y) \mapsto \|(x, y)\|$ is convex
   - The function
   $$\begin{array}{rcccl} \varphi & : & \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ & & t & \longmapsto & t^2 \end{array}$$
   is convex and increasing
   - Therefore $\varphi \circ f$ is convex
2. If f is convex, then f^2 is convex (when $f \geq 0$).
3. The p-norm $x \mapsto \|x\|^p$ is convex for all $p \geq 1$.
4. If f is convex, then $\exp(f)$ is convex.

1.4 Convexity and Differentiability

Theorem 1.4.1 *Let $U \subseteq \mathbb{R}^n$ be an open convex subset of $\mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be a differentiable function on U. Then the following three properties are equivalent:*

1. *f is convex on U.*
2. *For all $x, y \in U$: $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$.*
3. *For all $x, y \in U$: $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0$.*

Proof.
1) $\Rightarrow$ 2) For all $x, y \in U$ and all $t \in (0, 1]$, we have:

$$\begin{aligned} & f(ty + (1-t)x) \leq tf(y) + (1-t)f(x) \\ \Rightarrow\ & f(x + t(y-x)) \leq f(x) + t(f(y) - f(x)) \\ \Rightarrow\ & \frac{f(x + t(y-x)) - f(x)}{t} \leq f(y) - f(x) \\ \Rightarrow\ & \lim_{t \to 0^+} \frac{f(x + t(y-x)) - f(x)}{t} \leq f(y) - f(x) \\ & \frac{\partial f}{\partial v}(x) \leq f(y) - f(x) \quad \text{where } v = y - x \\ \Rightarrow\ & \langle \nabla f(x), y - x \rangle \leq f(y) - f(x). \end{aligned}$$

Therefore, 1) $\Rightarrow$ 2)

2) $\Rightarrow$ 3) For all $x, y \in U$:

$$\begin{aligned} & f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle \\ & f(x) - f(y) \geq \langle \nabla f(y), x - y \rangle \\ \Rightarrow\ & 0 \geq \langle \nabla f(x), y - x \rangle + \langle \nabla f(y), x - y \rangle \\ \Rightarrow\ & \langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0. \end{aligned}$$

Therefore, 2) $\Rightarrow$ 3)

3) $\Rightarrow$ 1): For all $x, y \in U$ and $t \in (0,1]$, define:

$$\varphi\colon [0,1] \to \mathbb{R}, \quad \varphi(t) = f(ty + (1-t)x)$$

The function φ is differentiable with:

$$\varphi'(t) = \langle \nabla f(x + t(y-x)), y - x \rangle$$

For any $t_1 \le t_2$:

$$\begin{aligned}
\varphi'(t_2) - \varphi'(t_1) &= \langle \nabla f(x + t_2(y-x)) - \nabla f(x + t_1(y-x)), y - x \rangle \\
&= \frac{1}{t_2 - t_1} \langle \nabla f(x + t_2(y-x)) - \nabla f(x + t_1(y-x)), (t_2 - t_1)(y - x) \rangle \ge 0 \\
&\Rightarrow \varphi' \text{ is monotonically increasing} \\
&\Rightarrow \varphi \text{ is convex}
\end{aligned}$$

Thus:

$$\begin{aligned}
\varphi(t) &\le t\varphi(1) + (1-t)\varphi(0) \\
\Rightarrow\ f(ty + (1-t)x) &\le tf(y) + (1-t)f(x) \\
&\Rightarrow\ f \text{ is convex.}
\end{aligned}$$

Theorem 1.4.2 *Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be a C^2 function where U is an open convex subset of $\mathbb{R}^n$. Then:*

1. *f is convex on U $\iff$ $\nabla^2 f(x) = H_f(x)$ is a positive semidefinite matrix for all $x \in U$.*
2. *If the matrix $\nabla^2 f(x) = H_f(x)$ is positive definite for all $x \in U$, then f is strictly convex on U.*
3. *f is strongly convex on U if and only if there exists $\alpha > 0$ such that:*

$$\langle \nabla^2 f(x)h, h \rangle \ge \alpha \|h\|^2 \quad \textit{for all } x \in U \textit{ and all } h \in \mathbb{R}^n.$$

Proof (Exercise).

Example. Consider the function:

$$\begin{array}{rccc}
f & : & \mathbb{R}^n & \to & \mathbb{R} \\
& & x & \mapsto & f(x) = \langle Ax, x \rangle - \langle b, x \rangle
\end{array}$$

where $A \in M_n(\mathbb{R})$ is symmetric positive definite and $b \in \mathbb{R}^n$.

1. Show that there exists $\alpha > 0$ such that:

$$\langle Ax, x \rangle \ge \alpha \|x\|^2, \quad \text{for all } x \in \mathbb{R}^n.$$

2. Deduce that f is strongly convex (and therefore strictly convex and convex).

Definition 1.4.1 *Let*

$$\begin{array}{rcccc} f & : & U \subset \mathbb{R}^n & \to & \mathbb{R} \\ & & x & \mapsto & f(x) \end{array}$$

We say that f is coercive on U if and only if

$$\lim_{\substack{\|x\|\to+\infty \\ x\in U}} f(x) = +\infty.$$

Thus, f is coercive on $\mathbb{R}^n$ if and only if:

$$\lim_{\|x\|\to+\infty} f(x) = +\infty.$$

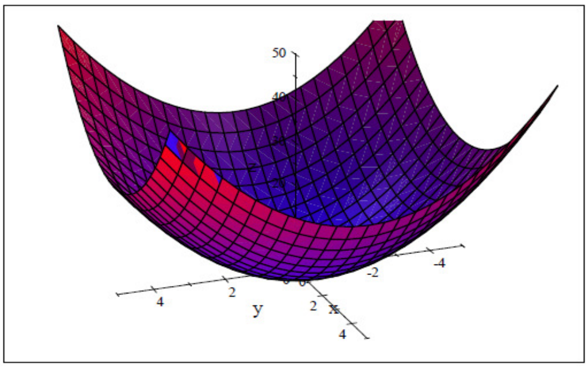

Figure 1.5: Example of coercive function of two variables

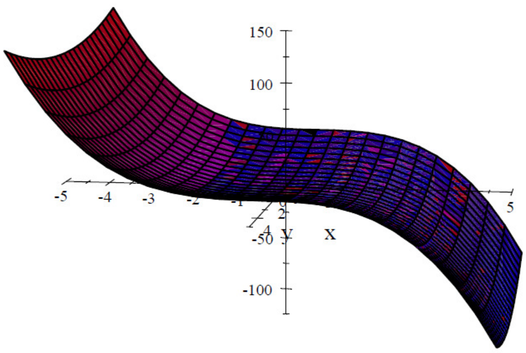

Figure 1.6: Example of non-coercive function of two variables

Examples.

1. The function $x \mapsto \|x\|$ is coercive on any normed vector space.
2. The function $t \mapsto e^t$ is coercive on $\mathbb{R}_+$ but not coercive on $\mathbb{R}$.
3. The function $f(x, y) = \sin(x + y)$ is not coercive.

Recall. (Matrix Calculus) Let $A \in M_n(\mathbb{R})$ be a real square matrix of order n:

1. A is positive semidefinite if and only if $\langle Ax, x\rangle = x^\top Ax \geq 0$ for all $x \in \mathbb{R}^n$.
   A is negative semidefinite if and only if $\langle Ax, x\rangle = x^\top Ax \leq 0$ for all $x \in \mathbb{R}^n$.
2. A is positive definite if and only if $\langle Ax, x\rangle = x^\top Ax > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$.
   A is negative definite if and only if $\langle Ax, x\rangle = x^\top Ax < 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$.
3. If $A \in M_n(\mathbb{R})$ is symmetric:
   $\mathrm{Sp}(A) = \{\lambda_i \mid \lambda_i \text{ is an eigenvalue of } A\} \subseteq \mathbb{R}$.
4. If $A \in M_n(\mathbb{R})$ is symmetric:
   (a) A is positive semidefinite $\Leftrightarrow \lambda_i \geq 0$ for all i.
   (b) A is positive definite $\Leftrightarrow \lambda_i > 0$ for all i.

Theorem 1.4.3 *(Sylvester's Criterion)*
A matrix A is positive definite if and only if $\det(A_k) > 0$ for all $k = 1, \dots, n$, where

$$A_k = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}$$

denotes the k-th leading principal submatrix of A.

1.5 Exercises

Exercise 1.1. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function defined by:

$$f(x,y) = \begin{cases} \frac{|x|^{3/2} y}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

1. Is f continuous on $\mathbb{R}^2$?
2. Compute the partial derivatives of f. Are they continuous?
3. Calculate the directional derivatives (if they exist) of f at the point $(0,0)$.
4. Is the function f differentiable at $(0,0)$?

Exercise 1.2

Expand the following functions around $(0,0)$ up to second order:

$$\begin{aligned} f(x,y) &= \arctan(x+y^2), \\ g(x,y) &= \ln(1+xy). \end{aligned}$$

Exercise 1.3

1. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be defined by:

$$f(x) = Ax, \quad \forall x \in \mathbb{R}^n,$$

where $A \in M_{m,n}(\mathbb{R})$.

(a) Show that $J_f(x) = A$ for all $x \in \mathbb{R}^n$.

(b) Compute $\nabla f(x)$ and $\nabla^2 f(x)$ for the linear function:

$$f(x) = \langle h, x \rangle, \quad \forall x \in \mathbb{R}^n.$$

2. Let $A \in M_n(\mathbb{R})$ and consider the function:

$$\begin{array}{rcl} f : \mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & \langle Ax, x \rangle \end{array}$$

(a) Show that for all $x \in \mathbb{R}^n$:

$$\nabla f(x) = (A + A^\top)x \quad \text{and} \quad \nabla^2 f(x) = A + A^\top.$$

(b) Show that if A is positive definite, then the diagonal elements of A are strictly positive.

Exercise 1.4

1. Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by:

$$f(x) = \begin{cases} x^2 & \text{if } x \geq 0 \\ x^3 & \text{if } x < 0 \end{cases}$$

(a) Study the convexity of f.

(b) Is the function f differentiable on $\mathbb{R}$?

2. Are the following functions coercive?

(a) $f(x_1, x_2) = -x_1^2 - x_2^2$.

(b) $f(x_1, x_2) = -x_2$.

(c) $f(x) = x^T b + c$, where $b, x \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

(d) $f(x) = \frac{1}{2}x^T Ax - b^T x$, where A is a symmetric positive definite square matrix of order n and b is a vector in $\mathbb{R}^n$.

Exercise 1.5

Consider the following subsets:

1. $A = \{(x, y)^T \in \mathbb{R}^2 : xy \leq 0\} \subseteq \mathbb{R}^2$.
2. $B = \{(x, y, z)^T \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 4 \text{ and } z - y = 0\}$.

Provide a graphical representation of each set and determine whether it is convex or not.

Exercise 1.6

Consider the function f defined on $\mathbb{R}^2$ by:

$$f(x, y) = x^4 + y^4 - 12xy.$$

Let Ω be the set defined by:

$$\Omega = \{(x, y)^T \in \mathbb{R}^2 : x > 0,\ y > 0,\ xy > 1\}.$$

1. Show that Ω is an open convex subset of $\mathbb{R}^2$.
2. Show that f is strictly convex on Ω.

Exercise 1.7

Analyze the convexity of the function $f : \mathbb{R}^3 \to \mathbb{R}$ defined by:

$$f(x, y, z) = |x| + |y| + z.$$

Exercise 1.8

Consider the two functions f and g defined by:

$$f(x, y) = x^4 + y^4 \quad \text{and} \quad g(x, y) = (x - y)^2.$$

1. Show that both f and g are convex on $\mathbb{R}^2$.
2. Let $h = f - g$. Show that h is neither convex nor concave on $\mathbb{R}^2$.
3. Is the function $f + g$ convex on $\mathbb{R}^2$?

Exercise 1.9

1. Let U be a convex subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}$ a function. Show that f is convex if and only if for all $x_1, x_2, \ldots, x_m \in U$ and all $\lambda_1, \lambda_2, \ldots, \lambda_m \in [0, 1]$ with $\sum_{i=1}^m \lambda_i = 1$, we have:
$$f\left(\sum_{i=1}^m \lambda_i x_i\right) \leq \sum_{i=1}^m \lambda_i f(x_i) \quad \text{(Jensen's inequality)}$$

2. Deduce that:

$$\ln\left(\sum_{i=1}^{n}\frac{1}{n}x_i\right) \geq \sum_{i=1}^{n}\frac{1}{n}\ln(x_i), \quad \text{with } x_i > 0 \text{ for all } 1 \leq i \leq n.$$

Exercise 1.10
Let f be the function defined on $\mathbb{R}^n$ by:

$$f(x) = (x^\top a)^2 + x^\top x$$

where $a \in \mathbb{R}^n$ $(a \neq 0)$.

1. Compute $\nabla f(x)$ and $H(x) = \nabla^2 f(x)$.
2. Deduce that f is a quadratic form on $\mathbb{R}^n$.
3. Show that f is strictly convex and coercive.

1.6 Exam style with detailed solutions

Exercise 1 (2021). Analyze the convexity (convex, strictly convex, concave, and strictly concave) of the following two functions.

1. $f : \mathbb{R} \to \mathbb{R}$ defined by:
$$f(x) = e^{(x^2+2021)^3}$$
2. $f : \mathbb{R}^2 \to \mathbb{R}$ defined by:
$$f(x, y) = -|x| + y$$

Exercise 2. Consider the following maximization problem:

$$\max_{(x,y)\in\mathbb{R}^2} \left[f(x, y) = -(y - x^2)^2\right].$$

1. Is the function f coercive on $\mathbb{R}^2$?
2. Find all critical points.
3. Show that all critical points satisfy the second-order necessary optimality condition.
4. Does the function f have a global maximum on $\mathbb{R}^2$?
5. Can the function f be strictly concave on $\mathbb{R}^2$?

Exercise 3. Consider the problem:

$$(P)\begin{cases}\min f(x, y, z) = e^{\cos(\pi - x)} + y^2 + 4y + z^2 + 5\\ (x, y, z)^\top \in \mathbb{R}^3\end{cases}$$

1. Is the function f coercive on $\mathbb{R}^3$?
2. Find all critical points of f.
3. Find all local minima of f on $\mathbb{R}^3$.
4. Show that f is bounded below on $\mathbb{R}^3$.
5. Determine the global minima of f on $\mathbb{R}^3$.
6. Can f be convex on $\mathbb{R}^3$? Justify your answer.
7. Can f be concave on $\mathbb{R}^3$? Justify your answer.

Solutions.

Exercise 1.

1. Consider $f(x) = e^{(x^2+2021)^3}$

   The second derivative is:

   $$f''(x) = \underbrace{\left[6(x^2+2021)^2 + 36x^2(x^2+2021)^4 + 24x^2(x^2+2021)\right]}_{>0} e^{(x^2+2021)^3}$$

   - $f''(x) > 0$ for all $x \in \mathbb{R}$, therefore f is strictly convex on $\mathbb{R}$
   - f being strictly convex implies it is convex on $\mathbb{R}$
   - f being strictly convex implies it is not concave on $\mathbb{R}$
   - f not being concave implies it is not strictly concave on $\mathbb{R}$

2. Consider $f(x,y) = -|x| + y$

   Let $U = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, V = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$ and $t \in [0,1]$, then:

   $$tU + (1-t)V = \begin{pmatrix} tu_1 + (1-t)v_1 \\ tu_2 + (1-t)v_2 \end{pmatrix}$$

   We have:

   $$\begin{aligned} f(tU + (1-t)V) &= -|tu_1 + (1-t)v_1| + tu_2 + (1-t)v_2 \\ &\geq -(|tu_1| + |(1-t)v_1|) + tu_2 + (1-t)v_2 \\ &= -t|u_1| - (1-t)|v_1| + tu_2 + (1-t)v_2 \\ &= t(-|u_1| + u_2) + (1-t)(-|v_1| + v_2) \\ &= tf(U) + (1-t)f(V) \end{aligned}$$

   Thus f is concave on $\mathbb{R}^2$.

   - f being concave implies it is not strictly convex on $\mathbb{R}^2$

- For $M_1 = (1,0)^\top$, $M_2 = (-1,0)^\top$ and $t = \frac{1}{2}$:
$$f\left(\tfrac{1}{2}M_1 + \tfrac{1}{2}M_2\right) = 0 \nleq -1 = \tfrac{1}{2}f(M_1) + \tfrac{1}{2}f(M_2)$$
Therefore f is not convex on $\mathbb{R}^2$
- For $M_1 = (1,0)^\top$, $M_2 = (2,0)^\top$ and $t = \frac{1}{2}$:
$$f\left(\tfrac{1}{2}M_1 + \tfrac{1}{2}M_2\right) = -\tfrac{3}{2} = \tfrac{1}{2}f(M_1) + \tfrac{1}{2}f(M_2)$$
Therefore f is not strictly concave on $\mathbb{R}^2$.

Exercise 2. Consider the problem: $\max\limits_{\mathbb{R}^2}\left[f(x,y) = -\left(y - x^2\right)^2\right]$.

1. For $X_n = \left(n, n^2\right)^t$, we have $\|X_n\| \longrightarrow +\infty$ if and only if $n \longrightarrow +\infty$, but $\lim\limits_{\|X_n\|\longrightarrow+\infty} f(X_n) = \lim\limits_{n\longrightarrow+\infty} -\left(n^2 - n^2\right) = 0 \neq -\infty$, hence f is not decreasing at infinity on $\mathbb{R}^2$.

2. $f \in C^\infty\left(\mathbb{R}^2\right)$ and $\nabla f(x,y) = \begin{pmatrix} 4x\left(y - x^2\right) \\ 2\left(x^2 - y\right) \end{pmatrix}$, thus $\nabla f(x,y) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Longleftrightarrow x^2 - y = 0$.

   Therefore, the critical points are $\left\{(x,y)^t \in \mathbb{R}^2 \text{ such that } y = x^2\right\} = \left\{\left(x, x^2\right)^t, x \in \mathbb{R}\right\}$.

3. We have: $H_f(x,y) = \nabla^2 f(x,y) = \begin{pmatrix} 4y - 12x^2 & 4x \\ 4x & -2 \end{pmatrix} \Longrightarrow \nabla^2 f\left(x, x^2\right) = \begin{pmatrix} -8x^2 & 4x \\ 4x & -2 \end{pmatrix}$.

   The eigenvalues of $\nabla^2 f\left(x, x^2\right)$ are 0 and $-\left(8x^2 + 2\right)$.

   Therefore, $\nabla^2 f\left(x, x^2\right)$ is a negative semi-definite matrix, meaning the critical points satisfy the necessary second-order condition.

4. We have: $f(x,y) \leq 0, \forall (x,y)^t \in \mathbb{R}^2$ and $f\left(\alpha, \alpha^2\right) = 0, \forall \alpha \in \mathbb{R}$, hence $\left(\alpha, \alpha^2\right)$ is a global maximum of f. Thus, the function f has an infinite number of global maxima on $\mathbb{R}^2$.

5. The function f has more than 2 maxima, so f is not strictly concave on $\mathbb{R}^2$.

Exercise 3. Consider the problem:
$$(p)\begin{cases} \min f(x,y,z) = \min\left[e^{\cos(\pi - x)} + y^2 + 4y + z^2 + 5\right] \\ (x,y,z)^T \in \mathbb{R}^3 \end{cases}$$

1. The function f is not coercive on $\mathbb{R}^3$ because for $X_n = (n,0,0)^t \in \mathbb{R}^3$, we have:
   $\lim\limits_{\|X_n\|\longrightarrow+\infty} f(X_n) = \lim\limits_{n\longrightarrow+\infty} e^{\cos(\pi - n)} + 5 \leq e^1 + 5 < +\infty$.

2. We have: $f(x,y,z) = e^{-\cos(x)} + y^2 + 4y + z^2 + 5$ and $f \in C^\infty(\mathbb{R}^3)$, thus
$$\nabla f(x,y,z) = \begin{pmatrix} (\sin x)\, e^{-\cos x} \\ 2y+4 \\ 2z \end{pmatrix}, \text{ then } \nabla f(x,y,z) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longleftrightarrow \begin{cases} x = k\pi, \forall k \in \mathbb{Z} \\ y = -2 \\ z = 0 \end{cases}.$$
Thus, the critical points are $\left\{(k\pi, -2, 0)^t \in \mathbb{R}^3 \text{ such that } k \in \mathbb{Z}\right\}$.

3. Since $f \in C^\infty(\mathbb{R}^3)$ and $\nabla^2 f(x,y,z) = H_f(x,y,z)$, we have
$$H_f(x,y,z) = \begin{pmatrix} e^{-\cos x}\sin^2 x + (\cos x)\, e^{-\cos x} & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \text{ then}$$
$$\nabla^2 f(k\pi, -2, 0) = \begin{cases} \begin{pmatrix} e^{-1} & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \text{ if } k = 2m \Longrightarrow \text{ positive definite} \\ \begin{pmatrix} -e & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \text{ if } k = 2m+1 \Longrightarrow \text{ not positive semi-definite.} \end{cases}$$
Thus, the points $(2m\pi, -2, 0)^t, m \in \mathbb{Z}$ are local minima of f on $\mathbb{R}^3$.

4. We have $f(x,y,z) = e^{\cos(\pi - x)} + y^2 + 4y + z^2 + 5 = e^{\cos(\pi - x)} + (y+2)^2 + z^2 + 1$.

   Since $-1 \leq \cos(\pi - x) \leq 1$, we have $f(x,y,z) \geq e^{-1} + 1, \forall (x,y,z)^t \in \mathbb{R}^3$.

   Thus, f is bounded below by $e^{-1} + 1$.

5. From 3) and 4) we have
$$\left.\begin{array}{r} f(x,y,z) \geq e^{-1} + 1, \forall (x,y,z)^t \in \mathbb{R}^3. \\ f(2m\pi, -2, 0) = e^{-1} + 1 \end{array}\right\} \Longrightarrow (2m\pi, -2, 0)^t, m \in \mathbb{Z} \text{ are global min-}$$
imizers of f on $\mathbb{R}^3$.

6. The function f is not convex on $\mathbb{R}^3$, because the points $((2m+1)\pi, -2, 0)^t, m \in \mathbb{Z}$ are critical points but not global minimizers.

7. The function f is not concave on $\mathbb{R}^3$, because the points $(2m\pi, -2, 0)^t, m \in \mathbb{Z}$ are critical points but not global maximizers.

Chapter 2

Optimization problem and optimality conditions

Contents

2.1 Introduction

In this chapter, we will first define the general formulation of optimization problems. We then present optimality conditions for unconstrained nonlinear optimization problems, beginning with first-order conditions followed by second-order necessary and sufficient conditions.

Definition 2.1.1 *An unconstrained optimization problem can be expressed in the following form:*

$$(P)\begin{cases}\min f(x);\\ x\in\mathbb{R}^n.\end{cases}\qquad\Longleftrightarrow\qquad (P)\begin{cases}\textit{Find } x^*\in\mathbb{R}^n \textit{ such that}\\ f(x^*)\leq f(x),\ \forall x\in\mathbb{R}^n.\end{cases}$$

where $x\in\mathbb{R}^n$ is called the optimization variable, and $f:\mathbb{R}^n\to\mathbb{R}$ is called the cost function or objective function or criterion.

Example. The problem of solving the equation $f(x)=0$ for $x\in\mathbb{R}^n$ is equivalent to the following optimization problem:

$$(P)\begin{cases}\text{Find } x^*\in\mathbb{R}^n;\\ |f(x^*)|=\min_{x\in\mathbb{R}^n}|f(x)|.\end{cases}$$

Definition 2.1.2 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a scalar function defined on $\mathbb{R}^n$ and let $x^* \in \mathbb{R}^n$.*

1. *We say that x^* is a local minimum point of f on $\mathbb{R}^n$ if there exists $r > 0$ such that:*

$$f(x^*) \leq f(x), \quad \forall x \in B(x^*, r).$$

2. *We say that x^* is a strict local minimum point of f on $\mathbb{R}^n$ if there exists $r > 0$ such that:*

$$f(x^*) < f(x), \quad \forall x \in B(x^*, r) \setminus \{x^*\}.$$

3. *We say that x^* is a global minimum (resp. strict global minimum) of f on $\mathbb{R}^n$ if:*

$$f(x^*) \leq f(x), \quad \forall x \in \mathbb{R}^n \quad (resp.\ f(x^*) < f(x), \forall x \in \mathbb{R}^n \setminus \{x^*\}).$$

4. *We can define local maximum (resp. strict local maximum), global maximum (resp. strict global maximum) points by reversing the inequalities above.*

5. *If $x^* \in \mathbb{R}^n$ is a local or global minimum (resp. maximum) point, the value $f(x^*)$ is called the minimal (resp. maximal) value.*

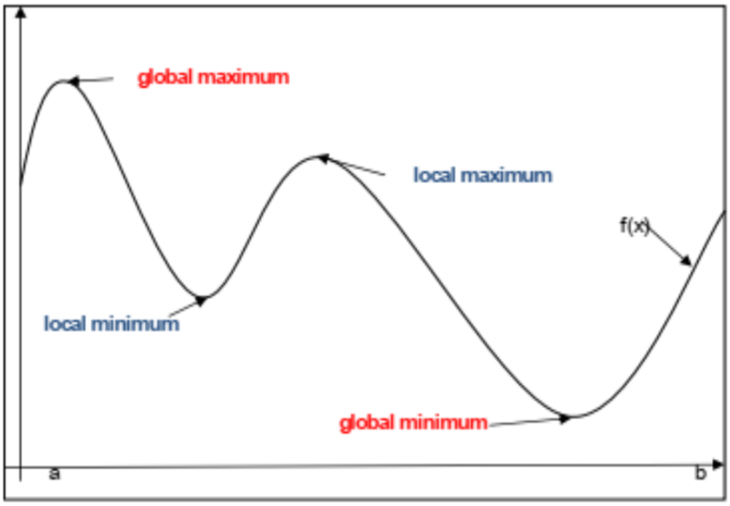

Figure 2.1: Local and Global minima and maxima of a function

Remark 2.1.1

1. *If x^* is a global minimum, then x^* is also a local minimum.*

2. *For a function $f : \mathbb{R}^n \to \mathbb{R}$, the infimum $\inf_{\mathbb{R}^n} f(x)$ may not equal the minimum $\min_{\mathbb{R}^n} f(x)$.*

Example. Consider the function

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} x^2 + 1 & \text{if } x \neq 0, \\ 2021 & \text{otherwise.} \end{cases}$$

For this function we have:

$$\inf_{x \in \mathbb{R}} f(x) = 1 \quad \text{but} \quad \min_{x \in \mathbb{R}} f(x) \text{ does not exist}$$

(since there is no $x^* \in \mathbb{R}$ such that $f(x^*) = 1$).

Lemma 2.1.1 *Let*

$$f : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto f(x)$$

1. *The optimization problems* $\min_{x \in \mathbb{R}^n} f(x)$ *and* $\max_{x \in \mathbb{R}^n}(-f(x))$ *are equivalent.*
2. $\min_{x \in \mathbb{R}^n} f(x) = -\left(\max_{x \in \mathbb{R}^n}(-f(x))\right)$.

2.2 Existence and Uniqueness Results

Theorem 2.2.1 *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be a continuous and coercive function on* $\mathbb{R}^n$*. Then there exists at least one point* $x^* \in \mathbb{R}^n$ *such that:*

$$f(x^*) \leq f(x), \quad \forall x \in \mathbb{R}^n.$$

The proof of this theorem relies on the following Weierstrass theorem:

Theorem 2.2.2 *(Weierstrass). Let* f *be a continuous function on a compact set* $K \subset \mathbb{R}^n$*. Then there exist points* $\underline{x}, \overline{x} \in K$ *such that:*

$$f(\underline{x}) \leq f(x) \leq f(\overline{x}) \quad \forall x \in K.$$

Proof. Let $a \in \mathbb{R}^n$ and define the subset:

$$K = \{x \in \mathbb{R}^n : f(x) \leq f(a)\}.$$

1. K is closed since $K = f^{-1}\big((-\infty, f(a)]\big)$ and f is continuous from $\mathbb{R}^n$ to $\mathbb{R}$.
2. K is bounded in $\mathbb{R}^n$. Indeed, if K were unbounded, there would exist a sequence $(x_m)_m \subset K$ such that:
$$\|x_m\| \to +\infty \quad \text{as } m \to +\infty.$$

The coercivity of f on $\mathbb{R}^n$ implies that $\lim_{m \to +\infty} f(x_m) = +\infty$, which contradicts

$$\lim_{m \to +\infty} f(x_m) \leq f(a).$$

From 1. and 2., the set K is compact in $\mathbb{R}^n$.

By the Weierstrass theorem, there exists $x^* \in K$ such that

$$f(x^*) \leq f(x) \text{ for all } x \in K.$$

Moreover,

$$f(x^*) \leq f(a) < f(x) \text{ for all } x \in \mathbb{R}^n \setminus K.$$

Therefore,

$$f(x^*) \leq f(x) \text{ for all } x \in \mathbb{R}^n.$$

.

Note that uniqueness is not guaranteed in this theorem.

Example. The function $f(x) = |x^2 - 1|$ is continuous and coercive on $\mathbb{R}$. Since $f(x) = |x^2 - 1| \geq 0 = f(1) = f(-1)$, there exist two global minima of f on $\mathbb{R}$.

Theorem 2.2.3 *(Uniqueness). Let $f : \mathbb{R}^n \to \mathbb{R}$ be strictly convex on $\mathbb{R}^n$. Then the optimization problem:*

$$(P) \begin{cases} \min f(x); \\ x \in \mathbb{R}^n \end{cases}$$

admits at most one solution, i.e., there exists at most one global minimum of f on $\mathbb{R}^n$.

Proof. Suppose there exist two distinct minimizers $\overline{x}_1 \neq \overline{x}_2$ of f on $\mathbb{R}^n$ such that:

$$f(\overline{x}_1) \leq f(x) \quad \forall x \in \mathbb{R}^n \quad \text{and} \quad f(\overline{x}_2) \leq f(x) \quad \forall x \in \mathbb{R}^n.$$

Let

$$y = \tfrac{1}{2}\overline{x}_1 + \tfrac{1}{2}\overline{x}_2.$$

Since $f(\overline{x}_1) = f(\overline{x}_2)$ (because $f(\overline{x}_1) \leq f(\overline{x}_2)$ and $f(\overline{x}_2) \leq f(\overline{x}_1)$), the strict convexity of f implies:

$$f(y) = f\left(\tfrac{1}{2}\overline{x}_1 + \tfrac{1}{2}\overline{x}_2\right) < \tfrac{1}{2}f(\overline{x}_1) + \tfrac{1}{2}f(\overline{x}_2) = f(\overline{x}_1).$$

This contradicts the assumption that $\overline{x}_1$ is a global minimizer of f on $\mathbb{R}^n$.

Theorem 2.2.4 *(Existence and Uniqueness). Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable on $\mathbb{R}^n$, and suppose there exists $\alpha > 0$ such that*

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \alpha \|x - y\|^2 \quad \forall x, y \in \mathbb{R}^n.$$

Then:

1. *f is strictly convex on $\mathbb{R}^n$.*
2. *f is coercive on $\mathbb{R}^n$.*

3. *The optimization problem:*

$$(P)\begin{cases}\min f(x);\\ x\in\mathbb{R}^n,\end{cases}$$

admits a unique solution, i.e., there exists a unique global minimum of f on $\mathbb{R}^n$.

Example. Let $f:\mathbb{R}^n\to\mathbb{R}$ be a quadratic form defined by:

$$f(x)=\frac{1}{2}x^\top Ax-b^\top x+c,$$

where A is a positive definite matrix.

1. Show that f satisfies:

$$\langle\nabla f(x)-\nabla f(y),x-y\rangle\geq\alpha\|x-y\|^2\quad\forall x,y\in\mathbb{R}^n.$$

2. Deduce that f has a unique global minimum $\bar{x}\in\mathbb{R}^n$.

3. Compute $\bar{x}$.

2.3 First-Order Necessary Condition for Optimality

Theorem 2.3.1 *(First-Order Necessary Condition). Let $f:\mathbb{R}^n\to\mathbb{R}$ be differentiable at a point $x^*\in\mathbb{R}^n$. Then:*

$$x^* \text{ is a local or global minimizer of } f \text{ on } \mathbb{R}^n \implies \nabla f(x^*)=0_{\mathbb{R}^n}.$$

Proof. Let $h\in\mathbb{R}^n\setminus\{0_{\mathbb{R}^n}\}$ and $t\in\mathbb{R}$. We have:

$$\lim_{t\to0}\frac{f(x^*+th)-f(x^*)}{t}=\langle\nabla f(x^*),h\rangle.$$

For sufficiently small $|t|$, we obtain:

$$0\leq\lim_{\substack{t\to0\\t>0}}\frac{f(x^*+th)-f(x^*)}{t}=\langle\nabla f(x^*),h\rangle,\tag{2.1}$$

$$0\geq\lim_{\substack{t\to0\\t<0}}\frac{f(x^*+th)-f(x^*)}{t}=\langle\nabla f(x^*),h\rangle.\tag{2.2}$$

From (2.1) and (2.2), we conclude $\langle\nabla f(x^*),h\rangle=0$ for all $h\in\mathbb{R}^n$, hence $\nabla f(x^*)=0$.

Remark 2.3.1

1. *The condition $\nabla f(x^*)=0$ is necessary but not sufficient. For example, consider $f(x)=x^3$ defined on $\mathbb{R}$. We have $\nabla f(x)=f'(x)=3x^2$ and $f'(0)=0$, but $x=0$ is not a local minimizer of f on $\mathbb{R}$.*

2. *This theorem doesn't apply when f is not differentiable. For instance:*

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto |x+1|$$

Here $\bar{x} = -1$ is a global minimizer of f on $\mathbb{R}$, but $\nabla f(-1) = f'(-1)$ does not exist.

Theorem 2.3.2 *(Necessary and Sufficient Condition). Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex and differentiable function in a neighborhood of x^*. Then:*

$$x^* \text{ is a global minimizer of } f \text{ on } \mathbb{R}^n \iff \nabla f(x^*) = 0.$$

Proof.
($\Rightarrow$) This follows directly from the previous theorem.
($\Leftarrow$) Assume $\nabla f(x^*) = 0$. Since f is convex, we have:

$$f(x) \geq f(x^*) + \langle \nabla f(x^*), x - x^* \rangle \quad \forall x \in \mathbb{R}^n \implies f(x) \geq f(x^*) \quad \forall x \in \mathbb{R}^n.$$

Therefore, x^* is a global minimizer of f on $\mathbb{R}^n$.

Example.

1. Consider $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = e^{x^2}$. We can show that f is convex on $\mathbb{R}$ (exercise). The gradient is:

$$\nabla f(x) = f'(x) = 2xe^{x^2} = 0 \iff x = 0.$$

Since:

$$\begin{cases} f'(0) = 0 \\ f \text{ is convex} \end{cases}$$

it follows that $\bar{x} = 0$ is a global minimizer of f on $\mathbb{R}$.

2. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by:

$$f(x, y) = (x - y)^2.$$

We have:

$$\nabla f(x, y) = \begin{pmatrix} 2(x-y) \\ 2(y-x) \end{pmatrix}, \quad \nabla^2 f(x, y) = H_f(x, y) = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}.$$

This shows f is a quadratic form since $H_f(x, y)$ doesn't depend on x and y.
For all $(x, y)^\top \in \mathbb{R}^2$, we have:

$$\text{Spec}(H_f(x, y)) = \{0, 4\} \subset [0, +\infty) \implies \begin{cases} H_f(x, y) \text{ is positive semidefinite} \\ f \text{ is convex on } \mathbb{R}^2 \end{cases}$$

The gradient condition:

$$\nabla f(x, y) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{cases} 2(x-y) = 0 \\ 2(y-x) = 0 \end{cases} \iff x = y$$

shows that for all $\alpha \in \mathbb{R}$, the point $\overline{X}_\alpha = (\alpha, \alpha)^\top$ is a global minimizer of f on $\mathbb{R}^2$.

Definition 2.3.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function on $\mathbb{R}^n$. A point $x^* \in \mathbb{R}^n$ satisfying $\nabla f(x^*) = 0$ is called a* ***stationary point*** *or* ***critical point****, and the equation $\nabla f(x) = 0_{\mathbb{R}^n}$ is called the* ***Euler equation****.*

Remark 2.3.2 *For a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, if $x^* \in \mathbb{R}^n$ is a local or global minimizer or maximizer, then x^* must be a critical point.*

2.4 Second-Order Necessary Conditions for Optimality

Proposition 2.4.1 *Proposition (Second-Order Necessary Condition). Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function in a neighborhood of $x^* \in \mathbb{R}^n$. If x^* is a local minimizer of f on $\mathbb{R}^n$, then:*

1. $\nabla f(x^*) = 0$
2. $\nabla^2 f(x^*)$ *is positive semidefinite on $\mathbb{R}^n$.*

Proof.

1. This follows from the first-order necessary condition (Theorem 2.5).
2. We proceed by contradiction. Suppose $\nabla^2 f(x^*)$ is not positive semidefinite. Then there exists a nonzero vector $P \in \mathbb{R}^n$ such that

   $$P^\top \nabla^2 f(x^*) P < 0.$$

   Using the second-order Taylor expansion:

   $$f(x^* + P) = f(x^*) + \langle \nabla f(x^*), P \rangle + \frac{1}{2} \langle \nabla^2 f(x^* + tP)P, P \rangle, \quad \text{for some } 0 < t < 1,$$

   which implies

   $$f(x^* + P) - f(x^*) = \frac{1}{2} \langle \nabla^2 f(x^* + tP)P, P \rangle.$$

   Since $\nabla^2 f(x)$ is continuous in a neighborhood of x^*, we have $\langle \nabla^2 f(x^* + tP)P, P \rangle < 0$ for sufficiently small t. Therefore:

   $$\begin{aligned} f(x^* + P) - f(x^*) < 0 &\implies f(x^* + P) < f(x^*) \\ &\implies x^* \text{ is not a local minimizer.} \end{aligned}$$

Remark 2.4.1

1. *The conditions 1 and 2 are necessary but not sufficient. Consider the example:*

   $$f : \mathbb{R} \to \mathbb{R}, \quad t \mapsto t^3$$

   We have:

   $$f'(t) = 3t^2, \quad f''(t) = 6t \quad \textit{with} \quad f'(0) = 0 \quad \textit{and} \quad f''(0) = 0$$

   Thus $f''(0)$ is positive semidefinite ($f''(0) \geq 0$), but $t^* = 0$ is neither a local minimizer nor a local maximizer on $\mathbb{R}$.

2. *The second-order necessary condition for x^* to be a local (or global) maximizer is as follows. Under the same assumptions on f:*

$$x^* \text{ is a local maximizer } \Longrightarrow \begin{cases} (i)\ \nabla f(x^*) = 0, \\ (ii)\ \nabla^2 f(x^*) \text{ is negative semidefinite.} \end{cases}$$

2.5 Sufficient second-order condition for Optimality

Proposition 2.5.1 *(Second-Order Sufficient Condition). Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a $\mathcal{C}^2$ function in a neighborhood of $x^* \in \mathbb{R}^n$ satisfying:*

1. $\nabla f(x^*) = 0$,
2. $\nabla^2 f(x^*)$ *is positive definite on* $\mathbb{R}^n$.

   Then x^* is a strict local minimizer of f on $\mathbb{R}^n$.

Remark 2.5.1 *Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a $\mathcal{C}^2$ function in a neighborhood of $x^* \in \mathbb{R}^n$ satisfying:*

1. $\nabla f(x^*) = 0$,
2. $\nabla^2 f(x^*)$ *is negative definite on* $\mathbb{R}^n$.

   Then x^* is a strict local maximizer of f on $\mathbb{R}^n$.

Example. Consider the function:

$$f\colon \mathbb{R}^3 \to \mathbb{R}, \quad X = (x, y, z)^\top \mapsto f(X) = x^4 - 2x^2 z + 2y^2 + 2yz + 2z^2 - 2y + 15.$$

1. Compute $\nabla f(x, y, z)$ and $\nabla^2 f(x, y, z)$ for all $(x, y, z)^\top \in \mathbb{R}^3$.
2. Determine the critical points of f.
3. Show that $f(x, y, z) \geq 14$ for all $(x, y, z) \in \mathbb{R}^3$.
4. Deduce the minimum of f on $\mathbb{R}^3$.

2.6 Exercises

Exercise 2.1. Let $f\colon \mathbb{R}^4 \to \mathbb{R}$ be defined by:

$$f(x) = f(x_1, x_2, x_3, x_4) = \sum_{i=1}^{4} 2x_i^2 + \left(\sum_{i=1}^{4} x_i \right)^2.$$

1. Compute $\nabla f(x)$ and $H(x) = \nabla^2 f(x)$.
2. Deduce that f is a quadratic form on $\mathbb{R}^4$.
3. Show that the minimization problem $\min_{x\in\mathbb{R}^4} f(x)$ admits a unique solution $\overline{x} \in \mathbb{R}^4$ and compute $\overline{x}$.

Exercise 2.2. Consider the function g defined by:

$$g(x,y) = ax^2y + bxy + 2xy^2 + c,$$

where $a, b, c \in \mathbb{R}$.

1. Find all critical points of g.
2. Determine the values of a, b, c for which g has a local minimum at the point $\left(\frac{2}{3}, \frac{1}{3}\right)^\top$ with $g\left(\frac{2}{3}, \frac{1}{3}\right) = -\frac{1}{9}$.

Exercise 2.3. Consider the function $f \colon \mathbb{R}^2 \to \mathbb{R}$ defined by:

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

1. Show that f is coercive (i.e., $f(x,y) \to +\infty$ as $\|(x,y)\| \to +\infty$).
2. Show that f has a global minimizer $(\overline{x}, \overline{y})^\top$ on $\mathbb{R}^2$ and compute it.
3. Can a coercive function have a global maximum?

Exercise 2.4. Let $f \colon \mathbb{R}^2 \to \mathbb{R}$ be defined by:

$$f(x,y) = \exp\left[(x-1)^2\right] + (y^2 - 4)^2.$$

1. Determine all critical points of f.
2. Find all local minima of f.
3. Show that f is coercive.
4. Deduce the global minima of f on $\mathbb{R}^2$.

Exercise 2.5. Consider the following sets:

1. $A = \left\{(x,y)^\top \in \mathbb{R}^2 \mid (x-1)(y+1) \leq 0\right\}$.
2. $B = \left\{(x,y)^\top \in \mathbb{R}^2 \mid y \leq 0 \text{ and } y \geq x^3\right\}$.

Sketch each set graphically and determine whether it is convex or not.

Exercise 2.6. Let $\{f_i\}_{i=1}^m$ be a family of convex functions defined on a convex set $U \subseteq \mathbb{R}^n$. Prove that:

1. The pointwise supremum $\sup_{1 \leq i \leq m} f_i$ is convex on U, where

$$\left(\sup_{1 \leq i \leq m} f_i\right)(x) = \sup_{1 \leq i \leq m} \left(f_i(x)\right).$$

2. For all $\lambda_i \geq 0$, the weighted sum $\sum_{i=1}^{m} \lambda_i f_i$ is convex on U.

Exercise 2.7. Analyze the convexity (convex, strictly convex, concave, and strictly concave) of the following functions:

1. The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by:

$$f(x) = e^{(x^2+2021)^3}$$

2. The function $f\colon \mathbb{R}^2 \to \mathbb{R}$ defined by:

$$f(x, y) = -|x| + y$$

Exercise 2.8. Prove that:

$$\ln\left(\frac{x+y}{2}\right) \geq \sqrt{\ln x \ln y}, \quad \forall x, y \in \,]1, +\infty[\,.$$

Exercise 2.9. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be defined by:

$$f(x) = \frac{1}{2}\langle x, x\rangle + \langle a, x\rangle^2$$

where $a \in \mathbb{R}^n$ $(a \neq 0)$.

1. Compute $\nabla f(x)$ and $H(x) = \nabla^2 f(x)$.
2. Deduce that f is a quadratic form on $\mathbb{R}^n$.
3. Show that f is strictly convex and coercive.

Chapter 3

Unconstrained optimization algorithms

Contents

3.1 Introduction

In this chapter, we present several algorithms for numerically approximating the solution(s) to an unconstrained nonlinear optimization problem:

$$(P) \begin{cases} \min f(x) \\ x \in \mathbb{R}^n \end{cases}$$

We introduce the most fundamental classical methods for solving such problems.

Definition 3.1.1 *An algorithm is defined by a mapping $\mathcal{A}\colon \mathbb{R}^n \to \mathbb{R}^n$ that generates a sequence of elements in $\mathbb{R}^n$ through the following iterative scheme:*

$$\text{(Alg.)} \quad \begin{cases} x^{(0)} \in \mathbb{R}^n & \text{(given initial point)} \\ x^{(k+1)} = \mathcal{A}(x^{(k)}) & \text{for } k \geq 0 \quad \text{(iteration } k\text{)} \end{cases}$$

Implementing an algorithm amounts to constructing such a sequence $(x^{(k)})_k$ in $\mathbb{R}^n$, and analyzing the algorithm's convergence means studying the convergence of this sequence.

Definition 3.1.2 *An algorithm $\mathcal{A}$ is said to converge if and only if the sequence generated by the algorithm (Alg.) converges to a limit point $x^* \in \mathbb{R}^n$.*

Definition 3.1.3 *Let $(x^{(k)})_k$ be a sequence with limit $x^* \in \mathbb{R}^n$ defined by a convergent algorithm $\mathcal{A}$. We say the convergence of $\mathcal{A}$ is:*

• ***Linear*** *if the error $e_k = \|x^{(k)} - x^*\|$ decreases linearly:*

$$\exists C \in [0,1), \exists k_0, \forall k \geq k_0 : e_{k+1} \leq C e_k.$$

• ***Superlinear*** *if the error e_k decreases as:*

$$e_{k+1} \leq \alpha_k e_k,$$

*where α_k is a positive sequence converging to 0. If α_k is a geometric sequence, the convergence is called **geometric**.*

• ***Of order p*** *if the error e_k decreases as:*

$$\exists C \geq 0, \exists k_0, \forall k \geq k_0 : e_{k+1} \leq C e_k^p.$$

*For $p = 2$, the convergence is called **quadratic**.*

• *Finally, the convergence is called **local** if it only occurs for initial points $x^{(0)}$ in some neighborhood of x^*. Otherwise, the convergence is **global**.*

3.2 Descent methods

Descent methods generally take the following form:

$$\begin{cases} X^0 \in \mathbb{R}^n & \text{given}, \\ X^{(k+1)} = X^{(k)} + \rho_k\, d^{(k)}, \end{cases}$$

where $d^{(k)} \in \mathbb{R}^n \setminus \{0\}$ is chosen such that $f\big(X^{(k+1)}\big) < f\big(X^{(k)}\big)$. Here, $d^{(k)}$ is called the descent direction, and $\rho_k > 0$ ($\rho_k \in \mathbb{R}_+^*$) is called the step size (or descent step).

Definition 3.2.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. A vector $d \in \mathbb{R}^n \setminus \{0\}$ is called a descent direction at the point x if there exists $\varepsilon > 0$ such that*

$$f(x + td) < f(x), \quad \forall t \in [0, \varepsilon]\,.$$

Lemma 3.2.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a $C^1(\mathbb{R}^n)$ function. Then, for every $x \in \mathbb{R}^n$, the vector $d = -\nabla f(x)$ is a descent direction at the point x.*

Proof. Since f is of class $C^1(\mathbb{R}^n)$, we use the first-order Taylor expansion of f around the point $x \in \mathbb{R}^n$:

$$f(x + td) = f(x) + t\langle d, \nabla f(x)\rangle + O(td).$$

For $t \in [0, \varepsilon]$ and $d = -\nabla f(x) \neq 0$, we obtain:

$$f(x + td) = f(x) - t\|\nabla f(x)\|^2 + O(td) \implies f(x + td) < f(x).$$

3.3 Gradient method

Methods where the descent direction at each iteration k is $d^k = -\nabla f(X^{(k)})$ are called gradient methods. The general gradient descent algorithm is given by:

1. **Initialization** $(k = 0)$

$$X^{(0)} \in \mathbb{R}^n,\ \rho_0 > 0 \text{ and } \varepsilon > 0$$

2. **Iteration** k

$$X^{(k+1)} = X^{(k)} - \rho_k \nabla f\left(X^{(k)}\right)$$

3. **Stopping criterion**

$$\|X^{(k)} - X^{(k+1)}\| \leqslant \varepsilon \text{ or } \|\nabla f\left(X^{(k+1)}\right)\| \leqslant \varepsilon \text{ then stop}$$

   Otherwise, set $k = k + 1$ and return to step 2.

Remark 3.3.1 *The step size ρ_k can be chosen as either fixed or variable.*

3.3.1 Gradient method with fixed step size

Let $f \in C(\mathbb{R}^n, \mathbb{R})$ and ρ be a strictly positive real number. The fixed-step gradient method (ρ fixed) is a descent method defined by the following algorithm:

• **Initialization** $(k = 0)$

$X^{(0)} \in \mathbb{R}^n$ given, $\rho > 0$ fixed and given, $\epsilon > 0$ (the precision is given).

• **Iteration** k

- Compute $\nabla f\left(X^{(k)}\right)$
- Compute $X^{(k+1)} = X^{(k)} - \rho \nabla f\left(X^{(k)}\right)$

• **Stopping criterion**

If $\|X^{(k+1)} - X^{(k)}\| < \varepsilon$, stop. Otherwise, set $k = k + 1$ and return to step 2.

Theorem 3.3.1 ***Theorem.*** *(Convergence) Let $f \in C^1(\mathbb{R}^n, \mathbb{R})$. Assume there exist two strictly positive real numbers α and M such that:*

$$\begin{aligned} i)\ \forall x, y &\in \mathbb{R}^n: \quad \langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \alpha \|x - y\|^2, \\ ii)\ \forall x, y &\in \mathbb{R}^n: \quad \|\nabla f(x) - \nabla f(y)\| \leq M \|x - y\|. \end{aligned}$$

Then:

1. *f is coercive on $\mathbb{R}^n$.*
2. *f is strictly convex on $\mathbb{R}^n$.*
3. *The function f has a unique minimum x^* such that $f(x^*) \leq f(x) \quad \forall x \in \mathbb{R}^n$.*
4. *If $0 < \rho < \dfrac{2\alpha}{M^2}$, the sequence generated by the fixed-step gradient method with step size ρ converges to the unique solution x^*, regardless of the starting point $x^{(0)}$.*

Example. Consider the problem:

$$(P) \begin{cases} \min f(x, y) = 2(x^2 + y^2) - 3xy \\ (x, y) \in \mathbb{R}^2 \end{cases}$$

1. Perform 4 iterations of the fixed-step gradient method ($\rho = \dfrac{1}{25}$) starting from the initial point $X^{(0)} = (1, 1)^t$.
2. Compute the exact solution $\bar{X}$ of (P).
3. Determine the error $\epsilon = e_4 = \|X^{(4)} - \bar{X}\|$.

Solution. Recall that the fixed-step gradient algorithm sequence is:

$$\begin{cases} X^{(0)} \quad \text{given} \\ X^{(k+1)} = X^{(k)} - \rho \nabla f(X^{(k)}) \end{cases}$$

1. We have $f \in C^1(\mathbb{R}^2)$. For all $X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, $\nabla f(X) = \nabla f(x, y) = \begin{pmatrix} 4x - 3y \\ -3x + 4y \end{pmatrix}$. Therefore, the sequence becomes:

$$\begin{cases} X^{(0)} = (1, 1)^t \quad \text{given} \\ X^{(k+1)} = \begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \end{pmatrix} = \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix} - \dfrac{1}{25} \begin{pmatrix} 4x^{(k)} - 3y^{(k)} \\ -3x^{(k)} + 4y^{(k)} \end{pmatrix} = \begin{pmatrix} \frac{21}{25}x^k + \frac{3}{25}y^k \\ \frac{3}{25}x^k + \frac{21}{25}y^k \end{pmatrix} \end{cases}$$

For:

- $k=0 \quad \Rightarrow \quad X^{(0)}=\begin{pmatrix}1\\1\end{pmatrix}, \quad k=1 \quad \Rightarrow \quad X^{(1)}=\begin{pmatrix}\dfrac{24}{25}\\ \dfrac{24}{25}\end{pmatrix}$

- $k=2 \quad \Rightarrow \quad X^{(2)}=\begin{pmatrix}\left(\dfrac{24}{25}\right)^2\\ \left(\dfrac{24}{25}\right)^2\end{pmatrix}, \quad k=3 \quad \Rightarrow \quad X^{(3)}=\begin{pmatrix}\left(\dfrac{24}{25}\right)^3\\ \left(\dfrac{24}{25}\right)^3\end{pmatrix}=$ $\left(\dfrac{24}{25}\right)^3\begin{pmatrix}1\\1\end{pmatrix}$

- $k=4 \quad \Rightarrow \quad X^{(4)}=\left(\dfrac{24}{25}\right)^4\begin{pmatrix}1\\1\end{pmatrix}$

2. We have $f \in C^2(\mathbb{R}^2)$ and the Hessian matrix $\nabla^2 f(x,y)=\begin{pmatrix}4 & -3\\ -3 & 4\end{pmatrix}=A$, so f is a quadratic form.

   Since $\Delta_1=\det(4)=4>0$ and $\Delta_2=\det\begin{pmatrix}4 & -3\\ -3 & 4\end{pmatrix}=7>0$, then f is strictly convex and coercive on $\mathbb{R}^2$.

   Therefore, f admits a unique strict global minimum $\overline{X}$ on $\mathbb{R}^2$ such that $\nabla f(\overline{X})=0_{\mathbb{R}^2}$.

   $\nabla f(\overline{X})=0_{\mathbb{R}^2}\Rightarrow A\overline{X}=0_{\mathbb{R}^2}\Rightarrow \overline{X}=A^{-1}0_{\mathbb{R}^2}=0_{\mathbb{R}^2}$.

3. The error $e_4=\|X^{(4)}-\overline{X}\|=\|\left(\dfrac{24}{25}\right)^4\begin{pmatrix}1\\1\end{pmatrix}-\begin{pmatrix}0\\0\end{pmatrix}\|=\left(\dfrac{24}{25}\right)^4\|\begin{pmatrix}1\\1\end{pmatrix}\|=\left(\dfrac{24}{25}\right)^4\sqrt{2}\approx 1.20.$

Remark 3.3.2

1. *Since* $X^{(k)}=\left(\dfrac{24}{25}\right)^k\begin{pmatrix}1\\1\end{pmatrix}$*, the error* $e_k=\|X^{(k)}-\overline{X}\|=\left(\dfrac{24}{25}\right)^k\sqrt{2}$ *satisfies* $\lim_{k\to\infty} e_k=0$*. Therefore, the fixed-step gradient method sequence converges to the unique solution* $\overline{X}=(0,0)^t$*, but the convergence is very slow.*

2. *The function can be expressed as* $f(X)=f(x,y)=\dfrac{1}{2}\begin{pmatrix}x & y\end{pmatrix}\begin{pmatrix}4 & -3\\ -3 & 4\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\dfrac{1}{2}X^tAX.$

   The gradient satisfies:

$$\begin{aligned}\langle\nabla f(X)-\nabla f(X');X-X'\rangle &\geq 1\|X-X'\|^2,\\ \|\nabla f(X)-\nabla f(X')\| &\leq 7\|X-X'\|.\end{aligned}$$

In this example, the step size satisfies $0 < \rho = \dfrac{1}{25} < \dfrac{2(1)}{(7)^2}$.

3.3.2 Gradient method with optimal step size

The general idea of this method is to compute at each iteration k the step size ρ_k (the displacement step) as the minimum over $\mathbb{R}_+$ of the function $\phi_k(t) = f(x^{(k)} - t\nabla f(x^{(k)}))$, that is,

$$\rho_k = \arg\min_{t\geq 0}\ \phi_k(t).$$

Thus, the algorithm of the gradient method with optimal step size takes the following form:

1. **Initialization** $(k = 0)$
   $X^{(0)}$ given and $\varepsilon > 0$
2. **Iteration** k

   while $\|\nabla f(X^{(k)})\| > \varepsilon$ do

   (a) Compute $\nabla f(X^{(k)}) = -d_k$.

   (b) Compute $\phi_k\,(t) = f\left(X^{(k)} - t\nabla f(X^{(k)})\right)$.

   (c) Compute ρ_k such that $f\left(X^{(k)} + \rho_k d^{(k)}\right) \leq f\left(X^{(k)} + td^{(k)}\right), \forall \rho > 0$, that is

   $$\rho_k = \arg\min_{\mathbb{R}_+}\ \phi_k(t)$$

   (d)
   $$X^{(k+1)} = X^{(k)} + \rho_k d^{(k)}$$
3. **Stopping criterion**
   $$\text{If } \|\nabla f\left(X^{(k)}\right)\| \leq \varepsilon \quad \text{stop.}$$

Example. Consider the problem:

$$(P)\begin{cases}\min\ 4(2x_1^2 + x_2^2)\\ \begin{pmatrix} x_1\\ x_2\end{pmatrix} \in \mathbb{R}^2\end{cases}$$

Starting from $X^{(0)} = \left(\dfrac{1}{10}, \dfrac{1}{10}\right)^t$, compute an approximate solution to (P) with precision $\varepsilon = 5\times 10^{-1}$.

Solution. The sequence of gradient method with optimal step size is given by:

$$\begin{cases} X^{(0)} = \frac{1}{10}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ et } \varepsilon = 10^{-2} > 0, \\ d_k = -\nabla f(X^{(k)}) \\ \rho_k = \rho_{opt} = \arg\min\limits_{\rho \geqslant 0} f\left(X^{(k)} + \rho d_k\right) \\ X^{(k+1)} = X^{(k)} + \rho_k d_k \end{cases}$$

$X^{(0)} = \frac{1}{10}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \Longrightarrow d_0 = -\nabla f\left(X^{(0)}\right) = -\nabla f\left(\frac{1}{10}, \frac{1}{10}\right)$, où $\nabla f(x_1, x_2) = \begin{pmatrix} 16x_1 \\ 8x_2 \end{pmatrix}$.
Then $d_0 = \frac{-1}{10}\begin{pmatrix} 16 \\ 8 \end{pmatrix}$.

$$\begin{aligned} \rho_0 = \arg\min_{\rho \geqslant 0} f\left(X^{(0)} + \rho d_0\right) &= \arg\min_{\rho \geqslant 0} f\left(\frac{1}{10}(1,1) - \rho\frac{1}{10}\begin{pmatrix} 16 \\ 8 \end{pmatrix}\right) \\ &= \arg\min_{\rho \geqslant 0} f\left(\frac{1-16\rho}{10}, \frac{1-8\rho}{10}\right) \\ &= \arg\min_{\rho \geqslant 0}\left[\tfrac{4}{100}\left(2(1-16\rho)^2 + (1-8\rho)^2\right)\right] \\ &= \arg\min_{\rho \geqslant 0} \tfrac{4}{100}\left[576\rho^2 + -80\rho + 3\right]. \end{aligned}$$

Therefore ρ_0 satisfies $2 \times 576\rho_0 - 80 = 0 \Longrightarrow \rho_0 = \frac{5}{72} > 0$.

$$X^{(1)} = X^{(0)} + \rho_0 d_0 = \frac{1}{10}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{5}{720}\begin{pmatrix} 16 \\ 8 \end{pmatrix} = \begin{pmatrix} -\frac{1}{90} \\ \frac{2}{45} \end{pmatrix}$$

$$\begin{aligned} X^{(1)} = \frac{1}{10}\begin{pmatrix} -\frac{1}{9} \\ \frac{4}{9} \end{pmatrix} &\Longrightarrow d_1 = -\nabla f\left(X^{(1)}\right) = -\nabla f\left(\frac{1}{10}\left(-\tfrac{1}{9}, \tfrac{4}{9}\right)\right) \\ &\Longrightarrow d_1 = \tfrac{-1}{90}\begin{pmatrix} -16 \\ 32 \end{pmatrix} = \tfrac{1}{45}\begin{pmatrix} 8 \\ -16 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \rho_1 = \arg\min_{\rho \geqslant 0} \phi_1(\rho) = \arg\min_{\rho \geqslant 0} f\left(X^{(1)} + \rho d_1\right) &= \arg\min_{\rho \geqslant 0} f\left(\frac{1}{90}(-1,4) + \rho\frac{1}{90}(16,-32)\right) \\ &= \arg\min_{\rho \geqslant 0}\left[\tfrac{4}{(90)^2}\left(2(-1+16\rho)^2 + (4-32\rho)^2\right)\right] \end{aligned}$$

ρ_1 is solution of $\phi_1'(\rho) = 0$

$$\begin{aligned} \phi_1'(\rho) = 0 &\Longrightarrow 4 \times 16(-1+16\rho) - 2 \times 32(4-32\rho) = 0 \\ &\Longrightarrow (-1+16\rho) - (4-32\rho) = 0 \Longrightarrow \rho = \tfrac{5}{48}. \end{aligned}$$

Thus $\rho_1 = \frac{5}{48}$

$$X^{(2)} = X^{(1)} + \rho_1 d_1 = \frac{1}{90}\begin{pmatrix} -1 \\ 4 \end{pmatrix} + \frac{5}{48}\frac{1}{45}\begin{pmatrix} 8 \\ -16 \end{pmatrix} = \begin{pmatrix} \frac{1}{135} \\ \frac{1}{135} \end{pmatrix}$$

$$d_2 = -\nabla f\left(X^{(2)}\right) = -\nabla f\left(\tfrac{1}{135}(1,1)\right) = -\tfrac{1}{135}\begin{pmatrix} 16 \\ 8 \end{pmatrix}.$$

$$\left\|\nabla f\left(X^{(2)}\right)\right\| = \tfrac{1}{135}\sqrt{(16)^2 + (8)^2} = \tfrac{8}{135}\sqrt{5} = 0.132\,51 < 0,5$$

whence $\overline{X} \simeq \begin{pmatrix} \frac{1}{135} \\ \frac{1}{135} \end{pmatrix}$.

Theorem 3.3.2 *(Convergence) Let $f \in C^1(\mathbb{R}^n, \mathbb{R})$ be a coercive function. Then:*

1. *The sequence $\{X^{(k)}\}_k$ generated by the gradient method with optimal step size is well-defined (i.e., ρ_k exists but is not necessarily unique).*

2. *The sequence $\{X^{(k)}\}_k$ is bounded and therefore admits a convergent subsequence. If $\{X^{(l)}\}_l$ is a convergent subsequence of $\{X^{(k)}\}_k$ with limit X^*, then $\nabla f(X^*) = 0_{\mathbb{R}^n}$.*

3. *If f is convex, then $X^* = \lim_{l\to\infty} X^{(l)}$ is a global minimum of f on $\mathbb{R}^n$.*

4. *If f is strictly convex, then the entire sequence $X^{(k)} \to X^*$ and:*

- $f(X^*) = \min_{x\in\mathbb{R}^n} f(x)$
- *X^* is the unique global minimum of f on $\mathbb{R}^n$*

3.3.3 Gradient method with variable step size

In this method, we do not necessarily take the optimal step size or a fixed step size; in other words, the step size varies from one iteration to another. Thus, the algorithm takes the following form:

1. **Initialization** $(k = 0)$
   $X^{(0)}$, $\varepsilon > 0$ given.

2. Compute $d_k = -\nabla f\left(X^{(k)}\right),\ t > 0$ such that $f\left(X^{(k)} + td_k\right) < f\left(X^{(k)}\right)$

3. Set $X^{(k+1)} = X^{(k)} + td_k$.

4. If $\|\nabla f\left(X^{(k)}\right)\| \leqslant \varepsilon$, stop; otherwise, set $k \longrightarrow k+1$ and return to step 2.

3.4 Conjugate gradient method

This method was discovered in 1952 for minimizing a quadratic function:

$$f(x) = \frac{1}{2}x^T Ax - b^T x + c$$

where A is a symmetric positive semi-definite matrix, $b, x \in \mathbb{R}^n$, and $c \in \mathbb{R}$. The minimum of f over $\mathbb{R}^n$ is x^* such that $Ax^* = b$.

Definition 3.4.1 *Let $A \in M_n(\mathbb{R})$ be a symmetric positive-definite matrix.*

1. *Two vectors $x, y \in \mathbb{R}^n - \{0_{\mathbb{R}^n}\}$ are said to be A-conjugate if*

$$\langle Ax, y\rangle = \langle x, y\rangle_A = 0.$$

2. *A family* $\{w_1, w_2, ..., w_n\} \subset \mathbb{R}^n - \{0_{\mathbb{R}^n}\}$ *is called A-conjugate if*

$$\langle w_i, w_j \rangle_A = w_i^t A w_j = 0 \ \forall i \neq j.$$

Remark 3.4.1

1. *If* A *is symmetric positive definite, then* $\langle \cdot, \cdot \rangle_A$ *defines an inner product on* $\mathbb{R}^n$.
2. *If a family* $\{w_1, w_2, \dots, w_n\}$ *is A-conjugate* $\Longrightarrow$ $\{w_1, w_2, \dots, w_n\}$ *is linearly independent.*

   The general idea of the conjugate gradient method is to construct directions $d_0, d_1, \dots, d_n$ *that are pairwise A-conjugate. At each iteration* k*, the direction* d_k *is obtained as a linear combination of* $g_k = \nabla f(X^{(k)})$ *and the previous direction. We set:*

   $$d_k = g_k + \alpha_k d_{k-1}$$

   where $g_k = \nabla f(X^{(k)}) = AX^{(k)} - b$ *and* $\alpha_k \in \mathbb{R}$ *is chosen such that:*

   $$\langle d_k, d_{k-1} \rangle_A = 0.$$

   Therefore, the conjugate gradient algorithm for minimizing the quadratic function:

1. ***Initialization***

   Choose $X^{(0)}$ *and* $\epsilon > 0$
   Compute $g_0 = \nabla f\left(X^{(0)}\right) = AX^{(0)} - b.$

2. ***Iteration***

   If $g_k = 0$ *or* $\|g_k\| \leqslant \epsilon$*, Stop*
   Otherwise

   (a) *Compute*

   $$d_k = \begin{cases} g_0 & \text{if} \quad k = 0 \\ g_k + \alpha_k d_{k-1} & \text{if} \quad k \geqslant 1 \end{cases} \quad , \text{ where } \alpha_k = -\frac{\langle g_k, \ Ad_{k-1} \rangle}{\langle Ad_{k-1}, \ d_{k-1} \rangle}$$

   (b) *Compute*

   $$\rho_k = \frac{\langle g_k, \ d_k \rangle}{\langle Ad_k, \ d_k \rangle}$$

   (c) $X^{(k+1)} = X^{(k)} - \rho_k d_k.$

   (d) *Compute* $g_{k+1} = AX^{(k+1)} - b.$

3. If $g_{k+1} = 0$, stop; otherwise set $k \longrightarrow k+1$ and return to step 2.

Example. Consider the function defined by:

$$f(x,y) = 3(x^2+y^2) = 3(x,y)I_2\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2}(x,y)6I_2\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2}X^tAX \quad (A = 6I_2).$$

$$\nabla f(x,y) = \begin{pmatrix} 6x \\ 6y \end{pmatrix} = 6I_2\begin{pmatrix} x \\ y \end{pmatrix}, \quad \nabla^2 f(x,y) = \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix} = 6I_2.$$

1. **Initialization**
   $k=0, \quad X^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad g_0 = \nabla f\left(X^{(0)}\right) = 6I_2X^{(0)} = 6\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \end{pmatrix} = d_0.$

2. **Iteration**
   $\rho_0 = \dfrac{\langle g_0,\ d_0\rangle}{\langle Ad_0,\ d_0\rangle} = \dfrac{1}{6}, \quad X^{(1)} = X^{(0)} - \dfrac{1}{6}d_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{6}\begin{pmatrix} 6 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$
   $g_1 = AX^{(1)} - b = 6I_2\begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. $g_1 = 0 \Longrightarrow X^* = X^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is the minimum of f on $\mathbb{R}^2$.

Theorem 3.4.1 *Let $f(x) = \dfrac{1}{2}x^tAx - b^tx + c$, where A is a symmetric matrix of order n. The sequence $\left(x^{(k)}\right)_k$ defined by the conjugate gradient method converges to the unique solution of $\min\limits_{x\in\mathbb{R}^n} f(x)$ in at most n iterations ($n = \dim \mathbb{R}^n$).*

3.5 Newton's method

Newton's method is not strictly speaking an optimization method.
In fact, it is a method used to solve nonlinear equations of the form $F(x) = 0$, where F is a function from $\mathbb{R}^n$ to $\mathbb{R}^n$. We will first describe it and then show how it can be applied to finding minima. This method directly searches for critical points, that is to say points x such that

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

We know Newton's method:

3.5.1 Newton's method in $\mathbb{R}$

Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be of class $C^1(\mathbb{R})$ with $f'(x) \neq 0, \forall x \in \mathbb{R}$. Newton's algorithm for solving $f(x) = 0$ is given by:

$$\begin{cases} x_0 \in \mathbb{R} \quad \text{given}, \epsilon > 0 \\ x_{k+1} = x_k - \dfrac{f(x_k)}{f'(x_k)} \\ \text{If } \left|x_{k+1} - x_k\right| < \epsilon \quad \text{or} \quad |f(x_k)| < \epsilon \text{ stop} \\ \text{Otherwise set} \quad k \longrightarrow k+1 \end{cases}$$

3.5.2 Newton's method in $\mathbb{R}^n$

Let

$$\begin{array}{rcl} F:\mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ X=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \longmapsto & F(X) \end{array}$$

1. The zeros of F are the solutions to the equation: $F(X)=0_{\mathbb{R}^n}$. We assume that $F\in C^1(\mathbb{R}^n)$ and that $J_F(X)$ is invertible $\forall X\in\mathbb{R}^n$. Then Newton's algorithm in this case is as follows:

$$\begin{cases} 1.\ k=0: \quad X^{(0)} \quad \text{given, } \epsilon>0 \quad \text{given} \\ 2.\ X^{(k+1)}=X^{(k)}-\left[J_F\left(X^{(k)}\right)\right]^{-1}F\left(X^{(k)}\right) \\ 3.\ \text{If } \|X^{(k+1)}-X^{(k)}\|\leqslant\epsilon \quad \text{or} \quad \|F\left(X^{(k)}\right)\|\leqslant\epsilon \text{ stop} \\ 4.\ \text{Otherwise, set } k=k+1 \text{ and return to step 2.} \end{cases}$$

2. **Application for solving an optimization problem.**

We have seen that a necessary optimality condition is $\nabla f(x^*)=0$. This is a nonlinear equation (or rather a system of nonlinear equations) in $\mathbb{R}^n$ and we will use Newton's method to solve it. However, we will only obtain critical points of f: we will then need to verify whether they are indeed minima.
Here, $F=\nabla f$ is indeed a function from $\mathbb{R}^n$ to $\mathbb{R}^n$. The derivative of F is nothing other than the Hessian matrix of f:
Let $f:\mathbb{R}^n\longrightarrow\mathbb{R}$ be of class $C^2(\mathbb{R}^n)$, we assume that the matrix $H_f(X)=\nabla^2 f(X)$ exists and is invertible for all x, and that the equation $\nabla f(X)=0_{\mathbb{R}^n}$ has at least one solution x^*. Then Newton's algorithm for finding the minimum of f on $\mathbb{R}^n$ is as follows:

1. **Initialization** $(k=0)$

$$X^{(0)} \text{ given in the neighborhood of } x^*, \quad \epsilon>0, \text{ a given precision}$$

2. **Iteration** k

$$X^{(k+1)}=X^{(k)}-\left[\nabla^2 f\left(X^{(k)}\right)\right]^{-1}\nabla f\left(X^{(k)}\right)$$

3. **Stopping criterion**

   while $\|X^{(k+1)}-X^{(k)}\|\leqslant\epsilon$ or $\|\nabla f\left(X^{(k)}\right)\|\leqslant\epsilon$ stop, otherwise set $k=k+1$ and return to 2.

Remark 3.5.1

1. *Let $F(X)=\nabla f(X)$, then $J_F(X)=\nabla^2 f(X)=H_f(X)$ is the Hessian matrix of f.*

2. *The first-order Taylor expansion for the function ∇f gives:*
$$\nabla f\left(X^{(k+1)}+d_k\right) \approx \nabla f\left(X^{(k)}\right)+\nabla^2 f\left(X^{(k)}\right) d_k .$$
We seek $X^{(k+1)}$ such that $F\left(X^{(k+1)}\right)=\nabla f\left(X^{(k+1)}\right)=0$, with $X^{k+1}=X^k+\rho_k d_k=X^k+d_k$. Therefore $d_k=-\left[\nabla^2 f\left(X^{(k)}\right)\right]^{-1} \nabla f\left(X^{(k)}\right)$, where d_k in this case is called the Newton direction.

Theorem 3.5.1 *(Local convergence of Newton's method) Let $f: \mathbb{R}^n \longmapsto \mathbb{R}$ be a function of class $C^2(\mathbb{R}^n, \mathbb{R})$ and let X^* be a critical point of f. Assume that $H_f\left(X^*\right)=\nabla^2 f\left(X^*\right)$ is invertible.*

1. *There exists a ball $B(X^*, \epsilon)$ such that $\forall X^{(0)} \in B(X^*, \epsilon)$, the Newton sequence is well-defined (i.e., $X^{(k)} \in B(X^*, \epsilon)$) and the sequence $\left(X^{(k)}\right)_k$ converges to the unique critical point in the ball.*
2. *If the matrix $\nabla^2 f$ is Lipschitz continuous (i.e., $\exists k>0$ such that: $\|\nabla^2 f(X)-\nabla^2 f(Y)\| \leqslant k\|X-Y\|, \forall\, X, Y \in \mathbb{R}^n$), then $\left(X^{(k)}\right)_k$ converges quadratically to X^*, meaning $\exists\, c>0$ such that:*
$$\|X^{(k+1)}-X^*\| \leqslant c\|X^{(k)}-X^*\|^2 .$$

Example. Let f be a function defined by:
$$\begin{array}{rcl} f: \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ X=\begin{pmatrix} x \\ y \\ z \end{pmatrix} & \longmapsto & f(X)=e^x+e^y-x-ey+(z+1)^2 \end{array}$$

1) Let us compute: $\nabla f(X)=\nabla f(x, y, z)=\begin{pmatrix} e^x-1 \\ e^y-e \\ 2(z+1) \end{pmatrix}$, $\nabla^2 f(x, y, z)=H(x, y, z)=$
$\begin{pmatrix} e^x & 0 & 0 \\ 0 & e^y & 0 \\ 0 & 0 & 2 \end{pmatrix}$

Therefore $[H(x, y, z)]^{-1}=\begin{pmatrix} e^{-x} & 0 & 0 \\ 0 & e^{-y} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$

2) **Application:**
The Newton's method iteration is given by:
$$\begin{cases} X^{(0)}=\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^3 \\ X^{(k+1)}=X^{(k)}-\left[H_f\left(X^{(k)}\right)\right]^{-1} \nabla f\left(X^{(k)}\right) \end{cases}$$

$$\Longrightarrow \begin{cases} X^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}; \\ \begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \\ z^{(k+1)} \end{pmatrix} = \begin{pmatrix} x^{(k)} \\ y^{(k)} \\ z^{(k)} \end{pmatrix} - \begin{pmatrix} e^{-x^{(k)}} & 0 & 0 \\ 0 & e^{-y^{(k)}} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} e^{x^{(k)}} - 1 \\ e^{y^{(k)}} - e \\ 2(z^{(k)} + 1) \end{pmatrix} \end{cases}$$

$$\Longrightarrow \begin{cases} X^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \\ z^{(k+1)} \end{pmatrix} = \begin{pmatrix} x^{(k)} - 1 + e^{-x^{(k)}} \\ y^{(k)} - 1 + e^{1-y^{(k)}} \\ -1 \end{pmatrix} \end{cases}$$

k	0	1	2	3	...
$x^{(k)}$	1	0.3678	0.060	1.7645×10^{-3}	...
$y^{(k)}$	0	1.7182	1.2058	1.01978	...
$z^{(k)}$	0	-1	-1	-1	...

Therefore $X^* \simeq \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.

3.6 Relaxation method

The idea of this method is to reduce a minimization problem in $\mathbb{R}^n$ to the successive solution of n one-dimensional minimization problems (at each iteration). Thus, the algorithm takes the following form:

1. **Initialization** $k = 0$

   Choose $X^{(0)} \in \mathbb{R}^n$ and $\varepsilon > 0$ as a given precision threshold.

2. **Iteration** k

   For each $i = 1, 2, ..., n$ compute $x_i^{(k+1)}$ as the solution to the problem:

   $$\min_{t \in \mathbb{R}} \left[\phi_i^{(k)}(t) = f\left(x_1^{(k+1)}, x_2^{(k+1)}, ..., x_{i-1}^{(k+1)}, t, x_{i+1}^{(k)}, ..., x_n^{(k)} \right) \right].$$

3. If $\|X^{(k+1)} - X^{(k)}\| \leqslant \varepsilon$ stop, otherwise set $k \longrightarrow k+1$ and return to step 2.

Example. Consider the problem

$$(p)\begin{cases}\min f(x,y)=y^2+x^2+xy-15\\ (x,y)^t\in\mathbb{R}^2\end{cases}$$

1. We have $f(x,y)=x^2+y^2+xy-15$, $\nabla f(x,y)=\begin{pmatrix}2x+y\\ x+2y\end{pmatrix}$ and $\nabla^2 f(x,y)=\begin{pmatrix}2&1\\1&2\end{pmatrix}$. Thus f is a quadratic form such that:

$$A=H_f(x,y)=\nabla^2 f(x,y)=\begin{pmatrix}2&1\\1&2\end{pmatrix}.$$

$\Delta_1=\det(2)=2>0,\ \Delta_2=\det\begin{pmatrix}2&1\\1&2\end{pmatrix}=4-1=3>0\Longrightarrow$ by Sylvester's criterion, $A=H_f(x,y)$ is positive definite.

$\begin{cases}f \text{ continuous}\\ f \text{ coercive}\end{cases}$ on $\mathbb{R}^2\Longrightarrow(P)$ has at least one solution $\bar{X}$.

f is strictly convex on $\mathbb{R}^2\Longrightarrow$ the problem (P) has at most one solution $\bar{X}$

Therefore the problem (P) has a unique solution denoted $\bar{X}=\begin{pmatrix}\bar{x}\\ \bar{y}\end{pmatrix}$.

1. Computing $\bar{X}=(\bar{x},\bar{y})^t$.

$$A\overline{X}=0_{\mathbb{R}^2}\Longrightarrow\bar{X}=0_{\mathbb{R}^2},\text{ since } A \text{ is invertible}$$

Thus $\bar{X}=(\bar{x},\bar{y})^t=(0,0)^t$.

2. **Relaxation Method**

$$\begin{cases}X^{(0)}=\begin{pmatrix}0\\1\end{pmatrix}\in\mathbb{R}^2,\text{ given}\\ X^{(k+1)}=\left(x_i^{(k+1)}\right)_{0\leqslant i\leqslant n}\quad\text{such that}\quad\forall i=1,...,n\\ x_i^{(k+1)}=\arg\min_{\mathbb{R}} f\left(x_1^{(k+1)},...,x_{i-1}^{(k+1)},t,x_{i+1}^{(k)},...,x_n^{(k)}\right)\end{cases}$$

a. Compute $X^{(1)}=\begin{pmatrix}x^{(1)}\\ y^{(1)}\end{pmatrix}$:

i. Compute $x^{(1)}$:

$f(t,y^{(0)})=f(t,1)=t^2+1+t-15,\ f'(t,1)=0\Longrightarrow 2t+1=0\Longrightarrow t=\dfrac{-1}{2}\Longrightarrow x^{(1)}=\dfrac{-1}{2}$.

ii. Compute $y^{(1)}$:

$f(x^{(1)},t)=f\left(\dfrac{-1}{2},t\right)=\dfrac{1}{4}+t^2-\dfrac{1}{2}t-15,\ f'\left(\dfrac{-1}{2},t\right)=0\Longrightarrow 2t-\dfrac{1}{2}=0\Longrightarrow t=\dfrac{1}{4}\Longrightarrow y^{(1)}=\dfrac{1}{4}$.

b. Compute $X^{(2)} = \begin{pmatrix} x^{(2)} \\ y^{(2)} \end{pmatrix}$:

i. Compute $x^{(2)}$:

$f(t, y^{(1)}) = f\left(t, \frac{1}{4}\right) = t^2 + \frac{1}{16} + t + \frac{1}{4}t - 15,\ f'\left(t, \frac{1}{4}\right) = 0 \Longrightarrow 2t + \frac{1}{4} = 0 \Longrightarrow t = \frac{-1}{8} \Longrightarrow x^{(2)} = \frac{-1}{8}$.

ii. Compute $y^{(2)}$:

$$f\left(\frac{-1}{8}, t\right) = \frac{1}{64} + t^2 - \frac{1}{8}t - 15,\ y^{(2)} = \frac{1}{16}.$$

k	0	1	2	3	4
$x^{(k)}$	0	$-1/2$	$-1/8$	?	?
$y^{(k)}$	1	$1/4$	$1/16$	?	?

3.7 Exercises

Exercise 3.1

Apply 4 iterations of the fixed-step gradient method ($\rho = \frac{1}{6}$) to solve the following problem:

$$\min_{\mathbb{R}^2} \left(2x^2 + 4y^2 + 4xy - 3y\right);$$

starting from the point $(2, 2)^T$.

Exercise 3.2

Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a function defined by:

$$f(x, y) = \frac{1}{2}\left(\alpha x^2 + y^2\right),$$

where $\alpha \in]1, +\infty[$.

1. Compute $\nabla f(x, y)$ and the Hessian matrix $A = \nabla^2 f(x, y)$.
2. Deduce that the function f has a unique global minimum X^* that we will determine.
3. Consider $(z^{(k)})_k$ a sequence of points in $\mathbb{R}^2$ defined by:

$$\begin{cases} z^{(0)} = (1, \alpha)^t \\ d^{(k)} = Az^{(k)} \\ \rho_k = \frac{\left\|d^{(k)}\right\|^2}{\left\langle Ad^{(k)}; d^{(k)}\right\rangle} \\ z^{(k+1)} = z^{(k)} - \rho_k d^{(k)} \end{cases}$$

   (a) Show that

$$z^{(k)} = \left(\tfrac{\alpha-1}{\alpha+1}\right)^k \begin{pmatrix} (-1)^k \\ \alpha \end{pmatrix}$$

(b) Calculate the number of iterations performed such that $e_k = \left\| z^{(k)} - X^* \right\| \leq 10^{-6}$.

Exercise 3.3

Let f be the function defined on $\mathbb{R}^3$ by:

$$f(x, y, z) = e^x + e^y - x - ey + (z+1)^2$$

1. Compute ∇f and $\nabla^2 f$ at any point (x, y, z) of $\mathbb{R}^3$.
2. Apply Newton's method to solve $\min\limits_{\mathbb{R}^3} f(x, y, z)$ using the initial point $X^{(0)} = (1, 0, 0)^T$.

Exercise 3.4

Consider the unconstrained optimization problem

$$\min_{\mathbb{R}^2} f(x, y) = 3x^2 + 3y^2$$

1. Let's examine two methods: Newton's method and the conjugate gradient method. Perform two iterations of each method using the starting point $X^{(0)} = (1, 1)^T$.
2. Analyze and compare the results.

Exercise 3.5

Consider the following unconstrained optimization problem:

$$(p) \begin{cases} \min f(x, y) = y^2 + x^2 + xy - 15 \\ (x, y)^T \in \mathbb{R}^2 \end{cases}$$

1. Show that problem (p) has a unique solution $\overline{X} = (\overline{x}, \overline{y})^T$.
2. Compute $\overline{X}$.
3. Consider the following algorithms:
   (a) Fixed-step gradient method $\left(\rho = \dfrac{1}{9}\right)$;
   (b) Optimal-step gradient method;
   (c) Relaxation method.

Perform two iterations for each method using the initial point $X^{(0)} = (0, 1)^T$.

1. Compute the error $e_2 = \left\| X^{(2)} - \overline{X} \right\|$ for each method.
2. Analyze and compare the results.

3.8 Exam style without solutions

Exercise 1. Let f be a function defined on $\mathbb{R}^3$ by $f(x,y,z) = x^4 - 2x^2y + 2y^2 - 2yz + 2z^2 - 4z + 5$

1. Determine the critical points of f on $\mathbb{R}^3$.
2. Show that the expression $f(x,y,z)$ can be written as a sum of squares.
3. Deduce the solutions of $\min\limits_{\mathbb{R}^3} f$.
4. Conclude that f is not strictly convex on $\mathbb{R}^3$.

Exercise 2. Consider the function defined on $\mathbb{R}^2$ by:

$$f_{\alpha,\beta}(X) = f_{\alpha,\beta}(x,y) = \frac{\alpha}{2}\left(x^2+y^2\right) - \beta xy$$

where $\alpha, \beta \in \mathbb{R}$.

1. Compute $\nabla f_{\alpha,\beta}(x,y)$ and $\nabla^2 f_{\alpha,\beta}(x,y)$.
2. Determine conditions on the values α and β such that the problem $\min\limits_{\mathbb{R}^2} f_{\alpha,\beta}(x,y)$ has a unique solution.
3. Let $\alpha = 2$ and $\beta = 1$, show that in this case $\min\limits_{\mathbb{R}^2} f_{2,1}(x,y)$ has a unique solution $\overline{X}$.
4. Compute $\overline{X}$.

Exercise 3. Consider the optimization problem

$$(p)\begin{cases} \min \left[f(x,y) = 20x^2 - 5xy + \dfrac{5}{2}y^2 \right] \\ \qquad (x,y) \in \mathbb{R}^2 \end{cases}$$

1. Show that problem (p) has a unique solution $\overline{X} = (\overline{x}, \overline{y})^T$.
2. Compute $\overline{X}$.
3. Consider the following algorithms:
   (a) Optimal-step gradient method;
   (b) Relaxation method.

   Perform two iterations for each method using the initial point $X^{(0)} = (1,1)^T$.
4. Compute the error $e_2 = \left\| X^{(2)} - \overline{X} \right\|$ for each method.
5. Analyze and compare the results.

3.9 Conclusion

In this lecture notes, we have reviewed some basic concepts of differential calculus and a bit of matrix calculus, as well as notions of convexity for sets and multivariable functions, which are necessary for studying unconstrained optimization problems. We have presented the fundamental algorithms for unconstrained optimization in accordance with the program of the Licence 3 degree in Mathematics. Various exercises accompany the document to help assimilate the more theoretical concepts covered in the course. The study of constrained optimization problems will be the subject of a future lecture notes.

Bibliography

[1] J. Céa, *Optimisation: théorie et algorithmes*, Dunod 1971.

[2] J.C. Culioli, *Introduction à l'optimisation*, Ellipses, Paris, 1994.

[3] P.Ciarlet, *Introduction à l'analyse matricielle et à l'optimisation*, Masson, Paris, 1988.

[4] M. S. Bazarra, H.D. Sherali and C.M. Shetty, *Nonlinear programming: Theory and Algorithms*, Wiley Interscience, Third Edition 2006.

[5] Dimitri P. Bertsekas, Nonlinear programming, Second Edition, 1999.