Optimization

Lecture Notes

Athanasios P. Liavas
School of Electrical and Computer Engineering
Technical University of Crete
Email: liavas@telecom.tuc.gr

Date: 15/10/2023
“Support for the Internationalization of Higher Education, School of Electrical and Computer Engineering, Technical University of Crete” (MIS 5150766), under the call for proposals “Support of the Internationalization of Higher Education- Technical University of Crete” (EDULLL 153). The project is co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme “Human Resources Development, Education and Lifelong Learning 2014-2020.”
Introduction

These notes are material that is developed during the teaching of the course “Optimization.” Optimization is an area of Applied Mathematics with an extremely large impact on Engineering Sciences, Economics, Operational Research etc.

The area of (Convex) Optimization is extremely active, with a wealth of contributions by members of various scientific communities, such as Mathematicians, Engineers, Economists.

In this course, we will cover a relatively small subset of the region, which could be characterized as introductory to the area of convex optimization.

Coverage will include

1. brief review of real function theory concepts,
2. elements of the theory of convex sets and convex functions,
3. definition of basic concepts of (convex) optimization problems,
4. extraction of optimality conditions,
5. definition of unconstrained convex optimization problems and the steepest descent and Newton algorithms,
6. definition of convex optimization problems with linear constraints and solution by a generalization of the Newton algorithm,
7. definition of convex optimization problems and solution by interior point method.

The following excellent books have significantly influenced the formatting of this material:


A very useful book covering vector calculus is the following:

Your comments and suggestions are most welcome.

The instructor,

Athanasios P. Liavas
Chapter 1

Euclidean spaces

In this chapter, we will briefly present some basic elements of the theory of Euclidean spaces. The book by Marsden and Tromba contains an extensive description and is an excellent source.

We assume that the reader is familiar with the $n$-dimensional Euclidean space $\mathbb{R}^n$, with $n \in \mathbb{N}$. If $n = 1$, then the corresponding space will be denoted as $\mathbb{R}$. In the remainder of the chapter, we generally assume that we are working on $\mathbb{R}^n$.

1.1 Points - Vectors

Definition 1.1.1. Vector is called every ordered $n$-tuple, $(x_1, x_2, \ldots, x_n)$, with $x_i \in \mathbb{R}$, for $i = 1, \ldots, n$.

![Figure 1.1: Point in three-dimensional space.](image)

Figure 1.1: Point in three-dimensional space.
Each vector will be denoted in three equivalent ways. More specifically, the three-dimensional vector \( \mathbf{v} = (x, y, z) \) will be denoted as

\[
\mathbf{v} = (x, y, z) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}^T = \begin{bmatrix} x \\ y \\ z \end{bmatrix},
\]

where the symbol \([\cdot]^T\) means transpose. Each vector corresponds to a point in \( n \)-dimensional space. For example, in Figure 1.1, we draw the point \( \mathbf{x}^* = (x^*_1, x^*_2, x^*_3) \).

A more “geometric” representation of a vector \( \mathbf{x}^* = (x^*_1, x^*_2, x^*_3) \) is given by the directed line segment, starting at the origin of the axes, i.e., the point \( \mathbf{0} = (0, 0, 0) \), and ending at the point \( \mathbf{x}^* \) (see Figure 1.2).

### 1.1.1 Vector Operations

Let \( \mathbf{x} = (x_1, x_2, \ldots, x_n) \) and \( \mathbf{y} = (y_1, y_2, \ldots, y_n) \) be vectors. Their sum \( \mathbf{z} = \mathbf{x} + \mathbf{y} \) is defined as the vector \( \mathbf{z} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \).

Schematically, the sum of the vectors \( \mathbf{x} \) and \( \mathbf{y} \) is represented by the diagonal of the parallelogram with sides the vectors \( \mathbf{x} \) and \( \mathbf{y} \) (see Fig. 1.3, for \( n = 2 \)).

The vector \( \mathbf{y} - \mathbf{x} = (y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n) \) is the vector we should add to \( \mathbf{x} \) to get \( \mathbf{y} \) (see Figure 1.3, for \( n = 2 \)).

Let the vector \( \mathbf{x} = (x_1, \ldots, x_n) \). Then, the vector \( a\mathbf{x} = (ax_1, \ldots, ax_n) \), with \( a \in \mathbb{R} \), is
called \textbf{scaling} of \( \mathbf{x} \). If \( a > 0 \), then \( a \mathbf{x} \) has the same direction as \( \mathbf{x} \) while, if \( a < 0 \), then \( a \mathbf{x} \) has direction opposite to that of \( \mathbf{x} \).

\subsection*{1.1.2 Inner product and Norm}

\textbf{Definition 1.1.2.} Let the vectors \( \mathbf{x} = (x_1, \ldots, x_n) \) and \( \mathbf{y} = (y_1, \ldots, y_n) \). Their inner product is defined as follows

\[
\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x} = \sum_{i=1}^{n} x_i y_i. \tag{1.2}
\]

\textbf{Definition 1.1.3.} As the Euclidean norm of the vector \( \mathbf{x} = (x_1, \ldots, x_n) \), we define the non-negative real number

\[
\| \mathbf{x} \|_2 = \langle \mathbf{x}, \mathbf{x} \rangle^{\frac{1}{2}} = (x_1^2 + \cdots + x_n^2)^{\frac{1}{2}}. \tag{1.3}
\]

The following important properties can be proved.

1. If \( \mathbf{x} \in \mathbb{R}^n \), then \( \| \mathbf{x} \|_2 \geq 0 \), with equality if, and only if, \( \mathbf{x} = \mathbf{0} \).

2. If \( \mathbf{x} \in \mathbb{R}^n \) and \( a \in \mathbb{R} \), then \( \| a \mathbf{x} \|_2 = |a| \| \mathbf{x} \|_2 \).

3. If \( \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \), then \( \| \mathbf{x} + \mathbf{y} \|_2 \leq \| \mathbf{x} \|_2 + \| \mathbf{y} \|_2 \), with equality if, and only if, \( \mathbf{x} = c \mathbf{y} \), with \( c \in \mathbb{R} \).

Any function \( \| \cdot \| : \mathbb{R}^n \rightarrow \mathbb{R}_+ \) satisfying the above conditions is called a norm function on \( \mathbb{R}^n \).
Figure 1.4: A line that passes through the point \( \mathbf{x} \) and is parallel to the vector \( \mathbf{v} \).

**Theorem 1.1.1.** Let the vectors \( \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \). Then,

\[
\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \cos \angle(\mathbf{x}, \mathbf{y}),
\]

where \( \angle(\mathbf{x}, \mathbf{y}) \) is the angle between the vectors \( \mathbf{x} \) and \( \mathbf{y} \).

**Proof.** For \( n = 3 \), see page 18 of the book by Marsden and Tromba. \( \square \)

**Corollary 1.1.1.** (Inequality Cauchy-Schwarz) Let \( \mathbf{x} \) and \( \mathbf{y} \) be vectors in \( \mathbb{R}^n \). Then

\[
|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2
\]

with equality if, and only if, \( \mathbf{x} \) and \( \mathbf{y} \) are collinear, that is, \( \mathbf{x} = a \mathbf{y} \) for \( a \in \mathbb{R} \).

**Proof.** The proof is obvious from Theorem 1.1.1 and the fact that \( |\cos(\theta)| \leq 1 \), for every angle \( \theta \). \( \square \)

**Definition 1.1.4.** The vectors \( \mathbf{x} \) and \( \mathbf{y} \) in \( \mathbb{R}^n \) are called orthogonal if \( \langle \mathbf{x}, \mathbf{y} \rangle = 0 \).

If the vectors \( \mathbf{x} \) and \( \mathbf{y} \) are orthogonal, then we write \( \mathbf{x} \perp \mathbf{y} \).

### 1.1.3 Lines

Let the vectors \( \mathbf{x}, \mathbf{v} \in \mathbb{R}^n \). The points described by the relationship

\[
\mathbf{l}(t) = \mathbf{x} + t\mathbf{v},
\]

for \( t \in \mathbb{R} \), define the line which passes through point \( \mathbf{x} \) and is parallel to the vector \( \mathbf{v} \) (see Figure 1.4).
1.1. POINTS - VECTORS

1.1.4 Straight line passing through the points $x$ and $y$

Let the vectors $x, y \in \mathbb{R}^n$. The points defined by the relation

$$l(t) = x + t(y - x), \quad \text{for } t \in \mathbb{R},$$

(1.6)

define the line that passes through the points $x$ and $y$ (see Figure 1.5).

We observe that, for $t = 0$, $l(t) = x$ while, for $t = 1$, $l(t) = y$. An alternative description of $l(t)$ is as follows:

$$l(t) = (1 - t)x + ty.$$  

(1.7)

For $t \in [0, 1]$, relation (1.7) defines the line segment connecting the points $x$ and $y$.

1.1.5 Plane perpendicular to a vector

A plane $\mathbb{P}$ is a set of points for which it holds true that all points of the straight lines connecting any two points of $\mathbb{P}$ belong to $\mathbb{P}$. That is, if $x_1, x_2 \in \mathbb{P}$ and $\theta \in \mathbb{R}$, then $\theta x_1 + (1 - \theta)x_2 \in \mathbb{P}$.

Let $\mathbb{P}$ be a plane in $\mathbb{R}^n$, $y$ a point of $\mathbb{P}$, and $a \in \mathbb{R}^n$ vector perpendicular to $\mathbb{P}$. Then, for every $x \in \mathbb{P}$, the vector $x - y$ is parallel to $\mathbb{P}$. Therefore, the vector $a$ is perpendicular to $x - y$.

The mathematical description of the points of $\mathbb{P}$ is as follows:

$$\mathbb{P} = \{x \in \mathbb{R}^n | \langle x - y, a \rangle = 0 \}$$

$$= \{x \in \mathbb{R}^n | x^T a = y^T a \}$$

$$= \{x \in \mathbb{R}^n | x^T a = c \},$$  

(1.8)

where $c := y^T a$.

In general, we say that the equation (1.8) defines the hyperplane in $\mathbb{R}^n$ which passes through the point $y$ and is perpendicular to the vector $a$. 

![Figure 1.5: A straight line that passes through the points x and y.](image-url)
If \( n = 2 \), then the hyperplane is simply a straight line, while, if \( n = 3 \), then the hyperplane is a plane. In general, a hyperplane in \( \mathbb{R}^n \) is a plane set with dimension \((n-1)\) (equivalently, with \((n-1)\) degrees of freedom). Expression (1.8) will prove extremely important later.

**Exercise:** Prove that the set \( \mathbb{P} \), defined by the relation (1.8), is a plane.

### 1.1.6 Open and Closed Sets

**Definition 1.1.5.** Let \( x \in \mathbb{R}^n \) and \( r \in \mathbb{R}_{++} \). The set
\[
\mathbb{B}(x, r) = \{ y \in \mathbb{R}^n : \| y - x \|_2 < r \}
\] (1.9)
is called the (Euclidean) ball with center \( x \) and radius \( r \).

The set \( \mathbb{B}(x, r) \) is also called a **neighborhood** of \( x \).

**Definition 1.1.6.** The point \( x \in S \) is called **interior point** of \( S \) if \( S \) contains a neighborhood of \( x \), that is, if there exists \( r_x \in \mathbb{R}_{++} \) such that \( \mathbb{B}(x, r_x) \subset S \).

**Definition 1.1.7.** The set of interior points of the set \( S \) is called the **interior** of \( S \).

**Definition 1.1.8.** A point \( x \) is called a **boundary point** of the set \( S \) if every neighborhood of \( x \) contains at least one point that belongs to \( S \) and at least one point which does not belong to \( S \). The set of all its boundary points \( S \) is called the **boundary** of \( S \).

A boundary point of the set \( S \) may not belong to \( S \).

**Definition 1.1.9.** A set is called **open** if it contains only its interior points or, equivalently, if it contains none of its boundary points.

**Definition 1.1.10.** A set is called **closed** if it contains its boundary.

It can be shown that a set is open if, and only if, its complement is closed.

A set may be neither open nor closed (give examples).

**Definition 1.1.11.** A set is called **bounded** if it is contained in a ball of finite radius.

**Definition 1.1.12.** A set is called **compact** if it is closed and bounded.
Theorem 1.1.2. (Theorem Weierstrass) Let \( U \subseteq \mathbb{R}^n \) be a compact set and let \( f : U \rightarrow \mathbb{R} \) be a continuous function. Then, there exists a point \( x_0 \in U \) such that \( f(x_0) \leq f(x) \), for every \( x \in U \).

Proof. The proof requires the use of some advanced Real Analysis concepts and is beyond the scope of this course.

In words, this important theorem says that, if \( U \) is compact and \( f \) is continuous, then there is a point of \( U \) which attains the minimum value of \( f \) in \( U \). If \( U \) is not closed or bounded, then we can easily construct examples in which the minimum value of \( f \) in \( U \) is not attained.

1.2 Elements of function theory

1.2.1 Functions

Definition 1.2.1. Let \( U \subseteq \mathbb{R}^n \). The transformation \( f : U \rightarrow \mathbb{R}^m \), which maps each element \( x \in U \) to one element of \( \mathbb{R}^m \), is called a function from \( U \) into \( \mathbb{R}^m \).

If in the definition 1.2.1 we have \( m = 1 \), then \( f \) is called a real function of \( n \) variables and is denoted as \( f(x) \) or as \( f(x_1, \ldots, x_n) \).

For example, the function
\[
f(x_1, x_2) = \sqrt{x_1^2 + x_2^2}, \quad \text{with } x_1, x_2 \in \mathbb{R},
\]
is a real function of two variables, while
\[
f(x_1, x_2, x_3) = x_1(x_2^2 + \log_2 x_3), \quad \text{with } x_1, x_2 \in \mathbb{R}, x_3 \in \mathbb{R}^+, \]
is a real function of three variables.

Definition 1.2.2. Let \( f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \). Then, the set
\[
G_f = \{(x, f(x)) \in \mathbb{R}^{n+1} \mid x \in U\}
\]
is called the graph of \( f \).

If \( n = 1 \), then the graph is a curve in \( \mathbb{R}^2 \) (see Figure 1.6), while, if \( n = 2 \), then the graph is a surface in \( \mathbb{R}^3 \).

If \( n > 2 \), then the graph is a hypersurface in \( \mathbb{R}^{n+1} \), which cannot be visualized.
1.2.2 Curves

Definition 1.2.3. A function $f : A \subseteq \mathbb{R} \to \mathbb{R}^n$ is called a curve in $\mathbb{R}^n$.

Example 1.2.1. The function $f : [0, 2\pi] \to \mathbb{R}^2$, with $f(t) = (\cos(t), \sin(t))$ defines the circle centered at point $(0, 0)$ with radius 1.

1.2.3 Level Sets

Definition 1.2.4. Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ and $c \in \mathbb{R}$. The set
\[ S_c := \{ x \in U \mid f(x) = c \} \] (1.11)
is called the level set of $f$, for value $c$.

If $n = 2$, then we are talking about a level curve (see Figure 1.7), while if $n = 3$ then we are talking about a level surface.

1.2.4 Limits of Sequences - Continuity of Functions

We assume that the reader is familiar with the concepts of vector sequences and convergence of sequences of vectors as well as the concept of continuity of functions. The book by Marsden and Tromba contains extensive coverage of these basic concepts.
1.2.5 Derivative

The concept of differentiation is about approximating functions by linear (actually, affine) functions. Next, we list basic definitions and properties of derivative functions.

**Definition 1.2.5.** Let $U$ be an open subset of $\mathbb{R}^n$ and $f : U \rightarrow \mathbb{R}^m$. $f$ is called **differentiable** at the point $x \in U$ if there exists an $m \times n$ matrix $A_f^x$ such that

$$
\lim_{y \to x} \left( \frac{\|f(y) - f(x) - A_f^x(y - x)\|}{\|y - x\|} \right) = 0. \tag{1.12}
$$

The matrix $A_f^x$ is called the **derivative** (or Jacobian) of $f$ at the point $x$ and is denoted as $Df(x)$. The function $f$ is called differentiable if it is differentiable at any point $x \in U$.

Next, we state three important properties of the derivative as theorems.

**Theorem 1.2.1.** If $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ are differentiable at the point $x \in \mathbb{R}^n$, then $(f + g)(x) = f(x) + g(x)$ is also differentiable at $x$ and

$$
D(f + g)(x) = Df(x) + Dg(x). \tag{1.13}
$$

**Proof.** See page 101 of the book by Marsden and Tromba.

**Theorem 1.2.2.** If $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $g : \mathbb{R}^n \rightarrow \mathbb{R}$ are differentiable at the point $x \in \mathbb{R}^n$, then $(f \cdot g)(x) = f(x) \cdot g(x)$ is also differentiable at $x$ and

$$
D(f \cdot g)(x) = g(x)Df(x) + f(x)Dg(x). \tag{1.14}
$$

**Proof.** See page 101 of the book by Marsden and Tromba.

**Theorem 1.2.3.** Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $h : \mathbb{R}^k \rightarrow \mathbb{R}^n$. If $h$ is differentiable at $x \in \mathbb{R}^k$ and $f$ is differentiable at $h(x) \in \mathbb{R}^n$, then $f \circ h$ is differentiable at $x \in \mathbb{R}^k$ and

$$
D(f \circ h)(x) = Df(h(x)) Dh(x), \tag{1.15}
$$

where in the right side of (1.15) we have matrix multiplication.

**Proof.** See page 102 of the book by Marsden and Tromba.

In the Appendix, at the end of the chapter, we prove in detail a special case of Theorem 1.2.3.

If a function $f$ is differentiable and its derivative, $Df$, is a continuous function, then we say that $f$ belongs to the family of functions $C^1$. 
**Definition 1.2.6.** Let $\mathbb{U}$ be an open subset of $\mathbb{R}^n$, $f : \mathbb{U} \to \mathbb{R}$, $x \in \mathbb{U}$, and $e_j$, for $j = 1, \ldots, n$, the $n$-dimensional vector with elements 0 everywhere except from the $j$-th position in which it has element 1. If the limit exists

$$
\lim_{h \to 0} \frac{f(x + he_j) - f(x)}{h},
$$

then it is called the **partial derivative** of $f$ with respect to the $j$-th coordinate, $x_j$, at the point $x$, and is denoted as $\frac{\partial f}{\partial x_j}(x)$ or as $\frac{\partial f}{\partial x_j}(x_1, \ldots, x_n)$.

Next, we state an important theorem that connects the derivative of a function $f$ with its partial derivatives.

**Theorem 1.2.4.** Let $\mathbb{U} \subseteq \mathbb{R}^n$ be an open set and $f : \mathbb{U} \to \mathbb{R}$.

1. If $f$ is differentiable at $x \in \mathbb{U}$, then the partial derivatives $\frac{\partial f}{\partial x_j}(x)$ exist and the derivative is equal to $Df(x) = \left[ \frac{\partial f(x)}{\partial x_1} \cdots \frac{\partial f(x)}{\partial x_n} \right]$.

2. If the partial derivatives of $f$ at the point $x$ exist and are continuous, then $f$ is differentiable at $x$ and $Df(x) = \left[ \frac{\partial f(x)}{\partial x_1} \cdots \frac{\partial f(x)}{\partial x_n} \right]$.

3. $f$ is $C^1$ in $\mathbb{U}$ if, and only if, the partial derivatives of $f$ exist and are continuous in $\mathbb{U}$.

**Proof.** Check out a good Calculus book. \(\square\)

The existence of the partial derivatives of $f$ does not automatically imply the differentiability of $f$. The continuity of the partial derivatives is extremely important.

**Definition 1.2.7.** If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable, then the function $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$, which is defined as

$$
\nabla f(x) = \begin{bmatrix}
\frac{\partial f}{\partial x_1}(x) \\
\vdots \\
\frac{\partial f}{\partial x_n}(x)
\end{bmatrix} = Df(x)^T,
$$

is called the **gradient** of $f$ at the point $x$.

Later, we will prove some extremely important properties of $\nabla f(x)$. 

1.2. ELEMENTS OF FUNCTION THEORY

1.2.6 Derivative of a curve

Definition 1.2.9. If \( f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}^n \), with \( f(t) = (f_1(t), \ldots, f_n(t)) \), then the \((n \times 1)\) vector \( Df(t_0) = (Df_1(t_0), \ldots, Df_n(t_0))\) is called the derivative of the curve \( f \) at the point \((f_1(t_0), \ldots, f_n(t_0))\).

The derivative \( Df \) can be expressed as

\[
Df(t_0) = \lim_{h \to 0} \frac{f(t_0 + h) - f(t_0)}{h}. \tag{1.19}
\]

The vector \( Df(t_0) \) is parallel to the line which is tangent to the curve \( f \) at the point \( f(t_0) \) (see Figure 1.8) and is expressed as follows:

\[
l(t) = f(t_0) + tDf(t_0). \tag{1.20}
\]

Figure 1.8: Curve derivative.
1.2.7 Directional Derivative

**Definition 1.2.10.** Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$. Let $x \in U$ and $0 \neq h \in \mathbb{R}^n$. If the limit

$$\lim_{t \to 0^+} \left( \frac{f(x + th) - f(x)}{t} \right)$$

exists, it is called the **directional derivative** of $f$, at the point $x$ in the direction $h$, and is denoted as $Df(x; h)$.

If the function $f$ is differentiable, then an equivalent expression for the directional derivative is

$$Df(x; h) = \left. \frac{d}{dt} f(x + th) \right|_{t=0}.$$  \hspace{1cm} (1.22)

This holds because

$$\left. \frac{d}{dt} f(x + th) \right|_{t=0} = \lim_{\Delta t \to 0} \left. \frac{f(x + (t + \Delta t)h) - f(x + th)}{\Delta t} \right|_{t=0} = \lim_{\Delta t \to 0} \frac{f(x + \Delta th) - f(x)}{\Delta t}.$$  \hspace{1cm} (1.23)

Therefore, in this case, the two-sided limit, $\lim_{t \to 0}$ exists, and is equal to the limit from the right, $\lim_{t \to 0^+}$.

**Theorem 1.2.5.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function and $x \in \mathbb{R}^n$. Then, for each $h \in \mathbb{R}^n$, the directional derivative $Df(x; h)$ exists and is given by

$$Df(x; h) = Df(x) h = \nabla f(x)^T h.$$  \hspace{1cm} (1.24)

**Proof.** Let $c(t) = x + th$. Then $f(x + th) = f(c(t))$. Using the chain rule, we have that

$$\frac{d}{dt} f(c(t)) = Df(c(t)) Dc(t).$$  \hspace{1cm} (1.25)

Furthermore, $c(0) = x$ and $Dc(t) = h$. Therefore,

$$Df(x; h) = \left. \frac{d}{dt} f(x + th) \right|_{t=0} = \left. \frac{d}{dt} f(c(t)) \right|_{t=0} = Df(c(0)) Dc(0) = Df(x)h = \nabla f(x)^T h,$$

completing the proof.
Choosing \( h \in \mathbb{R}^n \) with \( \|h\|_2 = 1 \), we can interpret \( Df(x; h) \) as the rate of change of \( f \) at the point \( x \) in the direction \( h \).

A direct consequence of the Theorem 1.2.5 is the following.

**Corollary 1.2.1.** If \( Df(x; h) \) exists, then \( Df(x; -h) \) also exists and

\[
Df(x; -h) = -Df(x; h).
\]

(1.27)

An extremely important theorem related to the rate of change of a function \( f \) is as follows.

**Theorem 1.2.6.** Let \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) be a smooth real function and \( x_0 \in \mathbb{R}^n \), with \( \nabla f(x_0) \neq 0 \). Then, the vector \( \nabla f(x_0) \) “points” to the direction along which the rate of increase of \( f \) at \( x_0 \) is maximum, while \( -\nabla f(x_0) \) “points” to the direction along which the rate of decrease of \( f \) at \( x_0 \) is maximum.

**Proof.** Let \( h \) be a vector with unit Euclidean norm in \( \mathbb{R}^n \). The rate of change of \( f \) at the point \( x_0 \) and in the direction \( h \) equals \( \nabla f(x_0)^T h \). Then

\[
\nabla f(x_0)^T h = \|\nabla f(x_0)\|_2 \cos \angle (\nabla f(x_0), h)
\]

(1.28)

and

\[
|\nabla f(x_0)^T h| = \|\nabla f(x_0)\|_2 |\cos \angle (\nabla f(x_0), h)| \leq \|\nabla f(x_0)\|_2,
\]

(1.29)

from which we obtain that

\[
-\|\nabla f(x_0)\|_2 \leq \nabla f(x_0)^T h \leq \|\nabla f(x_0)\|_2.
\]

(1.30)

The left inequality holds as equality for \( h = -\frac{1}{\|\nabla f(x_0)\|_2} \nabla f(x_0) \), while the right inequality holds as equality for \( h = \frac{1}{\|\nabla f(x_0)\|_2} \nabla f(x_0) \). \( \square \)

### 1.2.8 Gradient and function level sets

In this subsection, we prove another important result for the gradient.

We repeat that a function \( c : [a, b] \subseteq \mathbb{R} \rightarrow \mathbb{R}^m \) is a curve. The curve \( c(t) \) can be expressed as

\[
c(t) = \begin{bmatrix} c_1(t) \\ \vdots \\ c_m(t) \end{bmatrix},
\]

(1.31)
with $c_i : [a, b] \to \mathbb{R}$, for $i = 1, \ldots, m$. As we have seen, if $c$ is differentiable at $t \in [a, b]$, then its derivative is the $m$-dimensional vector
\begin{equation}
Dc(t) = \lim_{h \to 0} \frac{c(t + h) - c(t)}{h}.
\end{equation}

The $i$th element of $Dc(t)$ is equal to the derivative of $c_i$, for $i = 1, \ldots, m$, and $Dc(t)$ is parallel to the straight line which is tangent to the curve $c$ at the point $c(t)$.

**Theorem 1.2.7.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable real function and $x_0$ a point of the level set $S_c = \{x \in \mathbb{R}^n | f(x) = c\}$, with $\nabla f(x_0) \neq 0$. Then, the gradient vector $\nabla f(x_0)$ is orthogonal to the level set in the following sense (see Figure 1.9). If $c(t)$ is a smooth curve contained in the set $S_c$, with $c(0) = x_0$, and $v$ vector parallel to the line tangent to $c(t)$ at the point $x_0$, for example, $v = Dc(0)$, then
\begin{equation}
\nabla f(x_0)^T v = 0.
\end{equation}

**Proof.** Since the curve $c(t)$ is contained in the set $S_c$, we have that $f(c(t)) = c$. Moreover, $c(0) = x_0$. From the chain rule, we have that
\begin{align*}
0 &= \frac{d}{dt} f(c(t)) \bigg|_{t=0} = Df(c(t))Dc(t) \bigg|_{t=0} \\
&= Df(c(0))Dc(0) \\
&= Df(x_0)v \\
&= \nabla f(x_0)^T v,
\end{align*}
completing the proof.

As a consequence of the Theorem 1.2.7, it is reasonable to define the plane which is tangent to the level set of $f$, $S_c$, at point $x_0 \in S_c$, as follows.
1.3. APPENDIX

**Definition 1.2.11.** Let \( f : \mathbb{R}^n \to \mathbb{R} \) be differentiable and \( S_c = \{ x \in \mathbb{R}^n \mid f(x) = c \} \). Then, the plane tangent to \( S_c \) at the point \( x_0 \in S_c \), with \( \nabla f(x_0) \neq 0 \), is defined as

\[
P^f_{x_0} = \{ x \in \mathbb{R}^n \mid \nabla f(x_0)^T (x - x_0) = 0 \}.
\] (1.35)

Obviously, \( P^f_{x_0} \) contains the point \( x_0 \). If \( x \neq x_0 \) belongs to the tangent plane, then the vector \( x - x_0 \) is parallel to the tangent plane. But, we have proved that \( \nabla f(x_0) \) is perpendicular to every tangent vector of any curve which lies on the surface \( S_c \) and passes through the point \( x_0 \). Therefore, the definition (1.35) is satisfactory.

### 1.2.9 Taylor expansions

Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a smooth real function. Then, the following extremely important relationships hold true:

\[
f(x) = f(x_0) + \nabla f(x_0)^T (x - x_0) + O(\|x - x_0\|^2)
\]

\[
= f(x_0) + \nabla f(x_0)^T (x - x_0) + \frac{1}{2} (x - x_0)^T \nabla^2 f(x_0) (x - x_0) + O(\|x - x_0\|^3),
\] (1.36)

which we call first- and second-order Taylor expansions, respectively.

If, in the above expressions, we ignore the terms denoted by \( O(\cdot) \), then we have the first- and second-order Taylor approximations, respectively.

Moreover, it can be shown that (see any good book on Calculus of Several Variables)

\[
f(x) = f(x_0) + \nabla f(z)^T (x - x_0)
\] (1.37)

for some \( z = \theta x + (1 - \theta)x_0 \), with \( 0 \leq \theta \leq 1 \), and

\[
f(x) = f(x_0) + \nabla f(x_0)^T (x - x_0) + \frac{1}{2} (x - x_0)^T \nabla^2 f(w) (x - x_0),
\] (1.38)

for some \( w = \theta x + (1 - \theta)x_0 \), with \( 0 \leq \theta \leq 1 \).

### 1.3 Appendix

In this appendix, we use first-order Taylor expansions and give a proof of a special form of the chain rule.
Let $c : \mathbb{R} \to \mathbb{R}^n$ be a smooth curve and $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth real function. We define $g : \mathbb{R} \to \mathbb{R}$, with $g(t) = (f \circ c)(t) = f(c(t))$. Suppose we want to calculate the derivative $Dg(t) = D(f \circ c)(t)$. We remind you that

$$c(t + \Delta t) = c(t) + Dc(t)\Delta t + O(\Delta t^2),$$

(1.39)

$$f(x + \Delta x) = f(x) + Df(x)\Delta x + O(\|\Delta x\|^2)$$

(1.40)

Then

$$Dg(t) = \lim_{\Delta t \to 0} \frac{g(t + \Delta t) - g(t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{f(c(t + \Delta t)) - f(c(t))}{\Delta t} = \lim_{\Delta t \to 0} \frac{f(c(t) + Dc(t)\Delta t + O(\Delta t^2)) - f(c(t))}{\Delta t} = \lim_{\Delta t \to 0} \frac{f(c(t) + Df(c(t))Dc(t)\Delta t + O(\Delta t^2) - f(c(t))}{\Delta t}$$

(1.41)

$$= Df(c(t))Dc(t).$$

### 1.4 Brief overview of technical proofs

**Statements** are denoted by capital letters, eg, $P$, $Q$, and will be true or false. For example, the sentence

$$P = \text{“2 is an even number.”}$$

is true.

If $P$ is a proposition, then the **negation** of $P$ is a proposition, denoted by not $P$, and is true if $P$ is false and false if $P$ is true (see the truth table below).

<table>
<thead>
<tr>
<th>$P$</th>
<th>not $P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

An important way of constructing sentences is by implication. Symbolically, we have

$$(P \Rightarrow Q) \quad (\text{If } P, \text{ then } Q) \quad (P \text{ implies } Q).$$

(1.42)

The proposition $P$ is called **hypothesis** and the proposition $Q$ is called **conclusion**.

If the implication holds, we say that $P$ is sufficient for $Q$ and $Q$ is necessary for $P$. The truth table of the proposition $P \Rightarrow Q$ is as follows:
That is, if the hypothesis \( P \) is false, then the implication is true regardless of whether the conclusion \( Q \) is true or false (this, at first, might seem a bit strange). If the hypothesis \( P \) is true and the conclusion \( Q \) false, then the implication is false, while if both the hypothesis and the conclusion are true, then the implication is true.

The proposition \( Q \Rightarrow P \) is called the **inverse** of \( P \Rightarrow Q \) while the proposition \((\text{not } Q) \Rightarrow (\text{not } P)\) is called the **contrapositive** of \( P \Rightarrow Q \). The corresponding truth tables are as follows:

<table>
<thead>
<tr>
<th>( P )</th>
<th>( Q )</th>
<th>( P \Rightarrow Q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

We observe that the statement \( P \Rightarrow Q \) is logically equivalent (has the same truth table) with the proposition \((\text{not } Q) \Rightarrow (\text{not } P)\). Another proposition equivalent to \( P \Rightarrow Q \) is \( \text{not } (P \land (\text{not } Q)) \).

According to the above, in order to prove a proposition of the form \( P \Rightarrow Q \), we have the following basic options:

1. **Direct proof**: We start from the assumption \( P \) and, after “correct reasonings,” we arrive at conclusion \( Q \).

2. **Proof via contraposition**: We start from the hypothesis \((\text{not } Q)\) and, after “correct reasonings,” we conclude \((\text{not } P)\).

3. **Proof via reductio ad absurdum**: We start from the hypothesis \( P \) and \((\text{not } Q)\) and, after “correct reasonings,” we arrive at an assertion which is false.

1.4.1 Quantifiers

Important symbols that we use to construct sentences are the following: \( \forall \) and \( \exists \). The first one denotes the universal quantifier while the second denotes the existential quantifier.
For example, consider the statements

$$\forall x \exists y (x + y = 0)$$

and

$$\exists y \forall x (x + y = 0).$$

The first statement is true, while the second is false. Therefore, the “order” of the quantifiers matters.

The negation of simple sentences with quantifiers is constructed as follows:

$$\neg (\forall x P) \equiv (\exists x) \neg P,$$

$$\neg (\exists x P) \equiv (\forall x) \neg P.$$ 

If the sentences are more complex, we work as follows:

$$\neg [(\forall x)(\exists y) (x + y = 0)]$$

$$\equiv (\exists x) \neg [(\exists y) (x + y = 0)]$$

$$\equiv (\exists x) (\forall y) \neg (x + y = 0)$$

$$\equiv (\exists x) (\forall y) (x + y \neq 0).$$

The last proposition asserts that there exists $x$ such that for every $y$ the sum $x + y$ is nonzero. This proposition is false.
Chapter 2

Elements of the theory of convex sets

2.1 Affine sets

Definition 2.1.1. A set $C \subseteq \mathbb{R}^n$ is called affine if the points of the lines joining any two points of $C$ belong to $C$. That is, if $x, y \in C$ and $\theta \in \mathbb{R}$, then $\theta x + (1 - \theta)y \in C$.

In other words, the set $C$ is called affine if it contains every linear combination of any two of its points, under the condition that the coefficients of the linear combination sum to one.

Example 2.1.1. Prove that the solution set of the equation $Ax = b$ is an affine set.

The proof is as follows. If the system has no solutions, then the solution set is $\emptyset$, which is an affine set.

![Figure 2.1: The affine set $\theta x_1 + (1 - \theta)x_2$.](image-url)
Let \( x_1 \) and \( x_2 \) be solutions of the equation \( Ax = b \) and \( \theta \in \mathbb{R} \). Then
\[
Ax_1 = b, \quad Ax_2 = b \\
\Rightarrow A(\theta x_1) = \theta b, \quad A((1 - \theta) x_2) = (1 - \theta) b \\
\Rightarrow A(\theta x_1 + (1 - \theta)x_2) = b.
\]
Therefore, the point \( \theta x_1 + (1 - \theta)x_2 \) is a solution of the equation \( Ax = b \). So, the solution set is affine. \( \square \)

**Exercise:** Let \( x_1, x_2 \in \mathbb{R}^n \). Find the set \( \text{aff}\{x_1, x_2\} \).

**Definition 2.1.2.** If \( x_1, \ldots, x_k \in \mathbb{R}^n, \theta_1, \ldots, \theta_k \in \mathbb{R}, \) with
\[
\theta_1 + \cdots + \theta_k = 1,
\]
then every expression of the form \( \theta_1 x_1 + \cdots + \theta_k x_k \) is called an affine combination of \( x_1, \ldots, x_k \).

Using induction and the definition 2.1.1, it can be shown that an affine set contains every affine combination of its elements. That is, if \( C \) is an affine set, \( k \in \mathbb{N}, x_1, \ldots, x_k \in C, \) and \( \theta_1 + \cdots + \theta_k = 1, \) then \( \theta_1 x_1 + \cdots + \theta_k x_k \in C \) (prove it).

**Theorem 2.1.1.** If \( C \) is an affine set and \( x_0 \in C \), then the set
\[
V = C - x_0 = \{x - x_0 \mid x \in C\}
\]
is a linear subspace.
2.1. AFFINE SETS

Proof. We must prove that if \( v_1, v_2 \in V \) and \( a_1, a_2 \in \mathbb{R} \), then \( a_1 v_1 + a_2 v_2 \in V \).

Suppose that \( v_1, v_2 \in V \). Then, there exist \( x_1, x_2 \in C \) such that \( v_1 = x_1 - x_0 \) and \( v_2 = x_2 - x_0 \). Let \( a_1, a_2 \in \mathbb{R} \). Then

\[
\begin{align*}
    a_1 v_1 + a_2 v_2 &= a_1 (x_1 - x_0) + a_2 (x_2 - x_0) \\
    &= a_1 x_1 + a_2 x_2 - (a_1 + a_2) x_0 \\
    &= a_1 x_1 + a_2 x_2 - (a_1 + a_2) x_0 + x_0 - x_0 \\
    &= (a_1 x_1 + a_2 x_2 + (1 - a_1 - a_2) x_0) - x_0 \\
    &= x_* - x_0, \quad (2.3)
\end{align*}
\]

where \( x_* := a_1 x_1 + a_2 x_2 + (1 - a_1 - a_2) x_0 \in C \) (why?). So, any linear combination of \( v_1 \) and \( v_2 \) can be expressed as \( x_* - x_0 \), with \( x_* \in C \) and, thus, belongs to \( V \).

Similarly, it can be shown that every affine set \( C \) can be expressed as a translation of a linear subspace \( V \), that is

\[
C = V + x_0 = \{ v + x_0 | v \in V \}, \quad (2.4)
\]

with \( x_0 \in C \). Vector \( x_0 \) in the above definitions is not unique. Actually, it might be any point of \( C \).

The dimension of \( C \) is defined as the dimension of \( V \).

**Definition 2.1.3.** The set of the affine combinations of the elements of a set \( C \subseteq \mathbb{R}^n \) is called the **affine hull** of \( C \) and is denoted as \( \text{aff } C \).

That is,

\[
\text{aff } C = \{ \theta_1 x_1 + \cdots + \theta_k x_k | x_1, \ldots, x_k \in C, \theta_1 + \cdots + \theta_k = 1 \}. \quad (2.5)
\]

The affine hull of a set \( C \) is the smallest affine set that contains \( C \), in the following sense: if \( S \) is an affine set with \( C \subseteq S \), then \( \text{aff } C \subseteq S \).

2.1.1 Affine Dimension and Relative Interior

The **affine dimension** of a set \( C \subseteq \mathbb{R}^n \) is the dimension of its affine hull.

If the affine dimension of a set \( C \subseteq \mathbb{R}^n \) is less than \( n \), then \( C \) lies in the affine set \( \text{aff } C \neq \mathbb{R}^n \).

We define as **relative interior** of the set \( C \) the set

\[
\text{relint } C = \{ x \in C | B(x, r) \cap \text{aff } C \subseteq C, \text{ for some } r > 0 \}, \quad (2.6)
\]
where
\[ B(x, r) = \{ y \in \mathbb{R}^n \mid \|y - x\| \leq r \}, \]
and as relative boundary the set \( \text{cl } C \setminus \text{relint } C \).

**Example 2.1.2.** Consider the square
\[ C = \{ x \in \mathbb{R}^3 \mid -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1, x_3 = 0 \}. \]
The affine hull of \( C \) is the set \( \text{aff } C = \{ x \in \mathbb{R}^3 \mid x_3 = 0 \} \). The interior of \( C \) is empty (why?) while its boundary is set \( C \) itself. But the relative interior of \( C \) is the set
\[ \text{relint } C = \{ x \in \mathbb{R}^3 \mid -1 < x_1 < 1, -1 < x_2 < 1, x_3 = 0 \}; \]
and its relative boundary is the set
\[ \{ x \in \mathbb{R}^3 \mid \max\{|x_1|, |x_2|\} = 1, x_3 = 0 \}. \]

### 2.2 Convex sets

**Definition 2.2.1.** A set \( C \subseteq \mathbb{R}^n \) is called convex if the points of the line segments connecting any two points of \( C \) belong to \( C \). That is, if, for every \( x_1, x_2 \in C \) and \( 0 \leq \theta \leq 1 \), we have that
\[ \theta x_1 + (1 - \theta) x_2 \in C. \tag{2.7} \]

We can easily see that every affine set is convex.

**Definition 2.2.2.** Let \( k \in \mathbb{N}, x_1, \ldots, x_k \in \mathbb{R}^n \), and \( \theta_1, \ldots, \theta_k \in \mathbb{R} \), with \( \theta_i \geq 0 \), for \( i = 1, \ldots, k \), and \( \theta_1 + \cdots + \theta_k = 1 \). The point \( \theta_1 x_1 + \cdots + \theta_k x_k \) is called a convex combination of \( x_1, \ldots, x_k \).
2.3. CONES

Theorem 2.2.1. A set \( C \) is convex if, and only if, it contains all convex combinations of its points, that is, if \( k \in \mathbb{N}, x_1, \ldots, x_k \in C, \theta_1, \ldots, \theta_k \in \mathbb{R}, \) with \( \theta_i \geq 0, \) for \( i = 1, \ldots, k, \) and \( \theta_1 + \cdots + \theta_k = 1, \) then \( \theta_1 x_1 + \cdots + \theta_k x_k \in C. \)

Proof. (Inverse) The proof of the inverse is simple and we will start with it. Since \( C \) contains the convex combinations of its elements for any \( k \in \mathbb{N}, \) it also contains the convex combinations of its elements for \( k = 2. \) Therefore, it is a convex set.

(Direct) The proof will be done by induction.

For \( k = 2, \) the theorem holds by the definition of convex set.

We assume that the theorem holds for \( k = n. \)

Let \( k = n + 1, x_i \in C, \theta_i \in \mathbb{R}, \theta_i \geq 0, \) for \( i = 1, \ldots, n + 1, \) with \( \sum_{i=1}^{n+1} \theta_i = 1 \) and \( x = \sum_{i=1}^{n+1} \theta_i x_i. \) We assume that \( \theta_{n+1} \neq 0 \) and define \( \theta := \sum_{i=1}^{n} \theta_i. \) Then, \( \theta_{n+1} = 1 - \theta \) and

\[
x = \sum_{i=1}^{n+1} \theta_i x_i = \sum_{i=1}^{n} \theta_i x_i + \theta_{n+1} x_{n+1} = \theta \left( \sum_{i=1}^{n} \frac{\theta_i}{\theta} x_i \right) + \theta_{n+1} x_{n+1}.
\]

If we define \( \theta'_i := \frac{\theta_i}{\theta}, \) for \( i = 1, \ldots, n, \) we observe that the expression \( \sum_{i=1}^{n} \theta'_i x_i \) is a convex combination of \( n \) elements of \( C. \) We define \( x' = \sum_{i=1}^{n} \theta'_i x_i. \) From the induction hypothesis, we conclude that \( x' \in C. \) Finally, we find that \( x \in C \) because it can be expressed as a convex combination of two points of \( C, \) as follows

\[
x = \theta x' + (1 - \theta)x_{n+1}.
\]

Therefore, the proposition holds for \( k = n + 1 \) and the proof is complete.

Definition 2.2.3. The convex hull of a set \( C \) is the set of all convex combinations of its elements and is denoted as \( \text{conv} \ C. \) That is,

\[
\text{conv} \ C = \{ \theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C, \theta_i \geq 0, i = 1, \ldots, k, \theta_1 + \cdots + \theta_k = 1 \}.
\]

It can be shown that the set \( \text{conv} \ C \) is the smallest convex set containing \( C \) (prove it).

Exercise: Let \( x_1, x_2 \in \mathbb{R}^n. \) Find the set \( \text{conv}\{x_1, x_2\}. \)

2.3 Cones

Definition 2.3.1. A set \( C \subseteq \mathbb{R}^n \) is called cone if for every \( x \in C \) and \( \theta \geq 0, \) we have that \( \theta x \in C. \)
CHAPTER 2. ELEMENTS OF THE THEORY OF CONVEX SETS

Figure 2.4: Convex cone (a), non-convex cone (b).

Notice that if $C \neq \emptyset$, then $0 \in C$ (why?).

A set $C$ is a convex cone if it is, at the same time, convex set and cone (see Figure 2.4).

**Exercise:** Let $C$ be a convex cone. Prove that, if $x_1, x_2 \in C$ and $\theta_1, \theta_2 \geq 0$, then

$$\theta_1 x_1 + \theta_2 x_2 \in C. \quad (2.9)$$

**Proof:** Let $x_1, x_2 \in C$ and $\theta_1, \theta_2 \geq 0$. If $\theta_1 = \theta_2 = 0$, then $\theta_1 x_1 + \theta_2 x_2 = 0 \in C$.

Let $\theta_1 \neq 0$ or $\theta_2 \neq 0$. Then, due to the convexity of $C$, we have that

$$\frac{\theta_1}{\theta_1 + \theta_2} x_1 + \frac{\theta_2}{\theta_1 + \theta_2} x_2 \in C. \quad (2.10)$$

Moreover, $C$ is a cone, therefore,

$$(\theta_1 + \theta_2) \left( \frac{\theta_1}{\theta_1 + \theta_2} x_1 + \frac{\theta_2}{\theta_1 + \theta_2} x_2 \right) \in C. \quad (2.11)$$

That is, $\theta_1 x_1 + \theta_2 x_2 \in C$. $\square$

**Definition 2.3.2.** Let $k \in \mathbb{N}$, $x_1, \ldots, x_k \in \mathbb{R}^n$, and $\theta_1, \ldots, \theta_k \in \mathbb{R}$, with $\theta_i \geq 0$, for $i = 1, \ldots, k$. The vectors of the form

$$\theta_1 x_1 + \cdots + \theta_k x_k$$

are called **conic combinations** of $x_1, \ldots, x_k$.

A set $C$ is a convex cone if, and only if, it contains all conic combinations of its elements.

**Definition 2.3.3.** The **conic hull** of a set $C$ is the set of the conic combinations of its elements, that is, the set

$$\{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C, \theta_i \geq 0, i = 1, \ldots, k\}. \quad (2.12)$$

It can be shown that the conic hull of a set $C$ is the smallest convex cone which contains $C$. 
2.4 Examples of Convex Sets

$\emptyset$, the singleton $\{x_0\}$, and $\mathbb{R}^n$ are affine and, thus, convex sets.

Straight lines are affine and, therefore, convex sets, while line segments are convex but not affine sets.

2.5 Hyperplanes - Halfspaces

Let $\mathbf{0} \neq \mathbf{a} \in \mathbb{R}^n$. A hyperplane in $\mathbb{R}^n$ is a set of the form

$$\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{a}^T \mathbf{x} - b = 0\}. \quad (2.13)$$

A hyperplane divides $\mathbb{R}^n$ into two half-spaces, as follows:

$$\mathcal{P}^+ = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{a}^T \mathbf{x} - b \geq 0\}, \quad \mathcal{P}^- = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{a}^T \mathbf{x} - b \leq 0\}. \quad (2.14)$$

The set $\mathcal{P}$ is affine and, therefore, convex. The sets $\mathcal{P}^+$ and $\mathcal{P}^-$ are convex but not affine (prove it).

2.5.1 Euclidean balls

The Euclidean ball on $\mathbb{R}^n$, with center $\mathbf{x}_c$ and radius $r$, is defined as

$$\mathcal{B}(\mathbf{x}_c, r) = \{\mathbf{x} \in \mathbb{R}^n | \|\mathbf{x} - \mathbf{x}_c\|_2 \leq r\}. \quad (2.15)$$

Another representation of the Euclidean ball is as follows:

$$\mathcal{B}(\mathbf{x}_c, r) = \{\mathbf{x}_c + r\mathbf{x} | \|\mathbf{x}\|_2 \leq 1\}. \quad (2.16)$$
Euclidean balls are convex sets. The proof is as follows. Let \( x_1, x_2 \in B(x_c, r) \). This means that \( \|x_i - x_c\|_2 \leq r \), for \( i = 1, 2 \). We must prove that, for \( 0 \leq \theta \leq 1 \), we have that
\[
x := \theta x_1 + (1 - \theta)x_2 \in B(x_c, r),
\]
that is, \( \|x - x_c\|_2 \leq r \).

We proceed as follows:
\[
\|x - x_c\|_2 = \|\theta x_1 + (1 - \theta)x_2 - x_c\|_2 \\
= \|\theta(x_1 - x_c) + (1 - \theta)(x_2 - x_c)\|_2 \\
\leq \|\theta(x_1 - x_c)\|_2 + \|(1 - \theta)(x_2 - x_c)\|_2 \\
= \theta \|x_1 - x_c\|_2 + (1 - \theta) \|x_2 - x_c\|_2 \\
\leq \theta r + (1 - \theta) r \\
= r,
\]
therefore, \( x \in B(x_c, r) \).

### 2.5.2 Ellipsoids

If \( P = P^T \succ 0 \) and \( x_c \in \mathbb{R}^n \), then the set
\[
E(x_c, P) = \{ x \in \mathbb{R}^n \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1 \}
\]
is called **ellipsoid** with center \( x_c \).

An alternative description of \( E \) is as follows:
\[
E(x_c, A) = \{ x_c + Au \mid \|u\|_2 \leq 1 \},
\]
with \( A = P^{1/2} \).
Ellipsoids are convex sets. Next, we will give a proof for the case where \( \mathbf{x}_c = \mathbf{0} \) (the proof generalizes to every \( \mathbf{x}_c \)).

Let \( \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{E}(\mathbf{0}, \mathbf{P}) \), that is, \( \mathbf{x}_i^T \mathbf{P}^{-1} \mathbf{x}_i \leq 1 \), for \( i = 1,2 \). We must prove that, if \( 0 \leq \theta \leq 1 \) and \( \mathbf{x} := \theta \mathbf{x}_1 + (1 - \theta) \mathbf{x}_2 \), then \( \mathbf{x}^T \mathbf{P}^{-1} \mathbf{x} \leq 1 \). First, we have

\[
\mathbf{x}^T \mathbf{P}^{-1} \mathbf{x} = (\theta \mathbf{x}_1 + (1 - \theta) \mathbf{x}_2)^T \mathbf{P}^{-1} (\theta \mathbf{x}_1 + (1 - \theta) \mathbf{x}_2)
\]

\[
= \theta^2 \mathbf{x}_1^T \mathbf{P}^{-1} \mathbf{x}_1 + (1 - \theta)^2 \mathbf{x}_2^T \mathbf{P}^{-1} \mathbf{x}_2 + 2 \theta (1 - \theta) \mathbf{x}_1^T \mathbf{P}^{-1} \mathbf{x}_2
\]

\[
\leq \theta^2 + (1 - \theta)^2 + 2 \theta (1 - \theta) \mathbf{x}_1^T \mathbf{P}^{-1} \mathbf{x}_2. \tag{2.20}
\]

If we define \( \mathbf{y}_i := \mathbf{P}^{-\frac{1}{2}} \mathbf{x}_i \), for \( i = 1,2 \), then \( \|\mathbf{y}_i\|^2 = \mathbf{x}_i^T \mathbf{P}^{-1} \mathbf{x}_i \leq 1 \), for \( i = 1,2 \), and

\[
\mathbf{x}_1^T \mathbf{P}^{-1} \mathbf{x}_2 = \left( \mathbf{P}^{-\frac{1}{2}} \mathbf{x}_1 \right)^T \left( \mathbf{P}^{-\frac{1}{2}} \mathbf{x}_2 \right)
\]

\[
= \mathbf{y}_1^T \mathbf{y}_2 \tag{2.21}
\]

\[
\leq \|\mathbf{y}_1\| \|\mathbf{y}_2\| \]

\[
\leq 1,
\]

where at point (a) we used the Cauchy-Schwarz inequality. Since \( 0 \leq \theta \leq 1 \), we have that \( \theta (1 - \theta) \geq 0 \), and

\[
\theta (1 - \theta) \mathbf{x}_1^T \mathbf{P}^{-1} \mathbf{x}_2 \leq \theta (1 - \theta). \tag{2.22}
\]

Therefore,

\[
\mathbf{x}^T \mathbf{P}^{-1} \mathbf{x} \leq \theta^2 + (1 - \theta)^2 + 2 \theta (1 - \theta) = 1. \tag{2.23}
\]

Thus, the point \( \mathbf{x} \) belongs to the ellipsoid \( \mathbb{E}(\mathbf{0}, \mathbf{P}) \) and the set \( \mathbb{E}(\mathbf{0}, \mathbf{P}) \) is convex.
2.5.3 Polyhedra

We define a **polyhedron** as the solution of a finite set of linear inequalities and linear equalities (see Figure 2.7). For example, the set

$$\mathcal{D} = \{ x \in \mathbb{R}^n \mid a_j^T x - b_j \leq 0, j = 1, \ldots, m, \ c_j^T x - d_j = 0, j = 1, \ldots, p \}$$  \hspace{1cm} (2.24)

is a polyhedron.

Thus, a polyhedron is the intersection of a finite number of half-spaces and hyperplanes. Affine sets, line segments, and half-spaces are polyhedra.

**Exercise:** Prove that polyhedra are convex sets.

2.6 Set operations that preserve convexity

2.6.1 Intersection

If $S_1, S_2 \subseteq \mathbb{R}^n$ are convex sets, then their intersection $S_1 \cap S_2$ is a convex set. The above proposition generalizes to any number (finite or infinite) of convex sets.

More specifically, if $i \in I$, where $I$ is any set of indices, and $S_i$ is convex set, for each $i \in I$, then the set

$$S = \cap_{i \in I} S_i \hspace{1cm} (2.25)$$

is convex. The proof is as follows.

Let $x_1, x_2 \in S$. This means that $x_1, x_2 \in S_i$, for $i \in I$. If $0 \leq \theta \leq 1$, then, due to the convexity of $S_i$, for $i \in I$, we will have that $x := \theta x_1 + (1 - \theta) x_2 \in S_i$, for $i \in I$. Therefore, $x \in S$, proving that $S$ is a convex set.

2.6.2 Cartesian product

**Theorem 2.6.1.** If $S_1 \subseteq \mathbb{R}^n$ and $S_2 \subseteq \mathbb{R}^m$ are convex sets, then the Cartesian product $S = S_1 \times S_2 \subseteq \mathbb{R}^{n+m}$ is convex set.

**Proof.** We must prove that, if $x_1, x_2 \in S$ and $0 \leq \theta \leq 1$, then $\theta x_1 + (1 - \theta) x_2 \in S$.

Let $x_1, x_2 \in S$. Then, there exist $x_{11}, x_{12} \in S_1$ and $x_{21}, x_{22} \in S_2$ such that

$$x_1 = \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix}, \quad x_2 = \begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix}.$$
Since the sets $S_1$ and $S_2$ are convex, for each $0 \leq \theta \leq 1$, $\theta x_{11} + (1 - \theta)x_{12} \in S_1$ and $\theta x_{21} + (1 - \theta)x_{22} \in S_2$. Therefore,

$$\begin{bmatrix} \theta x_{11} + (1 - \theta)x_{12} \\ \theta x_{21} + (1 - \theta)x_{22} \end{bmatrix} \in S_1 \times S_2$$

$$\Rightarrow \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} + (1 - \theta) \begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix} \in S_1 \times S_2 \quad (2.26)$$

$$\Rightarrow \theta x_1 + (1 - \theta)x_2 \in S,$$

completing the proof. \qed

### 2.6.3 Image of affine function

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called affine if it is the sum of a linear function and a constant, that is, if it is of the form $f(x) = Ax + b$, with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

**Theorem 2.6.2.** Let $S \subseteq \mathbb{R}^n$ be a convex set and $f : \mathbb{R}^n \to \mathbb{R}^m$ an affine function. Then, the image of $S$ under $f$,

$$f(S) = \{ f(x) | x \in S \}, \quad (2.27)$$

is a convex set.

**Proof.** We must prove that, if $y_1, y_2 \in f(S)$ and $0 \leq \theta \leq 1$, then $y = \theta y_1 + (1 - \theta)y_2 \in f(S)$.

Let $y_1, y_2 \in f(S)$. This means that there exist $x_1, x_2 \in S$ such that $y_i = Ax_i + b$, for $i = 1, 2$.

Due to the convexity of $S$, we have that, for $0 \leq \theta \leq 1$, $x = \theta x_1 + (1 - \theta)x_2 \in S$. The image of $x$ is given by the relation

$$f(x) = Ax + b$$

$$= A (\theta x_1 + (1 - \theta)x_2) + b$$

$$= \theta (Ax_1 + b) + (1 - \theta) (Ax_2 + b) \quad (2.28)$$

$$= \theta y_1 + (1 - \theta)y_2$$

$$= y.$$

Therefore, $y = f(x) \in f(S)$ and the proof is complete. \qed

**Corollary 2.6.1.** If $f : \mathbb{R}^k \to \mathbb{R}^n$ is an affine function and $S \subseteq \mathbb{R}^n$ is a convex set, then the inverse image of $S$ under $f$, i.e., the set

$$f^{-1}(S) = \{ x | f(x) \in S \}, \quad (2.29)$$
is a convex set.

\textit{Proof.} The proof is based on the fact that the inverse of an affine function is an affine function. For example, if \( m > n \) and the columns of \( A \) are linearly independent, then if \( y = f(x) = Ax + b \), the inverse function is given by the relation

\[
x = f^{-1}(y) = (A^T A)^{-1} A^T (y - b)
= (A^T A)^{-1} A^T y - (A^T b f A)^{-1} A^T b
\]

(2.30)

with \( A' := (A^T A)^{-1} A^T \) and \( b' := - (A^T A)^{-1} A^T b \), which is an affine function. \qed

Next, we will present some important applications of the Theorem 2.6.2.

\section*{2.6.4 Set Scaling}

\textbf{Corollary 2.6.2.} If \( S \subseteq \mathbb{R}^n \) is a convex set and \( a \in \mathbb{R} \), then the set \( a S = \{ ax \mid x \in S \} \) is convex.

\textit{Proof.} This proposition is a consequence of Theorem 2.6.1, for the affine (actually, linear) function \( f : S \to \mathbb{R}^n \) with \( f(x) = ax \). \qed

\section*{2.6.5 Set translation}

\textbf{Corollary 2.6.3.} If \( S \subseteq \mathbb{R}^n \) is a convex set and \( a \in \mathbb{R}^n \), then the set \( S + a = \{ x + a \mid x \in S \} \) is convex.

\textit{Proof.} This proposition is a consequence of Theorem 2.6.1, for the affine function \( f : S \to \mathbb{R}^n \) with \( f(x) = x + a \). \qed

\section*{2.6.6 Projection}

\textbf{Corollary 2.6.4.} The projection of a convex set onto some of its coordinates results in a convex set. More specifically, if \( S \subseteq \mathbb{R}^{n+m} \) is a convex set, then the set

\[
T = \{ x_1 \in \mathbb{R}^n \mid (x_1, x_2) \in S \},
\]

(2.31)

is convex.

\textit{Proof.} The proof is left to the reader. \qed
2.6.7 Sum of sets

If $S_1, S_2 \subseteq \mathbb{R}^n$, then their sum is defined as

$$S_1 + S_2 = \{x_1 + x_2 \mid x_1 \in S_1, x_2 \in S_2\}. \quad (2.32)$$

**Corollary 2.6.5.** If $S_1$ and $S_2$ are convex, then $S_1 + S_2$ is convex.

**Proof.** One way to prove this corollary is to observe the following:

1. $S_1 \times S_2$ is convex (Cartesian product of convex sets),

2. the image of the convex set $S_1 \times S_2$ under the affine transformation $f : S_1 \times S_2 \to \mathbb{R}^n$
   defined as

   $$f(x_1, x_2) = [I_n, I_n] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 + x_2 \quad (2.33)$$

   is the sum $S_1 + S_2$.

2.7 Separating hyperplanes

**Theorem 2.7.1.** Let $C, D \subseteq \mathbb{R}^n$ be convex sets that do not intersect, that is, $C \cap D = \emptyset$.

Then, there exist $0 \neq a \in \mathbb{R}^n$ and $b \in \mathbb{R}$, such that $a^T x \geq b$ for every $x \in C$ and $a^T x \leq b$ for every $x \in D$ (see Figure 2.9).
CHAPTER 2. ELEMENTS OF THE THEORY OF CONVEX SETS

The hyperplane \( \{ x \mid a^T x = b \} \) is called **separating hyperplane** of \( C \) and \( D \).

**Proof.** See pages 46–48 of the book by Boyd and Vandenberghe. \( \square \)

### 2.8 Supporting hyperplanes

**Theorem 2.8.1.** Let \( C \subseteq \mathbb{R}^n \) be a nonempty convex set and \( x_0 \in \text{bd } C \). Then, there exists \( 0 \neq a \in \mathbb{R}^n \) such that \( a^T x \leq a^T x_0 \), for every \( x \in C \).

**Proof.** See page 51 of Boyd and Vandenberghe. \( \square \)

Theorem 2.8.1 basically says that the point \( x_0 \) and the set \( C \) are separated by the hyperplane \( P = \{ x \in \mathbb{R}^n \mid a^T x = a^T x_0 \} \).
Chapter 3

Convex functions

In this Chapter, we shall briefly consider significant properties of convex functions.

3.1 Convex Functions

**Definition 3.1.1.** A function $f : \text{dom} f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is called **convex** if the set $\text{dom} f$ is convex and if, for each $x, y \in \text{dom} f$ and $0 \leq \theta \leq 1$, it holds that

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

(3.1)

Geometrically, the inequality (3.1) implies that no point of the line segment connecting the points $(x, f(x))$ and $(y, f(y))$ lies under the graph of $f$ (see Fig. 3.1).

**Definition 3.1.2.** A function $f : \text{dom} f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is called **strictly convex** if the set $\text{dom} f$ is convex if, for every $x, y \in \text{dom} f$, with $x \neq y$, and $0 < \theta < 1$, it holds that

$$f(\theta x + (1 - \theta)y) < \theta f(x) + (1 - \theta)f(y).$$

(3.2)

Geometrically, the inequality (3.2) means that, with the exception of the points $(x, f(x))$ and $(y, f(y))$, the line segment connecting the points $(y, f(y))$ $(x, f(x))$ and $(y, f(y))$ lies above the graph of $f$ (see Fig. 3.1).

The function $f$ is called **concave** if $-f$ is convex.

An affine function is, at the same time, convex and concave.
The relation (3.1) is generalized as follows. Let \( f : \text{dom} f \subseteq \mathbb{R}^n \to \mathbb{R} \) be a convex function, \( x_i \in \text{dom} f \), for \( i = 1, \ldots, n \), and \( a_i \geq 0 \), for \( i = 1, \ldots, n \), with \( a_1 + \cdots + a_n = 1 \). Then

\[
f \left( \sum_{i=1}^{n} a_i x_i \right) \leq \sum_{i=1}^{n} a_i f(x_i). \tag{3.3}
\]

It can be shown that a function \( f \) is convex if, and only if, its restriction on the intersection of any straight line in \( \mathbb{R}^n \) and its domain \( \text{dom} f \) is a convex function.

**Theorem 3.1.1.** The function \( f : \text{dom} f \subseteq \mathbb{R}^n \to \mathbb{R} \) is convex if, and only if, for each \( x \in \text{dom} f \) and \( v \in \mathbb{R}^n \), the function \( g(t) = f(x+tv) \), with domain the set \( \{t \in \mathbb{R} | x+tv \in \text{dom} f \} \), is convex.

**Proof.** A proof appears in the Appendix at the end of this chapter.

Convex functions have some very important properties. For example, they are continuous on the interior of their domain. Discontinuities can only exist on the boundary of their domain.

The theory of convex functions is extensive. In this course, we will cover only the topics which are necessary for our developments.
3.2 First Order Conditions

Theorem 3.2.1. Let $\text{dom} f \subseteq \mathbb{R}^n$ be an open set and $f : \text{dom} f \to \mathbb{R}$ be a differentiable function. Then, $f$ is convex if, and only if, $\text{dom} f$ is a convex set and

$$f(y) \geq f(x) + \nabla f(x)^T(y - x),$$

(3.4)

for any $x, y \in \text{dom} f$.

The inequality (3.4) is very important. It states that the first-order Taylor approximation at any point of the domain of a convex function is a global underestimator of the function. That is, from local information (the values of the function and its derivative at a point), we get global information (a global underestimator of the function).

Proof. (Direct) Let $f$ be a convex function and $x, y \in \text{dom} f$. Then, for $0 < \theta \leq 1$ we have that $(1 - \theta)x + \theta y = x + \theta(y - x) \in \text{dom} f$ and

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

Dividing both sides of the above inequality by $\theta$, and after simple algebraic operations, we obtain

$$f(y) \geq f(x) + \frac{f(x + \theta(y - x)) - f(x)}{\theta}.$$

Taking the limit for $\theta \to 0^+$, we get (why?)

$$f(y) \geq f(x) + \nabla f(x)^T(y - x),$$
which is the relation to be proved.

(*Inverse*) Assume that (3.4) holds for every \(x, y \in \text{dom } f\). Let \(x_1, x_2 \in \text{dom } f\), with \(x_1 \neq x_2\), and \(0 \leq \theta \leq 1\) and set \(x = \theta x_1 + (1 - \theta) x_2\).

Using (3.4) twice, we get

\[
\begin{align*}
    f(x_1) &\geq f(x) + \nabla f(x)^T (x_1 - x), \\
    f(x_2) &\geq f(x) + \nabla f(x)^T (x_2 - x).
\end{align*}
\]

Multiplying (3.5) by \(\theta\) and (3.6) with \((1 - \theta)\) and adding the inequalities, we obtain

\[
\theta f(x_1) + (1 - \theta) f(x_2) \geq f(x) + \nabla f(x)^T (\theta x_1 + (1 - \theta) x_2 - x) = f(x),
\]

proving that \(f\) is convex.

\[\square\]

### 3.3 Second Order Conditions

**Theorem 3.3.1.** Let \(\text{dom } f \subseteq \mathbb{R}^n\) be an open set and \(f : \text{dom } f \to \mathbb{R}\) be continuous doubly differentiable function. Then, \(f\) is convex if, and only if, \(\text{dom } f\) is a convex set and

\[
\nabla^2 f(x) \succeq 0
\]

for any \(x \in \text{dom } f\).

*Proof. (Direct)* Let \(f\) be a convex function, \(x \in \text{dom } f\), and \(h \in \mathbb{R}^n\). Since \(\text{dom } f\) is an open set, we have that \(x + \lambda h \in \text{dom } f\), for \(|\lambda|\) sufficiently small. From (3.4), we get

\[
f(x + \lambda h) \geq f(x) + \lambda \nabla f(x)^T h.
\]

From the second-order Taylor expansion, we have

\[
f(x + \lambda h) = f(x) + \lambda \nabla f(x)^T h + \frac{\lambda^2}{2} h^T \nabla^2 f(x) h + O(\lambda^3).
\]

Combining the above two relations, we obtain that

\[
\frac{\lambda^2}{2} h^T \nabla^2 f(x) h + O(\lambda^3) \geq 0.
\]

Dividing by \(\lambda^2\) and taking the limit as \(\lambda \to 0\), we obtain

\[
h^T \nabla^2 f(x) h \geq 0
\]
for every $h \in \mathbb{R}^n$, which gives that, for every $x \in \text{dom} f$,

$$\nabla^2 f(x) \succeq 0.$$  

(Inverse) We must prove that, if $\nabla^2 f(x) \succeq 0$ for every $x \in \text{dom} f$, then $f$ is convex. From the second-order Taylor expansion, we get

$$f(y) = f(x) + \nabla f(x)^T (y - x) + \frac{1}{2} (y - x)^T \nabla^2 f(z) (y - x),$$

with $z = \lambda x + (1 - \lambda)y$, for some $\lambda \in [0, 1]$.

Since we have assumed that $\nabla^2 f(x) \succeq 0$ for every $x \in \text{dom} f$, we have that $\nabla^2 f(z) \succeq 0$, which implies that

$$(y - x)^T \nabla^2 f(z) (y - x) \geq 0.$$  

Therefore, for each $x, y \in \text{dom} f$,

$$f(y) \geq f(x) + \nabla f(x)^T (y - x),$$

and, by Theorem 3.4, we conclude that $f$ is convex.

\[ \square \]

3.4 Examples of Convex Functions

**Example 3.4.1.** In the sequel, we provide examples of scalar convex and concave functions.

1. $f(x) = e^{ax}$ is convex in $\mathbb{R}$, for every $a \in \mathbb{R}$.
2. $f(x) = x^a$ is convex in $\mathbb{R}_{++}$ if $a \geq 1$ or $a \leq 0$, and concave if $0 \leq a \leq 1$.
3. $f(x) = |x|^p$ is convex in $\mathbb{R}$, for $p \geq 1$.
4. $f(x) = \log(x)$ is concave in $\mathbb{R}_{++}$.

**Example 3.4.2.** Quadratic functions. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$, with

$$f(x) = \frac{1}{2} x^T P x + q^T x + r,$$

with $P \in \mathbb{R}^{n \times n}$ and $P = P^T$, $q \in \mathbb{R}^n$ and $r \in \mathbb{R}$.

It can be proved that

$$\nabla f(x) = Px + q$$

and

$$\nabla^2 f(x) = P.$$
and
\[ \nabla^2 f(x) = P. \]
From the second-order conditions, we conclude that \( f \) is convex (strictly convex) if, and only if, \( P \succeq 0 \) (\( P \succ 0 \)).

**Example 3.4.3.** In the sequel, we provide examples of vector convex and concave functions.

1. \( f(x) = \|x\| \) is convex in \( \mathbb{R}^n \), where \( \| \cdot \| \) is any norm in \( \mathbb{R}^n \).
2. \( f(x) = \max\{x_1, \ldots, x_n\} \) is convex in \( \mathbb{R}^n \).
3. The function \( f(x) = \log(e^{x_1} + \cdots + e^{x_n}) \) is convex in \( \mathbb{R}^n \).
4. The function \( f(x) = (\prod_{i=1}^{n} x_i)^{\frac{1}{n}} \) is concave in \( \mathbb{R}_{++}^n \).

### 3.5 Sublevel Sets

**Definition 3.5.1.** Let \( f : \text{dom} f \subseteq \mathbb{R}^n \to \mathbb{R} \). The set
\[ S_c = \{ x \in \text{dom} f \mid f(x) \leq c \} \tag{3.8} \]
is called the \( c \)-sublevel set of \( f \).

**Theorem 3.5.1.** Let \( f : \text{dom} f \subseteq \mathbb{R}^n \to \mathbb{R} \) be a convex function. Then, the \( c \)-sublevel set \( S_c \) is convex, for every \( c \in \mathbb{R} \).
3.6. Epigraph

**Figure 3.4:** Convex function epigraph.

Proof. [We must prove that if \( x_1, x_2 \in S_c \) and \( 0 \leq \theta \leq 1 \), then \( x = \theta x_1 + (1 - \theta) x_2 \in S_c \).]

Let \( x_1, x_2 \in S_c \) and \( 0 \leq \theta \leq 1 \). Then, \( f(x_1) \leq c \) and \( f(x_2) \leq c \), and \( x = \theta x_1 + (1 - \theta) x_2 \in \text{dom} f \). Furthermore,

\[
f(x) = f(\theta x_1 + (1 - \theta) x_2) \leq \theta f(x_1) + (1 - \theta) f(x_2) \\
\leq \theta c + (1 - \theta) c \\
= c.
\]

Therefore, \( x = \theta x_1 + (1 - \theta) x_2 \in S_c \). \( \square \)

We must point out that the inverse proposition does not hold. For example, the sublevel sets of the function \( f : \mathbb{R}_+ \to \mathbb{R} \), with \( f(x) = \log(x) \), are convex. But, the function \( f \) is concave.

### 3.6 Epigraph

**Definition 3.6.1.** Let \( f : \text{dom} f \subseteq \mathbb{R}^n \to \mathbb{R} \). The set

\[
\text{epif} = \{(x, t) \mid x \in \text{dom} f, f(x) \leq t\}
\]

is called the **epigraph** of \( f \).

**Theorem 3.6.1.** Let \( f : \text{dom} f \subseteq \mathbb{R}^n \to \mathbb{R} \). Function \( f \) is convex if, and only if, \( \text{epif} \) is a convex set.

**Proof.** (Direct). [We must prove that if \( f \) is a convex function, then \( \text{epif} \) is a convex set.]
We assume that \( f \) is convex. To prove that \( \text{epi} f \) is a convex set, we need to prove that if
\[
(x_1, t_1), (x_2, t_2) \in \text{epi} f, \quad 0 \leq \theta \leq 1,
\]
then
\[
\theta (x_1, t_1) + (1 - \theta) (x_2, t_2) \in \text{epi} f.
\]
Assume that \((x_1, t_1), (x_2, t_2) \in \text{epi} f\), which implies that
\[
x_1, x_2 \in \text{dom} f, \quad f(x_1) \leq t_1, \quad f(x_2) \leq t_2.
\] (3.9)

Since \( f \) is convex, we have, for \( 0 \leq \theta \leq 1 \),
\[
f(\theta x_1 + (1 - \theta) x_2) \leq \theta f(x_1) + (1 - \theta) f(x_2)
\] (3.9)
\[
\leq \theta t_1 + (1 - \theta) t_2,
\]
that is, \( \theta (x_1, t_1) + (1 - \theta) (x_2, t_2) \in \text{epi} f \).

(Inverse). [We must prove that if \( \text{epi} f \) is a convex set, then \( f \) is a convex function.]

We assume that the set \( \text{epi} f \) is convex. We know that the projection of \( \text{epi} f \) on its first \( n \) coordinates is a convex set. Therefore, the set \( \text{dom} f \) is convex.

In addition, if \((x_1, t_1), (x_2, t_2) \in \text{epi} f \) and \( 0 \leq \theta \leq 1 \), then \( \theta (x_1, t_1) + (1 - \theta) (x_2, t_2) \in \text{epi} f \).

We notice that we can set \( t_1 = f(x_1) \) and \( t_2 = f(x_2) \) (why?), taking that \( \theta (x_1, f(x_1)) + (1 - \theta)(x_2, f(x_2)) \in \text{epi} f \), that is,
\[
f(\theta x_1 + (1 - \theta) x_2) \leq \theta f(x_1) + (1 - \theta) f(x_2),
\]
that is, \( f \) is a convex function. \( \square \)

### 3.7 Function operations that preserve convexity

#### 3.7.1 Non-negative weighted sums

**Theorem 3.7.1.** Let \( f_k : \mathbb{R}^n \to \mathbb{R} \), for \( k = 1, \ldots, K \), be convex functions and \( w_k \geq 0 \), for \( k = 1, \ldots, K \). Then, the function
\[
f(x) = \sum_{k=1}^{K} w_k f_k(x)
\]
is convex.
Proof. From the convexity of $f_k$, for $k = 1, \ldots, K$, we have that, if $x_1, x_2 \in \mathbb{R}^n$ and $0 \leq \theta \leq 1$, then

$$f_k(\theta x_1 + (1 - \theta)x_2) \leq \theta f_k(x_1) + (1 - \theta)f_k(x_2), \text{ for } k = 1, \ldots, K.$$ 

Multiplying both sides of each inequality by $w_k$, for $k = 1, \ldots, K$, and summing, we get

$$\sum_{k=1}^{K} w_k f_k(\theta x_1 + (1 - \theta)x_2) \leq \sum_{k=1}^{K} w_k (\theta f_k(x_1) + (1 - \theta)f_k(x_2))$$

$$= \theta \sum_{k=1}^{K} w_k f_k(x_1) + (1 - \theta) \sum_{k=1}^{K} w_k f_k(x_2),$$

which gives that

$$f(\theta x_1 + (1 - \theta)x_2) \leq \theta f(x_1) + (1 - \theta)f(x_2).$$

\[\square\]

### 3.7.2 Synthesis with affine function

**Theorem 3.7.2.** Let $f : \mathbb{R}^n \to \mathbb{R}$, $A \in \mathbb{R}^{n \times m}$, and $b \in \mathbb{R}^n$. Let $g : \mathbb{R}^m \to \mathbb{R}$, with $g(x) = f(Ax + b)$ and domain $\text{dom } g = \{x \in \mathbb{R}^m | Ax + b \in \text{dom } f\}$. If $f$ is a convex function, then $g$ is also convex. If $f$ is a concave function, then $g$ is also concave.

*Proof. It is left to the reader.* \[\square\]

### 3.7.3 Point maxima and minima

**Theorem 3.7.3.** Let $f_k : \mathbb{R}^n \to \mathbb{R}$, for $k = 1, \ldots, K$, be convex functions. Then, the function

$$f(x) = \max\{f_1(x), \ldots, f_K(x)\}$$

is convex.

*Proof. From the convexity of $f_k$, for $k = 1, \ldots, K$, we have that, if $x_1, x_2 \in \mathbb{R}^n$ and $0 \leq \theta \leq 1$, then

$$f_k(\theta x_1 + (1 - \theta)x_2) \leq \theta f_k(x_1) + (1 - \theta)f_k(x_2), \text{ for } k = 1, \ldots, K.$$*
Therefore,
\[
\max_{k=1,\ldots,K} \{ f_k(\theta x_1 + (1-\theta)x_2) \} \leq \max_{k=1,\ldots,K} \{ \theta f_k(x_1) + (1-\theta)f_k(x_2) \}
\]
\[
\leq \max_{k=1,\ldots,K} \{ \theta f_k(x_1) \} + \max_{k=1,\ldots,K} \{ (1-\theta)f_k(x_2) \}
\]
\[
= \theta \max_{k=1,\ldots,K} \{ f_k(x_1) \} + (1-\theta) \max_{k=1,\ldots,K} \{ f_k(x_2) \}
\]
\[
= \theta f(x_1) + (1-\theta)f(x_2).
\]

The only non-trivial inequality is (!), which, essentially, is equivalent to the inequality
\[
\max_{k=1,\ldots,K} \{ a_k + b_k \} \leq \max_{k=1,\ldots,K} \{ a_k \} + \max_{k=1,\ldots,K} \{ b_k \},
\]
for any scalars \(a_k, b_k\), for \(k = 1, \ldots, K\) (try to prove this inequality).

### 3.7.4 Synthesis of functions

Let \( f, h : \mathbb{R} \rightarrow \mathbb{R} \) and \( g(t) = (f \circ h)(t) \). \( g \) is convex if, and only if, \( D^2 g(t) \geq 0 \) for every \( t \in \text{dom} g \).

We know that \( Dg(t) = Df(h(t)) \cdot Dh(t) \) and
\[
D^2 g(t) = D^2 f(h(t))(Dh(t))^2 + Df(h(t))D^2 h(t).
\]

With appropriate choice of \( f \) and \( h \), we can have \( D^2 g(t) \geq 0 \) for each \( t \in \text{dom} g \).

For example, if \( f \) is convex and non-decreasing, which imply that \( Df(t) \geq 0 \) and \( D^2 f(t) \geq 0 \), and \( h \) is convex, which implies that \( D^2 h(t) \geq 0 \), then \( g \) is convex (why?).

For more examples, see pages 83–87 of Boyd–Vandenberghe.

### 3.7.5 Partial minimization

**Theorem 3.7.4.** Let \( f : \mathbb{R}^{n+m} \rightarrow \mathbb{R} \), \( x \in \mathbb{R}^n \) and \( y \in \mathbb{R}^m \). Let \( f(x, y) \) be convex, with respect to \((x, y)\), and let \( C \subseteq \mathbb{R}^m \) be a convex set. Then, \( g(x) = \min_{y \in C} f(x, y) \) is convex, with respect to \( x \), if \( g(x) > -\infty \) for every \( x \) in
\[
\text{dom} \ g = \{ x \mid (x, y) \in \text{dom} f \text{ for some } y \in C \},
\]
and \( g(x) < \infty \), for some \( x \in \text{dom} g \).
Proof. We must prove that, if \(x_1, x_2 \in \text{dom} g\) and \(0 \leq \theta \leq 1\), then \(x = \theta x_1 + (1 - \theta) x_2 \in \text{dom} g\) and \(g(x) \leq \theta g(x_1) + (1 - \theta) g(x_2)\).

For each \(\epsilon > 0\), we can find \(y_1\) and \(y_2\) such that 
\[
(x_1, y_1), (x_2, y_2) \in \text{dom} f \\
g(x_1) + \epsilon \geq f(x_1, y_1), \quad g(x_2) + \epsilon \geq f(x_2, y_2).
\]

Multiplying the first inequality by \(\theta\) and the second with \((1 - \theta)\) and adding, we obtain
\[
\theta g(x_1) + (1 - \theta) g(x_2) + \epsilon \geq \theta f(x_1, y_1) + (1 - \theta) f(x_2, y_2)
\]
\[
\geq f(\theta x_1 + (1 - \theta) x_2, \theta y_1 + (1 - \theta) y_2)
\]
\[
= f(x, \theta y_1 + (1 - \theta) y_2)
\]
\[
\geq g(x).
\]

Therefore, \(x \in \text{dom} g\) and, since the inequality holds for every \(\epsilon > 0\), we have that \(g(x) \leq \theta g(x_1) + (1 - \theta) g(x_2)\). Therefore, \(g\) is a convex function (see, OPTIII, p. 68).

Appendix

Sketch of proof of Theorem 3.1.1, assuming \(f : \mathbb{R}^n \to \mathbb{R}\). (Work out the details for \(f : \text{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}\), with \(\text{dom} f\) a convex set).

Let \(f : \mathbb{R}^n \to \mathbb{R}, x, y \in \mathbb{R}^n\), and \(g_{x,y} : \mathbb{R} \to \mathbb{R}, \) with \(g_{x,y}(t) = f(tx + (1 - t)y)\). Hereafter, to simplify the notation, \(g_{x,y}\) will be denoted as \(g\).

Direct: If \(f\) is a convex function, then \(g\) is a convex function.

If \(f : \mathbb{R}^n \to \mathbb{R}\) is a convex function and \(x, y \in \mathbb{R}^n\), then \(g_{x,y} : \mathbb{R} \to \mathbb{R}, \) with \(g_{x,y}(t) = f(tx + (1 - t)y)\), is a convex function.

To prove that \(g\) is a convex function, we must prove that, for every \(t_1, t_2 \in \mathbb{R}\) and \(0 \leq \lambda \leq 1\), it must be true that
\[
g(\lambda t_1 + (1 - \lambda) t_2) \leq \lambda g(t_1) + (1 - \lambda) g(t_2).
\]

Equivalently, we must prove that
\[
f ((\lambda t_1 + (1 - \lambda) t_2)x + (1 - (\lambda t_1 + (1 - \lambda) t_2))y) \leq \\
\lambda f (t_1 x + (1 - t_1) y) + (1 - \lambda) f (t_2 x + (1 - t_2) y).
\]
CHAPTER 3. CONVEX FUNCTIONS

After simple algebraic operations, we get
\[
(\lambda t_1 + (1 - \lambda) t_2) x + (1 - (\lambda t_1 + (1 - \lambda) t_2)) y = \\
\lambda (t_1 x + (1 - t_1) y) + (1 - \lambda) (t_2 x + (1 - t_2) y). \tag{3.13}
\]

Therefore, (3.12) is equivalently written as
\[
f(\lambda (t_1 x + (1 - t_1) y) + (1 - \lambda) (t_2 x + (1 - t_2) y)) \leq \\
\lambda f(t_1 x + (1 - t_1) y) + (1 - \lambda) f(t_2 x + (1 - t_2) y), \tag{3.14}
\]
which is true, because we have assumed that \( f \) is a convex function.

**Inverse:** If \( g \) is a convex function, then \( f \) is a convex function.

We must prove that, if \( x, y \in \mathbb{R}^n \) and \( 0 \leq \lambda \leq 1 \), then
\[
f(\lambda x + (1 - \lambda) y) \leq \lambda f(x) + (1 - \lambda) f(y). \tag{3.15}
\]

From the definition of \( g \), we have
\[
g(\lambda) = f(\lambda x + (1 - \lambda) y), \quad g(1) = f(x), \quad g(0) = f(y). \tag{3.16}
\]

From the convexity of \( g \), we get
\[
g(\lambda) = g(\lambda \cdot 1 + (1 - \lambda) \cdot 0) \leq \lambda g(1) + (1 - \lambda) g(0), \tag{3.17}
\]
which is equivalent to (3.15), completing the proof.
Chapter 4

Convex Optimization Problems

4.1 Optimization Problems

Definition 4.1.1. An optimization problem is defined as

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0, \quad i = 1, \ldots, m, \\
& \quad h_i(x) = 0, \quad i = 1, \ldots, p,
\end{align*}
\]  \tag{4.1}

where

1. the vector \( x \in \mathbb{R}^n \) is called optimization variable,
2. the function \( f_0 : \text{dom} \ f_0 \subseteq \mathbb{R}^n \to \mathbb{R} \), is called cost function,
3. the inequalities \( f_i(x) \leq 0 \), with \( f_i : \text{dom} \ f_i \subseteq \mathbb{R}^n \to \mathbb{R} \), for \( i = 1, \ldots, m \), are called inequality constraints, and
4. the equalities \( h_i(x) = 0 \), with \( h_i : \text{dom} \ h_i \subseteq \mathbb{R}^n \to \mathbb{R} \), for \( i = 1, \ldots, p \), are called equality constraints.

The set of points for which all functions are defined is

\[
\mathbb{D} := \bigcap_{i=0}^{m} \text{dom} \ f_i \cap \bigcap_{i=1}^{p} \text{dom} \ h_i.
\]  \tag{4.2}

Definition 4.1.2. A point \( x \in \mathbb{D} \) is called feasible if it satisfies all constraints.

An optimization problem is called feasible if there exists a feasible point for that problem. Otherwise, it is called infeasible.
The set of feasible points of an optimization problem is called \textbf{feasible set} of the problem.

The feasible set of the optimization problem (4.1) is the set

\[ X := \{ x \in \mathbb{D} \mid f_i(x) \leq 0, \ i = 1, \ldots, m, \ h_i(x) = 0, \ i = 1, \ldots, p \}. \] (4.3)

The optimal value of the problem (4.1) is defined as

\[ p^* = \inf \{ f_0(x) \mid x \in X \}. \] (4.4)

\( p^* \) can take the values \( \pm \infty \). If the problem is infeasible, then \( p^* = \infty \). If there are feasible points \( x_k \in \mathbb{D} \), with \( f_0(x_k) \to -\infty \), when \( k \to \infty \), then \( p^* = -\infty \) and the problem is called unbounded from below.

\section*{4.1.1 Local and Global Minima}

\textbf{Definition 4.1.3.} A point \( x^* \in X \) is called \textbf{optimal point} if \( f_0(x^*) = p^* \). The set

\[ X_{\text{opt}} = \{ x \in X \mid f(x) = p^* \} \] (4.5)

is called \textbf{optimal set}.

Alternatively, we say that \( x \in X \) is an optimal point if \( f_0(x) \leq f_0(y) \) for every \( y \in X \).

If there is an optimal point for the problem (4.1), then we say that the problem has a solution and the optimal value is attained.

If the set \( X_{\text{opt}} \) is empty, then we say that the optimal value is not attained (this is always the case when the problem is unbounded from below).

\textbf{Definition 4.1.4.} The point \( x \in X \) is called \textbf{locally optimal} if there exists \( \epsilon > 0 \) such that \( f_0(x) \leq f_0(y) \) for every \( y \in X \), with \( ||y - x||_2 < \epsilon \).

If \( x \in X \) and \( f_i(x) = 0 \), for some \( i \) in the set \( \{1, \ldots, m\} \), then we say that the \( i \)-th inequality constraint is \textbf{active} at \( x \). On the other hand, if \( f_i(x) < 0 \), for some \( i \) in the set \( \{1, \ldots, m\} \), then we say that the \( i \)-th inequality constraint is \textbf{inactive} at \( x \).

A constraint is called \textbf{redundant} if its deletion does not change the feasible set.
4.2 Convex Optimization Problems

Definition 4.2.1. A problem of the form

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& \quad a_i^T x = b_i, \quad i = 1, \ldots, p,
\end{align*}
\]

(4.6)
is called convex optimization problem if the functions \( f_i \), for \( i = 0, \ldots, m \), are convex.

The feasible set of the optimization problem (4.6) is

\[
X := \{ x \in \mathbb{D} | f_i(x) \leq 0, \quad i = 1, \ldots, m, \quad a_i^T x = b_i, \quad i = 1, \ldots, p \}.
\]

(4.7)

Theorem 4.2.1. The feasible set of a convex optimization problem is convex.

Proof. The feasible set of a convex optimization problem is the intersection of the (convex) domains of the functions \( f_i \), for \( i = 0, \ldots, m \), the (convex) 0-sublevel sets \( \{ x | f_i(x) \leq 0 \} \), for \( i = 1, \ldots, m \), and the hyperplanes \( \{ x | a_i^T x = b_i \} \), for \( i = 1, \ldots, p \). Therefore, it is convex.

Theorem 4.2.2. Consider the convex optimization problem (4.6). Then, the set of points \( x \in X \) which minimize \( f_0 \) is convex. Moreover, every local minimum of \( f_0 \) is a global minimum.

Proof. Assume that the problem has a solution \( p^* \). Then, the set of optimal points is the convex \( p^* \)-sublevel set \( C_{p^*} := \{ x \in X | f_0(x) \leq p^* \} \).

\((\text{We shall prove the second statement by Reductio ad Absurdum).}\) Let \( x \in X \) be a local but not global minimum of \( f_0 \). Then, there exists \( y \in X \) such that \( f_0(y) < f_0(x) \). If we consider \( f_0 \) on the line segment \( \theta y + (1 - \theta) x \), with \( 0 \leq \theta \leq 1 \), we have

\[
f_0(\theta y + (1 - \theta) x) \leq \theta f_0(y) + (1 - \theta) f_0(x) < f_0(x),
\]

(4.8)
from which we conclude that \( x \) is not a local minimum (why?). This is false and the proof is complete.
4.3 An Optimality Criterion for Differentiable Cost Functions

Let $f_0 : \text{dom } f_0 \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ convex and differentiable function. We know that, for every $x, y \in \text{dom } f_0$, the following inequality holds:

$$f_0(y) \geq f_0(x) + \nabla f_0(x)^T(y - x). \quad (4.9)$$

**Theorem 4.3.1.** Let $X = \{x \in D \mid f_i(x) \leq 0, i = 1, \ldots, m, a_i^T x = b_i, \ i = 1, \ldots, p\}$. The point $x \in X$ is optimal for the convex optimization problem (4.6) if, and only if,

$$\nabla f_0(x)^T(y - x) \geq 0, \text{ for every } y \in X. \quad (4.10)$$

**Proof.** *(Direct, through reductio ad absurdum)* Assume that $x \in X$ is optimal but there exists a point $y \in X$ such that $\nabla f_0(x)^T(y - x) < 0$. If we define, for $0 \leq \theta \leq 1$,

$$z(\theta) := \theta y + (1 - \theta)x = x + \theta(y - x) \quad (4.11)$$

we have that $z(\theta) \in X$, for every $0 \leq \theta \leq 1$. For small and positive $\theta$, we should have $f_0(z(\theta)) < f_0(x)$ because

$$\left. \frac{d}{d\theta} f_0(z(\theta)) \right|_{\theta=0} = \nabla f_0(x)^T(y - x) < 0. \quad (4.12)$$

Therefore, $x$ is not optimal, which is false.

*(Inverse)* If (4.10) holds, then, due to (4.9), we have that

$$f_0(y) \geq f_0(x), \quad \text{for each } y \in X, \quad (4.13)$$

that is, $x$ is an optimal point for the problem (4.6).

The proof is complete. \qed
4.3.1 Unconstrained Problems

**Theorem 4.3.2.** Let \( f_0 : \mathbb{R}^n \to \mathbb{R} \) be a convex differentiable function. The point \( x \in \mathbb{R}^n \) is an optimal point for the unconstrained minimization problem if, and only if,

\[
\nabla f_0(x) = 0.
\]

(4.14)

**Proof.** By Theorem 4.3.1, we should have

\[
\nabla f_0(x)^T h \geq 0, \quad \text{for each } h \in \mathbb{R}^n.
\]

(4.15)

This means that \( \nabla f_0(x) = 0 \). Because if \( \nabla f_0(x) \neq 0 \) and we choose \( h = -\nabla f_0(x) \), then we get that

\[
-\|\nabla f_0(x)\|_2^2 \geq 0,
\]

(4.16)

which is false. \( \square \)

4.3.2 Problems with linear constraints

Consider the problem

\[
\text{minimize} \quad f_0(x) \\
\text{subject to} \quad Ax = b.
\]

(4.17)

The point \( x^* \in X \) is optimal if

\[
\nabla f_0(x^*)^T(y - x^*) \geq 0,
\]

(4.18)

for every \( y \), with \( Ay = b \).

4.3.3 Optimization in non-negative orthant

Let the problem

\[
\text{minimize} \quad f_0(x) \\
\text{subject to} \quad x \succeq 0.
\]

(4.19)

The point \( x^* \succeq 0 \) is optimal if

\[
\nabla f_0(x^*)^T(y - x^*) \geq 0, \quad \text{for every } y \succeq 0.
\]

(4.20)
4.4 Equivalent problems

In simple words, we would say that two optimization problems are equivalent when “from the solution of one we can compute the solution of the other.”

**Example.** The problem

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad \|Ax - b\|_2 \leq r,
\end{align*}
\]

(4.21)

is equivalent to the problem

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad \|Ax - b\|_2^2 \leq r^2.
\end{align*}
\]

(4.22)

Obviously, the two problems have the same solution. But the constraint of the first problem refers to a nondifferentiable function, while the corresponding function for the second problem is differentiable.

More complex (and more interesting) examples of equivalent problems will be mentioned later in the course.

4.4.1 Epigraph Form

The problem

\[
\begin{align*}
\text{minimize} & \quad t \\
\text{subject to} & \quad f_0(x) - t \leq 0 \\
& \quad f_i(x) \leq 0, \quad i = 1 \ldots, m, \\
& \quad Ax = b
\end{align*}
\]

(4.23)

is equivalent to the problem (4.6) and is called problem in epigraph form.

4.5 Linear optimization problems

The optimization problem with linear cost function and affine equality and inequality constraints

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Gx \preceq h \\
& \quad Ax = b
\end{align*}
\]

(4.24)
is called a linear programming problem or linear program (see Fig. 4.2). Two forms of linear programs are very common.

The standard form, in which the problem is expressed as follows

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax = b \\
& \quad x \succeq 0.
\end{align*}
\] (4.25)

The inequality form, in which the problem is expressed as

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax \preceq b.
\end{align*}
\] (4.26)

It can be proved that any linear program can be expressed in standard or inequality form (via slack variables).

### 4.6 Quadratic optimization problems

The optimization problem

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{2} x^T P x + q^T x + r \\
\text{subject to} & \quad Gx \preceq h \\
& \quad Ax = b,
\end{align*}
\] (4.27)

is called quadratic optimization problem. The problem is convex if \( P \succeq O \).
The optimization problem
\[
\begin{align*}
\text{minimize} & \quad \frac{1}{2} x^T P_0 x + q_0^T x + r_0 \\
\text{subject to} & \quad \frac{1}{2} x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m \\
& \quad A x = b,
\end{align*}
\]
(4.28)

is called \textbf{quadratic optimization problem with quadratic constraints}. The problem is convex if \( P_i \succeq O \), for \( i = 0, \ldots, m \).

### 4.6.1 Least Squares

The optimization problem
\[
\min_x \| A x - b \|_2^2
\]
(4.29)
4.6. QUADRATIC OPTIMIZATION PROBLEMS

is called linear least-squares problem.

The optimization problem

\[
\begin{align*}
\min_x & \quad \|Ax - b\|_2^2 \\
\text{subject to} & \quad l_i \leq x_i \leq u_i, \quad i = 1, \ldots, n,
\end{align*}
\]

(4.30)

is called linear least squares problem with box constraints.

Appendix

Let \(a \in \mathbb{R}\). We recall that \(f(x) = O(g(x))\), for \(x \to a\), if there are constants \(\delta > 0\) and \(K := K(a, \delta) > 0\) or \(K := K(\delta) > 0\) such that

\[
|f(x)| \leq K|g(x)|, \quad \text{for each } x \in \mathbb{R}, \quad \text{with } |x - a| < \delta.
\]

(4.31)

**Theorem 4.6.1.** Let \(f : \mathbb{R}^n \to \mathbb{R}\) be a differentiable function and \(x, \Delta x \in \mathbb{R}^n\) such that \(\nabla f(x)^T \Delta x < 0\). Then, for \(0 < t \in \mathbb{R}\) sufficiently small, we have that \(f(x + t\Delta x) < f(x)\).

**Proof.** From the first-order Taylor expansion and the definition of the \(O(t^2)\) expression, we have, for sufficiently small \(t\), i.e., \(0 < t < t_0\), that

\[
f(x + t\Delta x) = f(x) + t\nabla f(x)^T \Delta x + O(t^2)
\]

(4.32)

with \(K_{t_0} \in \mathbb{R}_{++}\). If

\[
t\nabla f(x)^T \Delta x + K_{t_0} t^2 < 0,
\]

(4.33)

then \(f(x + t\Delta x) < f(x)\). The inequality (4.33) is true for

\[
t < -\frac{1}{K_{t_0}} \nabla f(x)^T \Delta x,
\]

(4.34)

completing the proof.

Alternatively

\[
f(x + t\Delta x) = f(x) + t\nabla f(x)^T \Delta x + \frac{t^2}{2} \Delta x^T \nabla^2 f(w) \Delta x,
\]

(4.35)

with \(w = \theta x + (1 - \theta)(x + t\Delta x)\), for some \(\theta \in [0, 1]\). If we put

\[
z(\theta) = \Delta x^T \nabla^2 f(\theta x + (1 - \theta)(x + t\Delta x)) \Delta x
\]

(4.36)

and define \(z^* = \max_{\theta \in [0, 1]} z(\theta)\), then we have

\[
f(x + t\Delta x) \leq f(x) + t\nabla f(x)^T \Delta x + \frac{t^2}{2} z^*.
\]

(4.37)

The rest of the proof is the same as in the original proof, with \(K_{t_0} = \frac{z^*}{2}\). \(\square\)
Chapter 5

Unconstrained convex optimization

5.1 Unconstrained convex optimization problems

The problem we shall study in this chapter is the unconstrained convex optimization problem

$$\min_{x} f(x),$$

(5.1)

where $f : \mathbb{R}^n \to \mathbb{R}$ is a doubly differentiable convex real function.

We assume that the problem (5.1) has a solution and that the minimum value of $f$ equals $p_\star$. Therefore, there exists $x_\star \in \mathbb{R}^n$ such that $f(x_\star) = p_\star$.

As we have seen, a point $x_\star \in \mathbb{R}^n$ is a solution of the problem (5.1) if, and only if,

$$\nabla f(x_\star) = 0.$$

(5.2)

The relation (5.2) is a system of (usually, non-linear) equations, which rarely has a closed form solution, and which is usually solved through iterative procedures.

An iterative procedure for solving (5.2) is a procedure which generates a sequence of points $x_k \in \mathbb{R}^n$ such that $x_k \to x_\star$, when $k \to \infty$.

The system (5.2) can be solved through direct or indirect iterative procedures.

A direct iterative process tries to compute $x_\star$ by solving the system (5.2), while an indirect one tries to take advantage of properties of $f$ and compute $x_\star$ indirectly.
5.2 Starting point and sublevel sets

The methods we will present start from an initial point $x_0$ which belongs to $\text{dom} \ f$ and for which it should hold that the set

$$S = \{x \mid f(x) \leq f(x_0)\}$$

is closed.

This condition is satisfied if $x_0 \in \text{dom} \ f$ and $f$ is a closed function, that is, when all sublevel sets of $f$ are closed sets.

Two important cases of closed functions are as follows:

1. continuous functions $f$ with domain $\text{dom} \ f = \mathbb{R}^n$ are closed. Example: the function $f : \mathbb{R}^n \to \mathbb{R}$ with formula

$$f(x) = \frac{1}{2} x^T P x + q^T x + c,$$

with $P = P^T > 0$, is closed.

2. continuous functions with domain $\text{dom} \ f$ open set, for which function $f(x)$ tends to infinity when $x$ approaches $\text{bd} \ \text{dom} \ f$, are closed. Example: the function $f : \text{dom} \ f \subseteq \mathbb{R}^n \to \mathbb{R}$ defined as

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x).$$

The domain of $f$ is the set

$$\text{dom} \ f = \{x \in \mathbb{R}^n \mid a_i^T x < b_i, \ i = 1, \ldots, m\}.$$  

Notice that, as $x$ approaches the boundary of the domain, the function $f$ tends to infinity.

5.3 Descent Methods

The first method we will study is an indirect iterative process to solve (5.1) or, equivalently, (5.2) and is described as follows.

Suppose that at the $k$-th step of the process our estimate of the solution of (5.2) is the point $x_k$. 

5.4. GRADIENT METHOD

The next point is given by the relation

$$x_{k+1} = x_k + t_k \Delta x_k,$$  \hspace{1cm} (5.7)

for suitably chosen $t_k > 0$ and $\Delta x_k$.

It seems reasonable to choose $t_k$ and $\Delta x_k$ such that

$$f(x_{k+1}) \leq f(x_k),$$  \hspace{1cm} (5.8)

with equality only if $x_k = x_*$. Such iterative methods are called descent methods.

Due to the convexity of $f$, we know that if $\nabla f(x_k)^T(y - x_k) \geq 0$, then $f(y) \geq f(x_k)$. Therefore, in order to develop a descent method, we must choose $\Delta x_k$ such that

$$\nabla f(x_k)^T \Delta x_k < 0,$$  \hspace{1cm} (5.9)

that is, $\cos \angle (\Delta x_k, \nabla f(x_k)) < 0$. Such a direction of movement is called descent direction. In this case, $t_k$ is chosen as

$$t_k = \arg\min_{t > 0} f(x_k + t \Delta x_k).$$  \hspace{1cm} (5.10)

5.4 Gradient Method

Among the most common choices for $\Delta x_k$ is to set

$$\Delta x_k = -\nabla f(x_k).$$  \hspace{1cm} (5.11)

This direction is called negative gradient direction. Obviously, the choice (5.11) satisfies relation (5.9). Geometrically, we move from the point $x_k$ in the direction along which the rate of decrease of $f$ is maximum.

In Table 5.1, we present the gradient algorithm. Usually, for reasons we will explain later, as a termination criterion we set $\|\nabla f(x_k)\|_2 < \epsilon$, for some “small” $\epsilon > 0$.

5.4.1 Line Search

As we have seen, for a given descent direction $\Delta x_k$, the line search problem is expressed as

$$t_k = \arg\min_{t > 0} f(x_k + t \Delta x_k).$$  \hspace{1cm} (5.12)
\( x_0 \in \mathbb{R}^n, \ k = 0. \)

While (stopping criterion is FALSE)
1. \( \Delta x_k := -\nabla f(x_k). \)
2. Line search and choose \( t_k. \)
3. \( x_{k+1} = x_k + t_k \Delta x_k. \)
4. \( k := k + 1. \)

Table 5.1: Rank Algorithm.

5.4.2 Exact line search

When the solution of the problem (5.12) is expressed in closed form or is relatively easy to compute accurately, then we adopt the exact line search method to compute \( t_k. \)

Example 5.4.1. Exact line search for convex quadratic cost function and negative gradient direction.

Let the quadratic cost function \( f : \mathbb{R}^n \rightarrow \mathbb{R}, \) with
\[
    f(x) = \frac{1}{2} x^T P x + q^T x,
\]
with \( P \in \mathbb{R}^{n \times n}, \) \( P = P^T \succ O, \) and \( q \in \mathbb{R}^n. \)

As we have shown, the gradient of the function \( f \) is equal to
\[
    \nabla f(x) = Px + q.
\]

To solve the exact line search problem, we define the function
\[
    g(t) := f(x - t \nabla f(x))
\]
and we are looking for it
\[
    t_* = \arg \min_{t \geq 0} g(t).
\]

The function \( g \) is expressed as follows:
\[
    g(t) = \frac{1}{2} (x - t \nabla f(x))^T P (x - t \nabla f(x)) + q^T (x - t \nabla f(x))
    = \frac{1}{2} t^2 \nabla f(x)^T P \nabla f(x) - t (P x + q)^T \nabla f(x) + \frac{1}{2} x^T P x + q^T x
    = \frac{1}{2} t^2 \nabla f(x)^T P \nabla f(x) - t \| \nabla f(bfx) \|^2 + \frac{1}{2} x^T P x + q^T x
    = \frac{1}{2} at^2 - bt + c,
\]
5.4. GRADIENT METHOD

with \( a := \nabla f(x)^T P \nabla f(x) \), \( b := \| \nabla f(x) \|^2 \), and \( c := \frac{1}{2} x^T P x + q^T x \). The derivative of \( g \) is given by the relation

\[
\frac{dg(t)}{dt} = at - b.
\]

The solution of the equation \( \frac{dg(t)}{dt} = 0 \) is

\[
t^* = \frac{b}{a} = \frac{\| \nabla f(x) \|^2}{\nabla f(x)^T P \nabla f(x)} \geq 0.
\]

So, in this case, there is a closed form solution to the exact line search problem. \( \square \)

5.4.3 Backtracking line search

In many cases, solving the problem (5.12) is difficult. In these cases, we can adopt inexact line search techniques (note that there are many such techniques).

A common inexact line search technique is the **backtracking line search**. For each \( x \in \text{dom} f \), the backtracking line search takes as input the descent direction, \( \Delta x \), and returns a value \( t > 0 \) such that the value \( f(x + t\Delta x) \) is “sufficiently” smaller than \( f(x) \).

The method uses parameters \( \alpha \) and \( \beta \), with \( 0 < \alpha < 0.5 \) and \( 0 < \beta < 1 \), and is described in Table 5.4.3. It starts by setting \( t = 1 \), and decreases \( t \) by a multiplicative factor \( \beta \) until

\[
f(x + t\Delta x) \leq f(x) + \alpha t \nabla f(x)^T \Delta x.
\]

1. It can be shown that the condition (5.20) is satisfied for sufficiently small \( t \), therefore, the backtracking line search algorithm terminates always.

The proof is based on (1) the first-order Taylor expansion and (2) the fact that the vector \( \Delta x \) is a descent direction, that is, \( \nabla f(x)^T \Delta x < 0 \).

More specifically, from the first-order Taylor expansion, we have that, for sufficiently small \( t \),

\[
f(x + t\Delta x) \approx f(x) + t \nabla f(x)^T \Delta x < f(x) + \alpha t \nabla f(x)^T \Delta x.
\]

2. The first-order Taylor estimate of the decrease of the value of \( f \) is equal to

\[
f(x) - f(x + t\Delta x) = -t \nabla f(x)^T \Delta x.
\]

Therefore, the backtracking line search algorithm terminates when the decrease of \( f \) is at least \( -\alpha t \nabla f(x)^T \Delta x \), that is, at least \( \alpha \) times the first-order Taylor estimate of the decrease (see Figure 5.1).
Let $\Delta x$ be descent direction at $x$, $\alpha \in (0, 0.5)$, $\beta \in (0, 1)$.

$t := 1$.

While $( f(x + t\Delta x) > f(x) + \alpha t \nabla f(x)^T \Delta x )$

1. $t := \beta t$.

Table 5.2: Backtracking line search.

5.5 Convergence Analysis for Strongly Convex Functions

Convergence analysis of methods for solving optimization problems is extremely important because it provides information about the speed of convergence of the methods.

Additional features of the method, such as, for example, the computational complexity of each step of the method, complete the picture of the method.

Convergence analysis is usually accompanied by assumptions about the nature of the optimization problem, for example, assumptions about differentiability, strict or strong convexity of the cost function, etc.

In the sequel, we will study the convergence of the gradient method, assuming that the cost function is strongly convex.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a strongly convex doubly differentiable function. More specifically, we assume that there exist $0 < m \leq M < \infty$ such that

$$mI \preceq \nabla^2 f(x) \preceq MI, \quad \text{for every } x \in S,$$

where $I$ is the $(n \times n)$ identity matrix and $S := \{ x \in \mathbb{R}^n | f(x) \leq f(x_0) \}$. In practice, usually, the constants $m$ and $M$ are unknown. But, the assumption of their existence
allows convergence analysis of the gradient algorithm and leads to extremely important conclusions. For this reason, in what follows, we will assume the existence of these constants.

Relation (5.23) implies that, for any \( x \in S \) and \( y \in \mathbb{R}^n \),
\[
my^T I y \leq y^T \nabla^2 f(x) y \leq M y^T I y,
\]
(5.24)
or, equivalently,
\[
m\|y\|^2 \leq y^T \nabla^2 f(x)y \leq M\|y\|^2.
\]
(5.25)
Before proceeding, we mention that (5.23) is equivalent to
\[
\frac{1}{M} I \preceq (\nabla^2 f(x))^{-1} \preceq \frac{1}{m} I, \quad \text{for any } x \in S.
\]
(5.26)
We recall that, for any \( x, y \in S \),
\[
f(y) = f(x) + \nabla f(x)^T(y - x) + \frac{1}{2}(y - x)^T\nabla^2 f(z)(y - x),
\]
(5.27)
for some \( z \) on the line segment that connects the points \( x \) and \( y \).

Using (5.25), we can prove the inequalities, for any \( x, y \in S \):
\[
f(y) \geq f(x) + \nabla f(x)^T(y - x) + \frac{m}{2}\|y - x\|^2
\]
and
\[
f(y) \leq f(x) + \nabla f(x)^T(y - x) + \frac{M}{2}\|y - x\|^2,
\]
(5.28)
(5.29)
which will prove extremely useful in the sequel.

Their usefulness basically stems from the fact that, for a given \( x \), the expressions on the right-hand sides of (5.28) and (5.29) are quadratic functions of \( y \) which provide, respectively, a global underestimator and a global overestimator of \( f \), in the set \( S \).

Another inequality which will be useful later is
\[
p_* \geq f(x) - \frac{1}{2m}\|\nabla f(x)\|^2,
\]
(5.30)
which holds for every \( x \in S \). The proof is based on (5.28) and is as follows. For a given \( x \), the right-hand side of (5.28) is a quadratic function of \( y \), with minimum value \( f(x) - \frac{1}{2m}\|\nabla f(x)\|^2 \) (prove it). Hence, by (5.28), we have the inequality (why?)
\[
\min_{y \in S} f(y) \geq \min_{y \in S} \left( f(x) + \nabla f(x)^T(y - x) + \frac{m}{2}\|y - x\|^2 \right),
\]
(5.31)
which implies (5.30).

An interesting interpretation of (5.30) is as follows. If, for some \( x \in S \), we have that the value \( \|\nabla f(x)\|_2 \) is “small,” then the value of \( f(x) \) is “close” to the optimal value, \( p_* \). This interpretation offers a terminating condition of the gradient algorithm, for example, the condition \( \|\nabla f(x)\|_2 < \epsilon \), for a “small” positive \( \epsilon \).

Finally, using (5.28), it can be shown that, for every \( x \in S \),

\[
\|x - x_*\|_2 \leq \frac{2}{m} \|\nabla f(x)\|_2.
\]  

(5.32)

The proof is as follows. Setting \( y = x_* \) in (5.28), we get

\[
p_* = f(x_*) \geq f(x) + \nabla f(x)^T (x_* - x) + \frac{m}{2} \|x_* - x\|_2^2 \\
\geq f(x) - \|\nabla f(x)\|_2 \|x_* - x\|_2 + \frac{m}{2} \|x_* - x\|_2^2,
\]

(5.33)

where, in the second line, we used the Cauchy-Schwarz inequality. Since \( p_* \leq f(x) \), we have

\[
\frac{m}{2} \|x_* - x\|_2^2 - \|\nabla f(x)\|_2 \|x_* - x\|_2 \leq p_* - f(x) \leq 0,
\]

(5.34)

from which we get (5.32). This inequality states that, if for some \( x \) we have that the value \( \|\nabla f(x)\|_2 \) is “small,” then \( x \) is “close” to the optimal point \( x_* \). Therefore, we have another
5.5. CONVERGENCE ANALYSIS

argument for using the quantity $\| \nabla f(x) \|_2$ as a terminating criterion for the gradient algorithm.

During the analysis and in order to simplify notation, we will use expression $x_+ = x + t \Delta x$ instead of the expression $x_{k+1} = x_k + t_k \Delta x_k$.

5.5.1 Exact line search

In this subsection, we will study the convergence speed of the gradient method with exact line search.

We define $\tilde{f} : \mathbb{R} \to \mathbb{R}$, with $\tilde{f}(t) = f(x - t \nabla f(x))$. From (5.29), for $y = x - t \nabla f(x)$, we get

$$\tilde{f}(t) \leq f(x) - t \| \nabla f(x) \|_2^2 + \frac{Mt^2}{2} \| \nabla f(x) \|_2^2. \quad (5.35)$$

Next, we apply exact line search and minimize, with respect to $t$, both sides of (5.35).

The optimal value of the left side equals $\tilde{f}(t_{\text{exact}})$. The right-hand side is a quadratic function of $t$, which is minimized for $t = \frac{1}{2M}$, and has minimum value equal to $f(x) - \frac{1}{2M} \| \nabla f(x) \|_2^2$ (prove it). Therefore,

$$f(x_+) := \tilde{f}(t_{\text{exact}}) \leq f(x) - \frac{1}{2M} \| \nabla f(x) \|_2^2. \quad (5.36)$$

Subtracting $p_*$ from both sides, we get

$$f(x_+) - p_* \leq f(x) - p_* - \frac{1}{2M} \| \nabla f(x) \|_2^2. \quad (5.37)$$

From (5.30), we get

$$\| \nabla f(x) \|_2^2 \geq 2m(f(x) - p_*). \quad (5.38)$$

Combining (5.37) and (5.38), we get

$$f(x_+) - p_* \leq \left(1 - \frac{m}{M}\right) (f(x) - p_*). \quad (5.39)$$

Defining

$$c := 1 - \frac{m}{M}, \quad (5.40)$$

and applying the above inequality recursively, we obtain

$$f(x_k) - p_* \leq c^k (f(x_0) - p_*). \quad (5.41)$$

We observe that $0 \leq c < 1$. Therefore, the above relation proves that $f(x_k) \to p_*$, when $k \to \infty$. 
In particular, in order to derive a sufficient condition ensuring that we have achieved accuracy $\epsilon > 0$, we work as follows. Let

$$c^k(f(x_0) - p^*) \leq \epsilon. \quad (5.42)$$

Then

$$\frac{f(x_0) - p^*}{\epsilon} \leq \frac{1}{c^k} = \left(\frac{1}{c}\right)^k$$

$$\iff \log \left(\frac{f(x_0) - p^*}{\epsilon}\right) \leq k \log \left(\frac{1}{c}\right) \quad (5.43)$$

$$\iff k \geq \frac{\log \left(\frac{f(x_0) - p^*}{\epsilon}\right)}{\log \left(\frac{1}{c}\right)}. \quad (5.44)$$

Considering (5.41), we conclude that if

$$k \geq k_\epsilon := \frac{\log \left(\frac{f(x_0) - p^*}{\epsilon}\right)}{\log \left(\frac{1}{c}\right)}, \quad (5.44)$$

then $f(x_k) - p^* \leq \epsilon$.

We define as condition number of the problem the quantity

$$\mathcal{K} := \frac{M}{m}. \quad (5.45)$$

For large $\mathcal{K}$, we have that\(^1\)

$$\log \left(\frac{1}{c}\right) = -\log \left(1 - \frac{m}{M}\right) \approx \frac{m}{M} = \frac{1}{\mathcal{K}}. \quad (5.46)$$

In this case,

$$k_\epsilon \approx \mathcal{K} \log \left(\frac{f(x_0) - p^*}{\epsilon}\right), \quad (5.47)$$

that is, the maximum required number of iterations to get $\epsilon$-close to the optimal $p^*$ value increases linearly with the condition number of the problem. Moreover, it depends logarithmically on the starting point, $x_0$, and on the accuracy, $\epsilon$.

Another interpretation of (5.47) is as follows: to increase the accuracy of the solution $\epsilon$ by one decimal point, we must perform additional (at most) $O(1)\mathcal{K}$ iterations (why?).

\(^1\)Recall that $\log(1 - x) = -x + O(x^2)$. 


5.5.2 Another interpretation of the gradient descent method

Suppose that we are at the point $x_k$. The second-order approximation of $f$ around this point is as follows:

$$f(x) \approx f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2} (x - x_k)^T \nabla^2 f(x_k)(x - x_k). \quad (5.48)$$

If, in the above relationship, instead of the Hessian $\nabla^2 f(x_k)$ we put the matrix $\frac{1}{t}I$, we get

$$f(x) \approx f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2t} \|x - x_k\|^2 =: g(x). \quad (5.49)$$

Function $g(x)$ is a quadratic convex function of $x$, with derivative

$$\nabla g(x) = \frac{1}{t}(x - x_k) + \nabla f(x_k). \quad (5.50)$$

The point which minimizes $g(x)$ is

$$x^* = x_k - t\nabla f(x_k). \quad (5.51)$$

Therefore, the iteration of the gradient descent method at the point $x_k$ can be interpreted as the optimization of a simple quadratic approximation of $f$ around the point $x_k$. The smaller $t$ is, the larger the contribution of the second-order term, which means that the optimal point will be near $x_k$.

5.6 Newton method

Next, we describe an extremely important method for solving systems of nonlinear equations and optimization problems, the Newton method.

5.6.1 Newton step

The direction of movement

$$\Delta x_{Nt} = -\left(\nabla^2 f(x)\right)^{-1} \nabla f(x) \quad (5.52)$$

is called Newton step.

If $\nabla^2 f(x) \succ 0$, then

$$\nabla f(x)^T \Delta x_{Nt} = -\nabla f(x)^T \left(\nabla^2 f(x)\right)^{-1} \nabla f(x) \leq 0, \quad (5.53)$$
with equality if, and only if, \( \nabla f(x) = 0 \). Thus, the Newton step is a descent direction of \( f \) at the point \( x \).

The Newton step can be interpreted as follows.

1. *Second-order approximation minimization.* The second-order Taylor approximation of \( f \) at the point \( x \) is given by the relation

   \[
   \hat{f}_x(y) = f(x) + \nabla f(x)^T (y - x) + \frac{1}{2} (y - x)^T \nabla^2 f(x) (y - x). \quad (5.54)
   \]

   This quadratic function is minimized for \( y = x + \Delta x_N \) (see Figure 5.3) (prove it).

2. *Method of steepest descent, in terms of the Hessian norm.* The Newton step is the direction of steepest descent of \( f \) at \( x \) for the norm

   \[
   \|u\|_{\nabla^2 f(x)} = \left( u^T \nabla^2 f(x) u \right)^{1/2}. \quad (5.55)
   \]

3. *Solution of a linearized optimality condition.* The linear approximation of the equation \( \nabla f(x_*) = 0 \), around point \( x \), gives

   \[
   \nabla f(x + v) \approx \nabla f(x) + \nabla^2 f(x) v = 0, \quad (5.56)
   \]

   which gives \( v = - (\nabla^2 f(x))^{-1} \nabla f(x) \).
\( x \in \mathbb{R}^n, \epsilon > 0. \)

While (TRUE)
1. \( \Delta x_{Nt} := -(\nabla^2 f(x))^{-1} \nabla f(bfx). \)
2. \( \lambda^2 := \nabla f(x)^T (\nabla^2 f(x))^{-1} \nabla f(x). \)
3. quit if \( \frac{\lambda^2}{2} \leq \epsilon. \)
4. Perform backtracking line search and compute \( t. \)
5. \( x := x + t\Delta x_{Nt}. \)

Table 5.3: Newton method.

We present the Newton method in Table 5.3.

### 5.6.2 Newton decrement

The quantity
\[
\lambda(x) := \left( \nabla f(x)^T (\nabla^2 f(x))^{-1} \nabla f(x) \right)^{\frac{1}{2}}
\]
(5.57)
is called **Newton decrement**. It plays an important role in the analysis of the Newton method and can be used in the terminating condition of the algorithm.

If \( \hat{f}_x(y) \) is the second-order approximation of \( f \), at the point \( x \), that is,
\[
\hat{f}_x(y) = f(x) + \nabla f(x)^T (y - x) + \frac{1}{2} (y - x)^T \nabla^2 f(x)(y - x),
\]
(5.58)then
\[
f(x) - \inf_y \hat{f}_x(y) = f(x) - \hat{f}_x(x + \Delta x_{Nt})
= -\nabla f(x)^T \Delta x_{Nt} - \frac{1}{2} \Delta x_{Nt}^T \nabla^2 f(x) \Delta x_{Nt}
= \nabla f(x)^T (\nabla^2 f(x))^{-1} \nabla f(x) - \frac{1}{2} \nabla f(x)^T (\nabla^2 f(x))^{-1} \nabla f(x)
= \frac{1}{2} \lambda^2(x).
\]
(5.59)

That is, the quantity \( \frac{\lambda^2(x)}{2} \) is an estimate for the quantity \( f(x) - p_x \), based on the quadratic approximation of \( f \) at the point \( x \).

In addition, the Newton decrement appears in backtracking line search because
\[
-\lambda^2(x) = \nabla f(x)^T \Delta x_{Nt} = \left. \frac{d}{dt} f(x + t\Delta x_{Nt}) \right|_{t=0}.
\]
(5.60)
5.6.3 Local convergence of the Newton algorithm

In this subsection, we will study the convergence of the Newton algorithm assuming it starts from a point that is close enough to the optimal point.

The result we will prove states that, locally, the algorithm Newton has quadratic convergence.

In practice, this means that, if we start from a point which is sufficiently close to the solution, then the number of digits of the solution’s accuracy is doubled at each iteration.

A full analysis of the convergence of the Newton algorithm is given in the Appendix of the chapter.

Technical background

Before proceeding to the analysis of the Newton algorithm, we introduce some notation which we will use next.

Let $g = (g_1, \ldots, g_n) : \mathbb{R} \to \mathbb{R}^n$. We define

$$
\int_a^b g(t) dt := \begin{bmatrix}
\int_a^b g_1(t) dt \\
\vdots \\
\int_a^b g_n(t) dt
\end{bmatrix}.
$$

(5.61)

If $A : \mathbb{R} \to \mathbb{R}^{m \times n}$, then we define

$$
\int_a^b A(t) g(t) dt := \begin{bmatrix}
\int_a^b a_1^T(t) g(t) dt \\
\vdots \\
\int_a^b a_m^T(t) g(t) dt
\end{bmatrix} = \begin{bmatrix}
\sum_{i=1}^n \int_a^b a_{1,i}(t) g_i(t) dt \\
\vdots \\
\sum_{i=1}^n \int_a^b a_{m,i}(t) g_i(t) dt
\end{bmatrix}.
$$

(5.62)

Furthermore, notice that we can write

$$
c = \int_0^1 c dt, \quad Ac = \int_0^1 Ac dt.
$$

(5.63)

We define the norm

$$
\left\| \int_a^b g(t) dt \right\|_2 = \sqrt{\sum_{i=1}^n \left( \int_a^b g_i(t) dt \right)^2}.
$$

(5.64)

The following inequalities will be useful later

$$
\left\| \int_a^b A(t) g(t) dt \right\|_2 \leq \int_a^b \|A(t)g(t)\|_2 dt \leq \int_a^b \|A(t)\|_2 \|g(t)\|_2 dt.
$$

(5.65)
The left inequality is a generalization of the triangle inequality, while the right inequality results from an application of the relationship
\[ \|Ag\|_2 \leq \|A\|_2\|g\|_2, \quad \text{for} \ A \in \mathbb{R}^{m \times n}, \ g \in \mathbb{R}^n. \] (5.66)

At this point, we have all the technical tools at our disposal to proceed with the local convergence analysis of the Newton algorithm.

**Theorem 5.6.1.** Let \( f : \mathbb{R}^n \to \mathbb{R} \) be doubly continuously differentiable. Let us assume that

1. \( \exists m > 0 \) such that \( \nabla^2 f(x) \succeq mI \), for every \( x \in \mathbb{R}^n \). Notice that this assumption leads to the inequality
   \[ (\nabla^2 f(x))^{-1} \preceq \frac{1}{m} I, \quad \text{for any} \ x \in \mathbb{R}^n. \]

2. \( \exists L > 0 \) such that \( \|\nabla^2 f(x) - \nabla^2 f(y)\|_2 \leq L\|x - y\|_2 \), for any \( x, y \in \mathbb{R}^n \).

Let \( \{x_k\} \) be the sequence produced by the Newton method and \( x^* \) the optimal point. Then, for \( k = 0, 1, \ldots \), the following inequality holds true
\[ \|x_{k+1} - x^*\|_2 \leq \frac{L}{2m}\|x_k - x^*\|^2_2. \] (5.67)

Moreover, if \( \|x_0 - x^*\|_2 \leq \frac{m}{L} \), then, for \( k = 0, 1, \ldots \), we have that
\[ \|x_k - x^*\|_2 \leq \frac{2m}{L} \left( \frac{1}{2} \right)^{2^k}. \] (5.68)

**Proof.** We study the algorithm with \( t_k = 1 \). From the definition of the algorithm, we have
\[
x_{k+1} - x^* = x_k - (\nabla^2 f(x_k))^{-1} \nabla f(x_k) - x^* \\
\overset{(!!)}{=} x_k - x^* + (\nabla^2 f(x_k))^{-1} (\nabla f(x^*) - \nabla f(x_k)) \\
= x_k - x^* + (\nabla^2 f(x_k))^{-1} \int_0^1 \nabla^2 f(x_k + t(x^* - x_k)) (x^* - x_k) dt \\
= (\nabla^2 f(x_k))^{-1} \int_0^1 [\nabla^2 f(x_k + t(x^* - x_k)) - \nabla^2 f(x_k)] (x^* - x_k) dt,
\]
where at point (!!) we worked as follows. We define the function \( g : \mathbb{R} \to \mathbb{R}^n \), with
\[ g(t) := \nabla f(x_k + t(x^* - x_k)). \]
If $Dg$ is the derivative of $g$ (note that $Dg : \mathbb{R} \to \mathbb{R}^n$), then
\[ g(1) - g(0) = \int_0^1 Dg(t)dt. \quad (5.70) \]
Using the chain rule, we get
\[ Dg(t) = D\nabla f(x_k + t(x_* - x_k)) \]
\[ = \nabla^2 f(x_k + t(x_* - x_k))D(x_k + t(x_* - x_k)) \]
\[ = \nabla^2 f(x_k + t(x_* - x_k))(x_* - x_k). \quad (5.71) \]
Therefore, (5.70) is written as
\[ \nabla f(x_*) - \nabla f(x_k) = \int_0^1 \nabla^2 f(x_k + t(x_* - x_k))(x_* - x_k)dt. \quad (5.72) \]
This relationship was used at point (!!).

From (5.69), we get
\[ \|x_{k+1} - x_*\|_2 \]
\[ \leq \| (\nabla^2 f(x_k))^{-1} \|_2 \]
\[ \left\| \int_0^1 [\nabla^2 f(x_k + t(x_* - x_k)) - \nabla^2 f(x_k)](x_* - x_k)dt \right\|_2 \]
\[ \leq \| (\nabla^2 f(x_k))^{-1} \|_2 \]
\[ \int_0^1 \| [\nabla^2 f(x_k + t(x_* - x_k)) - \nabla^2 f(x_k)](x_* - x_k)\|_2 dt \]
\[ \leq \| (\nabla^2 f(x_k))^{-1} \|_2 \]
\[ \int_0^1 \| \nabla^2 f(x_k + t(x_* - x_k)) - \nabla^2 f(x_k)\|_2 \| (x_* - x_k)\|_2 dt \]
\[ \leq \frac{1}{m} \int_0^1 Lt\| x_k - x_* \|_2^2 dt \]
\[ = \frac{L}{2m}\| x_k - x_* \|_2^2, \]
where, at point (*), we used the two assumptions. The first part of the theorem is proved. The second part is proved by induction. For $k = 0$, we assume that
\[ \|x_0 - x_*\|_2 \leq \frac{m}{L} = \frac{2m}{L} \left( \frac{1}{2} \right)^{2^0}. \quad (5.73) \]
5.6. NEWTON METHOD

Let the assumption hold for some \( k \), that is,

\[
\|x_k - x_\ast\|_2 \leq \frac{2m}{L} \left( \frac{1}{2} \right)^{2^k}.
\]  

(5.74)

We will prove that it holds for \( k + 1 \). From the first part of the theorem, we have that

\[
\|x_{k+1} - x_\ast\|_2 \leq \frac{L}{2m} \|x_k - x_\ast\|_2^2
\]

\[
\leq \frac{L}{2m} \left( \frac{2m}{L} \left( \frac{1}{2} \right)^{2^k} \right)^2
\]

\[
= \frac{2m}{L} \left( \frac{1}{2} \right)^{2^{k+1}},
\]  

(5.75)

completing the proof. \( \square \)
Chapter 6

Optimality Conditions

In this chapter, we will derive necessary and sufficient conditions characterizing optimal solutions of convex optimization problems. First, we will prove Farkas's Lemma and use it to prove the Fritz John conditions. Then, making an additional assumption, we will prove the Karush-Kuhn-Tucker conditions.

Our approach, which is primarily geometric, is different from that of Boyd and Vanderberghe,\(^1\) and is based on material from the book by Bazaraa, Sherali, Shetty and notes by M. Epelman (available online) (see also Chapters 10 and 11 of the book by A. Beck).

The optimality conditions (FJ, KKT) are extremely important because

1. they offer geometric interpretation and deeper understanding of the problem,

2. important constrained optimization algorithms search the optimal point by, essentially, searching for the point which satisfies the optimality conditions.

6.1 Necessary optimality conditions

Consider the convex optimization problem

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& \quad h_i(x) = 0, \quad i = 1, \ldots, p,
\end{align*}
\]

(6.1)

with \( h_i(x) = a_i^T x - b_i \), for \( i = 1, \ldots, p \).

\(^1\)Which can be seen as complementary to the geometric approach and deserves independent study.
We assume that the functions $f_i$, for $i = 0, \ldots, m$, are differentiable, with $f_i : \text{dom} f_i \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, with $\text{dom} f_i$ an open convex set.

Equality constraints can be expressed in the form
\[ Ax = b, \quad (6.2) \]
with
\[
A := \begin{bmatrix}
    a_1^T \\
    \vdots \\
    a_p^T
\end{bmatrix}, \quad b := \begin{bmatrix}
    b_1 \\
    \vdots \\
    b_p
\end{bmatrix}.
\quad (6.3)
\]

Let $D := \bigcap_{i=0}^m \text{dom} f_i$.

The feasible set of problem (6.1) is the set
\[ X := \{ x \in D \mid f_i(x) \leq 0, \ i = 1, \ldots, m, \ Ax = b \}. \quad (6.4) \]

**Definition 6.1.1.** Let $x \in X$.

1. The set $F_0 := \{ d \in \mathbb{R}^n \mid \nabla f_0(x)^T d < 0 \}$ is called cone of descent directions of $f_0$ at the point $x$.

2. The set $I := \{ i \in \{1, \ldots, m\} \mid f_i(x) = 0 \}$ is called set of indices of the active inequality constraints of problem (6.1) at the point $x$.

3. The set $G_0 := \{ d \in \mathbb{R}^n \mid \nabla f_i(x)^T d < 0, \ \text{for} \ i \in I \}$ is called cone of “interior” directions of active inequality constraints of problem (6.1) at the point $x$.

4. The set $H_0 := \{ d \in \mathbb{R}^n \mid Ad = 0 \}$ is called set of allowed directions of equality constraints of problem (6.1) at the point $x$.

The first important result for the characterization of optimal solutions of problem (6.1) is as follows.

**Theorem 6.1.1.** If $x \in X$ is an optimal point of the problem (6.1), then
\[ F_0 \cap G_0 \cap H_0 = \emptyset. \quad (6.5) \]
6.2. FARKAS’S LEMMA

Proof. The proof will be made by reductio ad absurdum.

Let \( x \in X \) be an optimal point for the problem (6.1) and \( F_0 \cap G_0 \cap H_0 \neq \emptyset \). Let \( d \in F_0 \cap G_0 \cap H_0 \) and \( \theta > 0 \). Then, for all sufficiently small \( \theta \) and \( i \in I \), we have that

\[
    f_i(x + \theta d) < f_i(x) = 0 \quad \text{and} \quad A(x + \theta d) = b.
\]

Furthermore, since for \( i \not\in I \) we have that \( f_i(x) < 0 \), then, for \( i \not\in I \) and sufficiently small \( \theta \), due to the continuity of \( f_i \), we will have that \( f_i(x + \theta d) < 0 \).

That is, for sufficiently small \( \theta \), the points \( x + \theta d \) are feasible points of the problem (6.1).

At the same time, however, for sufficiently small \( \theta \), we have that \( f_0(x + \theta d) < f_0(x) \), which is false, because we have assumed that the point \( x \) is optimal.

Theorem 6.1.1 indicates that, if the point \( x \in X \) is optimal for the problem (6.1), then there is no strictly feasible direction \( d \) which, at the same time, is decreasing direction of \( f_0 \).

6.2 Farkas’s Lemma

Next, we will translate the geometric condition (6.5) into an algebraic relationship between the gradients of the cost and constraint functions.

First, we prove the best-known result from a class of results that are known as Theorems of Alternatives.

Lemma 6.2.1. Farkas’ Lemma. Let \( A \in \mathbb{R}^{m \times n} \) and \( c \in \mathbb{R}^n \). Then, exactly one of the following systems has a solution:

\[
\begin{align*}
(1) \quad & Ax \leq 0, \quad c^T x > 0, \\
(2) \quad & A^T y = c, \quad y \geq 0.
\end{align*}
\]

Proof. Assume that system (2) has a solution, that is, there exists \( y \geq 0 \) such that so that \( A^T y = c \).

Next, we consider the system (1) and more specifically the inequality \( Ax \leq 0 \). The inequality is satisfied for \( x = 0 \). But this solution does not satisfy the inequality \( c^T x > 0 \). If there is no \( x \neq 0 \) satisfying this inequality, then the system (1) has no solution. If there
exists \( x \neq 0 \) which satisfies this inequality, then, since \( y \geq 0 \) and \( Ax \leq 0 \), we get that \( c^T x = y^T Ax \leq 0 \). Therefore, in this case too, system (1) has no solution.

Next, we assume that system (2) has no solution.

We define the set \( \mathbb{M} := \{ x \in \mathbb{R}^n \mid x = A^T y, y \geq 0 \} \). We know that the set \( \mathbb{M} \) is a closed convex cone, because if

\[
A := \begin{bmatrix}
a_1^T \\
\vdots \\
a_m^T
\end{bmatrix},
\]

then \( \mathbb{M} = \{ \sum_{i=1}^m y_i a_i \mid y_i \geq 0, \text{ for } i = 1, \ldots, m \} \). That is, \( \mathbb{M} \) is the conic hull of \( a_1, \ldots, a_m \).

In addition, from the assumption about the non existence of a solution for the system (2), we know that \( c \notin \mathbb{M} \). Therefore, there exists a hyperplane which separates the point \( c \) and the convex set \( \mathbb{M} \). That is, there exist \( p \in \mathbb{R}^n \) and \( \alpha \in \mathbb{R} \) such that \( p^T c > \alpha \) and \( p^T x \leq \alpha \) for every \( x \in \mathbb{M} \). Then

1. Since \( 0 \in \mathbb{M} \), we have that \( 0 = p^T 0 \leq \alpha \) and \( p^T c > \alpha \geq 0 \), so, \( p^T c > 0 \).

2. We recall that for each \( x \in \mathbb{M} \), we have that \( \alpha \geq p^T x \). Furthermore, \( x \in \mathbb{M} \) implies that \( x = A^T y \), with \( y \geq b f 0 \). Therefore, for every \( y \geq 0 \), we have

\[
\alpha \geq p^T x = p^T A^T y = y^T A p.
\]
6.2. FARKAS’S LEMMA

But, since \( y \geq 0 \) can become arbitrarily large, the inequality \( \alpha \geq y^T A p \), for each \( y \geq 0 \), it follows that \( A p \leq 0 \).

This happens because, if some element of \( A p \) is greater than zero, let \( (A p)_i \), then we can choose a \( y \) which has zeros everywhere except the \( i \)-th position where it has arbitrarily large \( y_i \), so that the inequality \( \alpha \geq y^T A p \) is not satisfied.

Therefore, we constructed a vector \( p \) such that so that \( A p \leq 0 \) and \( c^T p > 0 \).

So, we found a solution for the system (1), completing the proof.

The following result is an important application of the Farkas’s Lemma.

**Lemma 6.2.2. (Basic Lemma)** Exactly one of the following two systems has a solution.

\[
\begin{align*}
(1) \quad A x < 0, & \quad B x \leq 0, \quad H x = 0 \\
(2) \quad A^T u + B^T w + H^T v = 0, & \quad u \geq 0, \quad w \geq 0, \quad 1^T u = 1,
\end{align*}
\]  

where \( 1 := (1, \ldots, 1) \).

**Proof.** The system (1) is equivalently expressed as

\[
\begin{align*}
A x + e 1 \leq 0, & \quad e > 0 \\
B x \leq 0 \\
H x \leq 0 \\
-H x \leq 0.
\end{align*}
\]

A more condensed form of it is as follows:

\[
\begin{bmatrix}
A & 1 \\
B & 0 \\
H & 0 \\
-H & 0
\end{bmatrix}
\begin{bmatrix}
x \\
e
\end{bmatrix}
\leq 0,
\begin{bmatrix}
0 & \cdots & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x \\
e
\end{bmatrix}
> 0.
\]

The system (6.10) has exactly the form of system (1) of Lemma 6.2.1.

The system corresponding to the system (2) of Lemma 6.2.1 has the form

\[
\begin{bmatrix}
A & 1 \\
B & 0 \\
H & 0 \\
-H & 0
\end{bmatrix}^T
\begin{bmatrix}
u \\
w \\
v_1 \\
v_2
\end{bmatrix}
= \begin{bmatrix}
0 \\
\vdots \\
0 \\
1
\end{bmatrix}, \quad (u, w, v_1, v_2) \geq 0,
\]

where \( 1 := (1, \ldots, 1) \).
or, equivalently,
\[ A^T u + B^T w + H^T (v_1 - v_2) = 0, \quad 1^T u = 1, \quad (u, w, v_1, v_2) \geq 0. \] (6.12)

Setting \( v := v_1 - v_2 \), we complete the proof. \( \square \)

Note: The relation \( 1^T u = 1 \) essentially implies that \( u \neq 0 \). Therefore, in the statement of the theorem, the relation \( 1^T u = 1 \) can be replaced by the relation \( u \neq 0 \).

### 6.3 First-order optimality conditions

Next, we use Lemma 6.2.2 and prove important algebraic optimality conditions for the optimization problem (6.1).

**Theorem 6.3.1. (Fritz John necessary conditions)** Let \( x \in X \) be an optimal point for the problem (6.1). Then, there exists a vector \((u_0, u, v)\) such that
\[
\begin{align*}
&u_0 \nabla f_0(x) + \sum_{i=1}^m u_i \nabla f_i(x) + A^T v = 0, \\
&(u_0, u) \geq (0, 0), \quad (u_0, u) \neq (0, 0), \\
&u_i f_i(x) = 0, \quad i = 1, \ldots, m.
\end{align*}
\] (6.13)

**Proof.** Without loss of generality, we assume \( I = \{1, \ldots, l\} \) and define
\[
\mathcal{A} := \begin{bmatrix}
\nabla f_0(x)^T \\
\nabla f_1(x)^T \\
\vdots \\
\nabla f_l(x)^T
\end{bmatrix}.
\] (6.14)

Since \( x \in X \) is an optimal point for the problem (6.1), from Theorem 6.1.1, we conclude that \( F_0 \cap G_0 \cap H_0 = \emptyset \).

That is, there is no \( d \) which satisfies the relations \( \mathcal{A}d < 0 \) and \( \mathcal{A}d = 0 \). Therefore, from Lemma 6.2.2 for \( B = 0 \), we conclude that there exist vectors \((u_0, u_1, \ldots, u_l)\) and \( v \) such that
\[
u_0 \nabla f_0(x) + \sum_{i=1}^l u_i \nabla f_i(x) + A^T v = 0,
\] (6.15)
with \((u_0, u_1, \ldots, u_l) \geq 0\) and \((u_0, u_1, \ldots, u_l) \neq 0\). We set \(u_{l+1} = \cdots = u_m = 0\).

Therefore, we get that \((u_0, u) \geq (0, 0), (u_0, u) \neq (0, 0)\) and \(u_i f_i(x) = 0\), for \(i = 1, \ldots, m\). The proof is complete.

Next, we use an additional assumption and prove the necessary KKT conditions.

**Theorem 6.3.2. (Necessary conditions KKT)** Let \(x \in X\) be an optimal point of the optimization problem (6.1). Moreover, assume that the gradients of all active constraints of problem (6.1)\(^2\) at the point \(x\) are linearly independent. Then, there exist vectors \(u\) and \(v\) such that

\[
\nabla f_0(x) + \sum_{i=1}^{m} u_i \nabla f_i(x) + A^T v = 0, \\
u \geq 0, \quad u_i f_i(x) = 0, \quad \text{for } i = 1, \ldots, m.
\]

(6.16)

**Proof.** We have assumed that \(x\) is an optimal point of the problem (6.1). Therefore, it must satisfy the Fritz John conditions of (6.13).

If \(u_0 > 0\), then we can re-parameterize the Fritz John conditions as follows: \(u_i := \frac{u_i}{u_0}\), for \(i = 0, \ldots, m\), and \(v_i := \frac{v_i}{u_0}\), for \(i = 1, \ldots, p\), taking (6.16).

If \(u_0 = 0\), then there exist \(u \geq 0\), with \(u \neq 0\), and \(v\) such that

\[
\sum_{i \in I} u_i \nabla f_i(x) + \sum_{i=1}^{p} v_i a_i = 0.
\]

(6.17)

This, however, is impossible, because we have assumed that the gradients of the active constraints at point \(x\) are linearly independent. The proof is complete.

Next, we list sufficient optimality conditions.

**Theorem 6.3.3. (Sufficient Conditions KKT)** Let the point \(x \in X\) satisfy the Karush-Kuhn-Tucker conditions, that is, for some \(u \in \mathbb{R}^m\) and \(v \in \mathbb{R}^p\), the following relations hold:

\[
\nabla f_0(x) + \sum_{i=1}^{m} u_i \nabla f_i(x) + A^T v = 0, \\
u \geq 0, \quad u_i f_i(x) = 0, \quad \text{for } i = 1, \ldots, m.
\]

(6.18)

\(^2\)As active constraints of the problem are defined the active inequality constraints and all the equality constraints.
Then, \( x \) is an optimal point for the problem (6.1).

**Proof.** We know that the feasible set \( X \) of the problem (6.1) is convex. Let \( \hat{x} \in X \) be a point different from \( x \). The points \( \theta \hat{x} + (1 - \theta)x \), with \( 0 \leq \theta \leq 1 \), are feasible. This means that, for every \( i \in I \) and \( 0 \leq \theta \leq 1 \),

\[
f_i(\theta \hat{x} + (1 - \theta)x) = f_i(x + \theta(\hat{x} - x)) \leq 0 = f_i(x).
\]

(6.19)

Since the value of \( f_i \) does not increase when moving from \( x \) along the direction \( \hat{x} - x \), we have that, for each \( i \in I \),

\[
\nabla f_i(x)^T(\hat{x} - x) \leq 0.
\]

(6.20)

Also, since \( \hat{x}, x \in X \), we have that

\[
A(\hat{x} - x) = 0.
\]

(6.21)

Using (6.18), (6.20), and (6.21), we obtain that, \( \forall \hat{x} \in X \),

\[
\nabla f_0(x)^T(\hat{x} - x) = -\left( \sum_{i=1}^{m} u_i \nabla f_i(x)^T + v^TA \right)(\hat{x} - x)
\]

\[
= -\sum_{i=1}^{m} u_i \nabla f_i(x)^T(\hat{x} - x) + v^TA(\hat{x} - x)
\]

\[
\geq 0.
\]

(6.22)

We recall that a sufficient and necessary condition for the point \( x^* \) to be a solution for the problem \( \min_{x \in X} f_0(x) \) is as follows:

\[
\nabla f_0(x^*)^T(y - x^*) \geq 0, \ \forall \ y \in X.
\]

(6.23)

Therefore, from (6.22), we conclude that, for any \( \hat{x} \in X \), \( f_0(\hat{x}) \geq f_0(x) \). That is, \( x \) is an optimal point of the problem (6.1).

\[\square\]

### 6.4 Constraint qualification

In Theorem 6.3.2, we assumed that

1. the point \( x \) is optimal for problem (6.1),

2. “a certain requirement” is satisfied by the constraints,
and we proved the necessity of the KKT conditions.

The extra “requirement,” which allows us to prove the KKT conditions, is called constraint qualification.

We have already proved that if the gradients of the active constraints of the problem (6.1) at the optimal point $\mathbf{x}$ are linearly independent, then the KKT conditions hold true.

However, this requirement is local, because it only concerns the optimal point and can be checked only if we have in advance some optimal point in mind.

An extremely useful global condition is the following.

**Definition 6.4.1. (Slater Condition)** Consider the optimization problem (6.1). The Slater condition is satisfied if there exists $\bar{\mathbf{x}} \in \mathbb{X}$ such that $f_i(\bar{\mathbf{x}}) < 0$, for $i = 1, \ldots, m$.

**Theorem 6.4.1. (Slater condition and necessity of the KKT conditions)** Let the Slater condition be satisfied for the optimization problem (6.1). Then, the KKT conditions are necessary to characterize the optimal point of the problem, that is, if $\mathbf{x} \in \mathbb{X}$ is an optimal point of the problem (6.1), then it satisfies the KKT conditions.

*Proof.* Let $\mathbf{x} \in \mathbb{X}$ be an optimal point of the problem (6.1). Then, the Fritz John conditions hold, i.e., there is a vector $(u_0, \mathbf{u}, \mathbf{v})$, with $(u_0, \mathbf{u}) \geq \mathbf{0}$ and $(u_0, \mathbf{u}) \neq \mathbf{0}$, such that

\[
u_0 \nabla f_0(\mathbf{x}) + \sum_{i=1}^{m} u_i \nabla f_i(\mathbf{x}) + \mathbf{A}^T \mathbf{v} = \mathbf{0},
\]

with $u_i f_i(\mathbf{x}) = 0$, for $i = 1, \ldots, m$. If $u_0 \neq 0$, then we can divide by $u_0$ and get the KKT conditions.

Assume that $u_0 = 0$. Since we have assumed that the Slater condition is satisfied, there exists $\bar{\mathbf{x}} \in \mathbb{X}$ such that, for each $i \in I$,

\[0 = f_i(\mathbf{x}) > f_i(\bar{\mathbf{x}}).
\]

From the convexity of $f_i$, we have that

\[f_i(\bar{\mathbf{x}}) \geq f_i(\mathbf{x}) + \nabla f_i(\mathbf{x})^T (\bar{\mathbf{x}} - \mathbf{x}).
\]

From the relations (6.25) and (6.26), we conclude that, for every $i \in I$, the following inequality holds

\[\nabla f_i(\mathbf{x})^T (\bar{\mathbf{x}} - \mathbf{x}) < 0.
\]

\footnote{Obviously, since $\mathbf{x} \in \mathbb{X}$, the relation $\mathbf{A} \mathbf{x} = \mathbf{b}$ will also be satisfied.}
Furthermore, since $\bar{x}$ and $x$ are feasible points, we have that $A(\bar{x} - x) = 0$. From (6.24) and the assumption $u_0 = 0$, we get that one (or some) of $u_i$, for $i \in \mathbb{I}$, should be positive, therefore

$$0 = 0^T(\bar{x} - x) = \left( \sum_{i=1}^{m} u_i \nabla f_i(x)^T + v^T A \right)(\bar{x} - x)$$

$$= \sum_{i \in \mathcal{I}} u_i \nabla f_i(x)^T (\bar{x} - x)$$

$$< 0,$$

which is false. So, if the Slater condition holds, then we must have that $u_0 > 0$, completing the proof. □

Figure 6.2: KKT conditions. (a) Feasible set of inequality constraints $\mathbb{X} = \{x \in \mathbb{R}^n : f_i(x) \leq 0, i = 1, 2, 3\}$. (b) Optimal point $x_*$ inside the feasible set (c) Optimal point $x_*$ with one active constraint (d) Optimal point $x_*$ with two active constraints.
6.5 Geometric interpretation for problems with inequality constraints

In order to give a geometric interpretation to the KKT conditions, we consider the convex optimization problem with only inequality constraints. Furthermore, we assume that Slater’s condition holds. In this case, the KKT conditions are sufficient and necessary optimality conditions. If \( I = \{1, \ldots, l\} \), then the point \( x \in X \) is optimal if, and only if, there exists a vector \( (u_1, \ldots, u_l) \) with \( (u_1, \ldots, u_l) \geq 0 \) such that

\[
\sum_{i=1}^{l} u_i \nabla f_i(x) = -\nabla f_0(x) \tag{6.29}
\]

That is, the point \( x \in X \) is optimal for the optimization problem if, and only if, the negative gradient of \( f_0 \) at the point \( x \), \( -\nabla f_0(x) \), lies in the cone generated by the gradients of the active inequality constraints at the point \( x \) (see Figure 6.2).

In Figure 6.3, we depict the optimization problem of the convex function \( f_0(x) \) under the affine equality constraint \( h_1(x) = a^T x - b = 0 \). Notice that \( \nabla f_0(x_*) \) is collinear with \( \nabla h_1(x_*) \) but can have the same or opposite direction (i.e., \( v \in \mathbb{R} \)).
6.6 Examples

First, we give an example where the conditions KKT *do not* hold.

**Example 6.6.1.** Consider the optimization problem

\[
\begin{aligned}
&\text{minimize} & & x_1 \\
&\text{subject to} & & (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1 \\
& & & (x_1 - 1)^2 + (x_2 + 1)^2 \leq 1.
\end{aligned}
\]  

(6.30)

The only feasible point is \( x = (1, 0) \). Prove that, at this point, the Fritz John conditions hold true but the KKT conditions do not hold true. Try to understand why this happens. Are the gradients of the active constraints independent or does the Slater condition hold true? \( \square \)

We continue with two very important examples.

**Example 6.6.2.** Let \( f(x) = \frac{1}{2} x^T P x + q^T x \) and consider the quadratic problem with affine equality constraints

\[
\begin{aligned}
&\text{minimize} & & f(x) \\
&\text{subject to} & & Ax = b.
\end{aligned}
\]  

(6.31)

The KKT equations are

\[
\begin{aligned}
\nabla f(x) &+ A^T v = 0 \\
Ax & = b
\end{aligned}
\]  

(6.32)

and can be expressed as the system of linear equations

\[
\begin{bmatrix}
P & A^T \\
A & O
\end{bmatrix}
\begin{bmatrix}
x \\
v
\end{bmatrix}
=
\begin{bmatrix}
-q \\
b
\end{bmatrix}.
\]  

(6.33)

If there exists a solution of the system (6.33), say, \((x_*, v_*)\), then \(x_*\) is called **primal optimal** and \(v_*\) **dual optimal**. \( \square \)

We will say more about this very important problem later.

**Example 6.6.3.** For data \( x_k \in \text{dom} f \), with \( \nabla f(x_k) \neq 0 \), \( A \succ 0 \), and \( t > 0 \), compute the solution of the problem

\[
\begin{aligned}
&\text{minimize} & & \nabla f(x_k)^T x \\
&\text{subject to} & & (x - x_k)^T A(x - x_k) \leq t^2.
\end{aligned}
\]  

(6.34)
The problem (6.34) can be written as

\[
\begin{align*}
\text{minimize} & \quad f_0(x) := \nabla f(x_k)^T x \\
\text{subject to} & \quad f_1(x) \leq 0,
\end{align*}
\] (6.35)

with \( f_1(x) := (x - x_k)^T A(x - x_k) - t^2 \).

The KKT conditions are

\[
\begin{align*}
(a) & \quad \nabla f(x_k) + 2 \lambda A(x_* - x_k) = 0 \\
(b) & \quad f_1(x_*) = (x_* - x_k)^T A(x_* - x_k) - t^2 \leq 0 \\
(c) & \quad \lambda \geq 0, \quad \lambda f_1(x_*) = 0.
\end{align*}
\] (6.36)

From the relation (a), we find that we must have \( \lambda > 0 \) (why?). From this, we conclude that \( f_1(x_*) = 0 \) (why?).

From (a), we get

\[
(x_* - x_k) = -\frac{1}{2 \lambda} A^{-1} \nabla f(x_k).
\] (6.37)

Substituting this value to (b) and noting that \( f_1(x_*) = 0 \), we have

\[
\frac{1}{4 \lambda^2} \nabla f(x_k)^T A^{-1} \nabla f(x_k) = t^2
\Rightarrow \lambda = \frac{1}{2t} \left( \nabla f(x_k)^T A^{-1} \nabla f(x_k) \right)^{\frac{1}{2}}.
\] (6.38)

Substituting this value of \( \lambda \) into (6.37), we obtain

\[
x_* = x_k - \frac{t}{(\nabla f(x_k)^T A^{-1} \nabla f(x_k))^\frac{1}{2}} A^{-1} \nabla f(x_k).
\] (6.39)

We observe that, indeed, \( f_1(x_*) = 0 \).

In addition, we can make the following important observations:

1. Setting \( A = I \), we get the step of the gradient descent method.

2. Setting \( A = \nabla^2 f(x_k) \), we obtain the step of the Newton method.

3. The solution would be exactly the same if instead of \( f_0(x) = \nabla f(x_k)^T x \) we had set \( f_0(x) = f(x_k) + \nabla f(x_k)^T (x - x_k) \) (why?), which is the first order Taylor approximation of \( f(x) \) around the point \( x_k \).
Therefore, we get a new interpretation of gradient descent and Newton methods, for unconstrained problems, as minimizations of the first order approximation of the cost function at point $x_k$, subject to an inequality constraint of suitably chosen norms in $\mathbb{R}^n$ (see Figure 6.6).

\hfill \square

**Example 6.6.4.** Let $S = \{x \in \mathbb{R}^2 | x_1^2 + x_2^2 \leq 1, x_1 \geq 0, x_2 \geq 0\}$. At least $f_0(x) = c^T x$, with (1) $c = [1 \ 1]^T$ and (2) $c = [-2 \ -1]^T$. Compute the point

$$ x_* = \arg \min_{x \in S} f_0(x). $$
We observe that $x_*$ is the solution of the convex problem (see Figure 6.7)

$$\begin{align*}
\text{minimize} & \quad f_0(x) = c^T x, \\
\text{subject to} & \quad f_1(x) = x_1^2 + x_2^2 - 1 \leq 0 \\
& \quad f_2(x) = -x_1 \leq 0 \\
& \quad f_3(x) = -x_2 \leq 0.
\end{align*}$$

(6.40)

The KKT conditions for this problem are as follows:

$$\nabla f_0(x_*) + \lambda_1 \nabla f_1(x_*) + \lambda_2 \nabla f_2(x_*) + \lambda_3 \nabla f_3(x_*) = 0$$

$$\lambda_i \geq 0, \quad i = 1, 2, 3,$$

$$f_i(x_*) \leq 0, \quad i = 1, 2, 3,$$

$$\lambda_i f_i(x_*) = 0, \quad i = 1, 2, 3,$$

(6.41)

with

$$\nabla f_0(x_*) = c, \quad \nabla f_1(x_*) = \begin{bmatrix} 2x_{*,1} \\ 2x_{*,2} \end{bmatrix}, \quad \nabla f_2(x_*) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \nabla f_3(x_*) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$  

1. First, we consider the case where $c = [1 1]^T$. Substituting into (6.45), we get

$$\begin{bmatrix} 2x_{*,1}\lambda_1 - \lambda_2 \\ 2x_{*,2}\lambda_1 - \lambda_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$$\lambda_i \geq 0, \quad i = 1, 2, 3,$$

$$f_i(x_*) \leq 0, \quad i = 1, 2, 3,$$

$$\lambda_i f_i(x_*) = 0, \quad i = 1, 2, 3.$$

(6.42)

If $\lambda_1 > 0$, then

$$x_{*,1} = \frac{-1 + \lambda_2}{2\lambda_1}, \quad x_{*,2} = \frac{-1 + \lambda_3}{2\lambda_1}.$$  

(6.43)

**Example 6.6.5.** Let $S = \{x \in \mathbb{R}^2 | x_1^2 + x_2^2 \leq 1, x_1 \geq 0, x_2 \geq 0\}$. At least $f_0(x) = c^T x$, with (1) $c = [1 1]^T$ and (2) $c = [-2 \ -1]^T$. Compute the point

$$x_* = \arg \min_{x \in S} f_0(x).$$

We observe that $x_*$ is the solution of the convex problem (see Figure 6.7)

$$\begin{align*}
\text{minimize} & \quad f_0(x) = c^T x, \\
\text{subject to} & \quad f_1(x) = x_1^2 + x_2^2 - 1 \leq 0 \\
& \quad f_2(x) = -x_1 \leq 0 \\
& \quad f_3(x) = -x_2 \leq 0.
\end{align*}$$

(6.44)
CHAPTER 6. OPTIMALITY CONDITIONS

\[ \nabla f_0(x_k) \]

\[ \{ x \mid f_0(x) = f_0(x_k) \} \]

Figure 6.6: Problem solution (6.34) for $A = I$ and $A = \nabla^2 f_0(x_k)$.

\[ c^T x = a_1 \]

\[ c^T x = a_2 \]

Figure 6.7: Feasible set and isosurfaces of $f_0(x)$, with $c = [1 \ 1]^T$, and $a_1 < a_2$. 
The KKT conditions for this problem are as follows:

\[
\nabla f_0(x^*) + \lambda_1 \nabla f_1(x^*) + \lambda_2 \nabla f_2(x^*) + \lambda_3 \nabla f_3(x^*) = 0
\]

\[
\lambda_i \geq 0, \quad i = 1, 2, 3,
\]

\[
f_i(x^*) \leq 0, \quad i = 1, 2, 3,
\]

\[
\lambda_i f_i(x^*) = 0, \quad i = 1, 2, 3,
\]

with

\[
\nabla f_0(x^*) = \mathbf{c}, \quad \nabla f_1(x^*) = \begin{bmatrix} 2x_{s,1} \\ 2x_{s,2} \end{bmatrix}, \quad \nabla f_2(x^*) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \nabla f_3(x^*) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.
\]

1. First, we consider the case where \( \mathbf{c} = [1 \ 1]^T \). Substituting into (6.45), we get

\[
\begin{bmatrix} 2x_{s,1} \lambda_1 - \lambda_2 \\ 2x_{s,2} \lambda_1 - \lambda_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}
\]

\[
\lambda_i \geq 0, \quad i = 1, 2, 3,
\]

\[
f_i(x^*) \leq 0, \quad i = 1, 2, 3,
\]

\[
\lambda_i f_i(x^*) = 0, \quad i = 1, 2, 3.
\]

If \( \lambda_1 > 0 \), then

\[
x_{s,1} = \frac{-1 + \lambda_2}{2\lambda_1}, \quad x_{s,2} = \frac{-1 + \lambda_3}{2\lambda_1}.
\]

For \( x^* \) to be feasible, \( x_{s,1} \geq 0 \) and \( x_{s,2} \geq 0 \) should hold true. For this to happen, we must have \( \lambda_2 \geq 1 \) and \( \lambda_3 \geq 1 \).

Since, however, it should be true that \( \lambda_i f_i(x^*) = 0 \), for \( i = 2, 3 \), we should have that \( x_{s,1} = x_{s,2} = 0 \).

But, in this case, we have that \( x_{s,1}^2 + x_{s,2}^2 - 1 < 0 \) and the relation \( \lambda_1 f_1(x^*) = 0 \) does not hold true.

So, we cannot have \( \lambda_1 > 0 \).

If \( \lambda_1 = 0 \), then, from the first equation of (6.46), we obtain that \( \lambda_2 = 1 \) and \( \lambda_3 = 1 \), which implies that \( x_{s,1} = x_{s,2} = 0 \). We observe that, if we set these values, we satisfy the KKT conditions.

So, the optimal solution is \( x^* = (0, 0) \).
For $x_*$ to be feasible, $x_{*,1} \geq 0$ and $x_{*,2} \geq 0$ should hold true. For this to happen, should $\lambda_2 \geq 1$ and $\lambda_3 \geq 1$.

Since, however, it should be true that $\lambda_i f_i(x_*) = 0$, for $i = 2, 3$, we should have that $x_{*,1} = x_{*,2} = 0$.

But, in this case, we have that $x_{*,1}^2 + x_{*,2}^2 - 1 < 0$ and not the relation $\lambda_1 f_1(x_*) = 0$ holds.

So, we cannot have $\lambda_1 > 0$.

If $\lambda_1 = 0$, then, from the first equation of (6.46), we obtain that $\lambda_2 = 1$ and $\lambda_3 = 1$, which implies that $x_{*,1} = x_{*,2} = 0$. We observe that if we set these values, then we satisfy the KKT conditions.

So, the optimal solution is $x_* = (0, 0)$.

2. Next, we consider the case where $c = [-2 \quad -1]^T$. From Figure 6.9, we conclude that only the constraint $f_1(x) \leq 0$ will be active at the optimal point. Therefore, we have
Figure 6.9: Feasible set and isosurfaces of $f_0(x)$, with $c = [-2 - 1]^T$, and $a_1 < a_2$.

$\lambda_2 = \lambda_3 = 0$, and

$$
\begin{bmatrix}
2x_{s,1}\lambda_1 \\
2x_{s,2}\lambda_1
\end{bmatrix} = \begin{bmatrix}
2 \\
1
\end{bmatrix}
$$

$\lambda_i \geq 0$, $i = 1, 2, 3$, \hspace{1cm} (6.48)

$$
\begin{align*}
&f_i(x_s) \leq 0, \hspace{1cm} i = 1, 2, 3, \\
&\lambda_i f_i(x_s) = 0, \hspace{1cm} i = 1, 2, 3.
\end{align*}
$$

From the first equation, we get that $\lambda_1 = \frac{1}{x_{s,1}} = \frac{1}{2x_{s,2}}$. Therefore, $x_{s,1} = 2x_{s,2}$.

Substituting in the relationship $f_1(x_s) = 0$, we get

$$
x_{s,1}^2 + x_{s,2}^2 = 1 \Rightarrow 4x_{s,2}^2 + x_{s,2}^2 = 1 \Rightarrow x_{s,2} = \frac{1}{\sqrt{5}}.
$$

So, we get that $x_{s,1} = \frac{2}{\sqrt{5}}$, $x_{s,1} = \frac{1}{\sqrt{5}}$, $\lambda_1 = \frac{\sqrt{5}}{2}$, $\lambda_2 = \lambda_3 = 0$. \hspace{1cm} \square

## 6.7 Projection onto a convex set

Next, we turn to the extremely important problem of finding the projection onto a convex set. Let $S \subset \mathbb{R}^n$ be a (non-empty) convex set and point $x_0 \in \mathbb{R}^n$. The solution of the problem

$$
\begin{align*}
\text{minimize} \hspace{1cm} & f_0(x) := \|x_0 - x\|^2_2 \\
\text{subject to} \hspace{1cm} & x \in S.
\end{align*}
$$

(6.50)
Figure 6.10: Projection of point $x_0$ onto convex set $S$.

is called the **projection** of the point $x_0$ onto the set $S$ and is denoted by $P_S(x_0)$.

We know that the point $x_*$ minimizes the convex function $f(x)$ over the set $X$ if, and only if,

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

(6.51)

Therefore, the point $P_S(x_0)$ is a solution of the problem (6.50) if, and only if,

$$(x_0 - P_S(x_0))^T(x - P_S(x_0)) \leq 0, \quad \forall x \in S.
$$

(6.52)

Next, we will study an extremely important example, the solution of which is given in closed form.

**Example 6.7.1.** Find the projection of the point $x_0 \in \mathbb{R}^n$ onto the convex set $P := \{x \in \mathbb{R}^n \mid x \geq 0\}$.

This problem is written as follows:

minimize $f_0(x) := \frac{1}{2} \|x_0 - x\|_2^2$

subject to $f_i(x) = -x_i \leq 0, \quad i = 1, \ldots, n.$

(6.53)

The KKT conditions are as follows:

$$
\nabla f_0(x_*) + \sum_{i=1}^n \lambda_i \nabla f_i(x_*) = 0
$$

$$
\lambda_i \geq 0, \quad i = 1, \ldots, n
$$

$$
f_i(x_*) \leq 0, \quad i = 1, \ldots, n
$$

$$
\lambda_i f_i(x_*) = 0, \quad i = 1, \ldots, n.
$$

(6.54)

Defining $\lambda := (\lambda_1, \ldots, \lambda_n)$, we write the first relation of (6.54) as

$$(x_* - x_0) - \lambda = 0
$$

$$
\implies x_{*,i} = x_{0,i} + \lambda_i, \quad i = 1, \ldots, n.
$$

(6.55)
Therefore, for \( i = 1, \ldots, n \), we have
\[
(a) \quad \lambda_i \geq 0, \quad (b) \quad x_{*,i} = x_{0,i} + \lambda_i \geq 0, \quad (c) \quad \lambda_i x_{*,i} = \lambda_i (x_{0,i} + \lambda_i) = 0. \tag{6.56}
\]

We will consider the above relations separately for each \( i \).

1. Let \( x_{0,i} \geq 0 \). Then, due to (a) and (c), we have that \( \lambda_i = 0 \). Therefore, from (b) we get \( x_{*,i} = x_{0,i} \).

2. Let \( x_{0,i} < 0 \). Then, due to (b), we have that \( \lambda_i > 0 \). Therefore, due to (c), we have that \( x_{*,i} = x_{0,i} + \lambda_i = 0 \).

So, \( P(x_0) = (x_0)_+ = \max\{0, x_0\} \), where the operator \( \max\{\cdot, \cdot\} \) is applied elementwise. \( \square \)

**Exercise:** Find the projection of the point \( x_0 \in \mathbb{R}^n \) onto the set
\[
B(0, r) := \{x \in \mathbb{R}^n \mid \|x\|_2 \leq r\}.
\]

**Example 6.7.2.** Let \( 0 \neq a \in \mathbb{R}^n \) and hyperplane \( \mathbb{H} := \{x \in \mathbb{R}^n \mid a^T x = b\} \). Find the distance of a point \( x_0 \in \mathbb{R}^n \) from \( \mathbb{H} \).

We consider the problem
\[
\begin{align*}
\text{minimize} & \quad f_0(x) := \frac{1}{2} \|x_0 - x\|_2^2 \\
\text{subject to} & \quad f_1(x) = a^T x - b = 0. \tag{6.57}
\end{align*}
\]

The KKT conditions are as follows:
\[
\nabla f_0(x_*) + v_1 \nabla f_1(x_*) = 0 \implies x_* - x_0 + v_1 a = 0, \quad a^T x_* - b = 0. \tag{6.58}
\]

Multiplying the first equation from the left by the vector \( a^T \), we get
\[
a^T x_* - a^T x_0 + v_1 \|a\|_2^2 = 0 \implies v_1 = \frac{1}{\|a\|_2^2} (a^T x_0 - b). \tag{6.59}
\]

Therefore, the projection of \( x_0 \) onto \( \mathbb{H} \) is as follows:
\[
x_* = x_0 - \frac{a^T x_0 - b}{\|a\|_2^2} a, \tag{6.60}
\]
and the distance of \( x_0 \) from \( \mathbb{H} \) is equal to
\[
\|x_* - x_0\|_2 = \frac{|a^T x_0 - b|}{\|a\|_2}. \tag{6.61}
\]
Exercise: Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Find the projection of the point $x_0 \in \mathbb{R}^n$ onto the set

$$S := \{x \in \mathbb{R}^n \mid Ax = b\}.$$

Exercise: Let $a, b \in \mathbb{R}^n$ with $a \leq b$. Find the projection of the point $x_0 \in \mathbb{R}^n$ onto the set

$$S := \{x \in \mathbb{R}^n \mid a \leq x \leq b\}.$$
Chapter 7

Duality

For every optimization problem there is another optimization problem which is closely related to the original. The original is called \textbf{primal} problem while the connected one is called \textbf{Lagrange dual} problem.

Under some conditions, usually, convexity and some constraint qualification, the primal and the dual problem have the same optimal cost value and, moreover, it is possible to compute the solution of the primal problem via the solution of the dual.

\textbf{7.1 Primal and Dual Problem}

Let us consider the minimization problem

\[
\begin{align*}
\text{minimize} \quad & f_0(x) \\
\text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \ldots, m, \\
& h_i(x) = 0, \quad i = 1, \ldots, p,
\end{align*}
\]  

(7.1)

with \(x \in \mathbb{R}^n\).

Let \(\mathcal{D} := \cap_{i=0}^m \text{dom } f_i \cap \cap_{i=1}^p \text{dom } h_i\). The feasible set is given by

\[
\mathcal{X} := \{x \in \mathcal{D} \mid f_i(x) \leq 0, \quad i = 1, \ldots, m, \quad h_i(x) = 0, \quad i = 1, \ldots, p\}. 
\]  

(7.2)

At present, we do not assume that the problem (7.1) is a convex optimization problem. We assume that there is a solution to the problem, say, \(x_* \in \mathcal{X}\), with \(f_0(x_*) = p_*\).

We set \(\lambda := [\lambda_1 \cdots \lambda_m]^T\) and \(v := [v_1 \cdots v_p]^T\) and define the Lagrangian function \(L :\)
$D \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, with

$$L(x, \lambda, v) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x).$$  \hspace{1cm} (7.3)

The vectors $\lambda$ and $v$ are called dual variables or Lagrange multiplier vectors.

First, we note that

$$\max_{\lambda \geq 0, v} L(x, \lambda, v) = \begin{cases} f_0(x), & \text{if } x \in X, \\ +\infty, & \text{if } x \notin X. \end{cases}$$  \hspace{1cm} (7.4)

Therefore, the problem (7.1) is equivalent to the problem

$$\min_{x \in D} \max_{\lambda \geq 0, v} L(x, \lambda, v).$$  \hspace{1cm} (7.5)

The dual problem is the following:

$$\max_{\lambda \geq 0, v} \min_{x \in D} L(x, \lambda, v).$$  \hspace{1cm} (7.6)

Next, we will study the dual problem and its relation to the primal problem.

### 7.2 The Lagrangian dual function

We define the Lagrangian dual function, or, simply, dual function, $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, as

$$g(\lambda, v) := \inf_{x \in D} L(x, \lambda, v) = \inf_{x \in D} \left( f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x) \right).$$  \hspace{1cm} (7.7)

When at a point $(\lambda, v)$ the Lagrangian is not bounded below, then we set $g(\lambda, v) = -\infty$.

We define the domain of $g$ as follows:

$$\text{dom} g = \{(\lambda, v) \in \mathbb{R}^m_+ \times \mathbb{R}^p : g(\lambda, v) > -\infty\}.$$  \hspace{1cm} (7.8)

**Theorem 7.2.1.** Let $g$ be the Lagrange dual function of (7.7). Then

1. $\text{dom} g$ is a convex set.

2. The function $g$ is concave.
7.2. THE DUAL FUNCTION

Proof. First, we prove that the domain of \( g \) is a convex set. Let \((\lambda_1, v_1), (\lambda_2, v_2) \in \text{dom} g\). Then

\[
g(\lambda_1, v_1) = \inf_{x \in D} L(x, \lambda_1, v_1) > -\infty,
\]

\[
g(\lambda_2, v_2) = \inf_{x \in D} L(x, \lambda_2, v_2) > -\infty.
\]

(7.9)

Assume that \( \alpha \in [0, 1] \). Then, we get

\[
g(\alpha \lambda_1 + (1 - \alpha) \lambda_2, \alpha v_1 + (1 - \alpha) v_2) = \inf_{x \in D} \alpha L(x, \lambda_1, v_1) + (1 - \alpha) L(x, \lambda_2, v_2)
\]

\[
\geq \inf_{x \in D} \alpha L(x, \lambda_1, v_1) + \inf_{x \in D} (1 - \alpha) L(x, \lambda_2, v_2)
\]

\[
= \alpha \inf_{x \in D} L(x, \lambda_1, v_1) + (1 - \alpha) \inf_{x \in D} L(x, \lambda_2, v_2)
\]

\[
= \alpha g(\lambda_1, v_1) + (1 - \alpha) g(\lambda_2, v_2)
\]

\[
> -\infty.
\]

Thus, \( \text{dom} g \) is a convex set. During the proof, we also proved the concavity of \( g \).\( \square \)

Theorem 7.2.2. Let \( g \) be the Lagrange dual function of (7.7). Then, for each \( \lambda \geq 0 \) and \( v \), we have

\[
g(\lambda, v) \leq p_*.\]

(7.10)

Proof. Let \( \bar{x} \) be a feasible point of the problem (7.1), that is, \( \bar{x} \in X \). Then

\[
\sum_{i=1}^{m} \lambda_i f_i(\bar{x}) + \sum_{i=1}^{p} v_i h_i(\bar{x}) \leq 0,
\]

because every term of the first sum is a non-positive number and every term of the second sum equals zero.

Therefore,

\[
L(\bar{x}, \lambda, v) = f_0(\bar{x}) + \sum_{i=1}^{m} \lambda_i f_i(\bar{x}) + \sum_{i=1}^{p} v_i h_i(\bar{x}) \leq f_0(\bar{x})
\]

(7.11)

and

\[
g(\lambda, v) = \inf_{x \in D} L(x, \lambda, v) \leq L(\bar{x}, \lambda, v) \leq f_0(\bar{x}).
\]

(7.12)
The above relation holds for every $\bar{x} \in X$, so it also holds for the optimal point $x_*$. Therefore
\[ g(\lambda, v) \leq f_0(x_*) = p_* \] (7.13)
proving the theorem.

Through the dual function, we derive a non-trivial lower bound for $p_*$ only when $\lambda \geq 0$ and $(\lambda, v) \in \text{dom } g$, that is, $g(\lambda, v) > -\infty$.

The pairs $(\lambda, v)$ with $\lambda \geq 0$ and $(\lambda, v) \in \text{dom } g$ are called dual feasible points.

**Example 7.2.1.** Consider the optimization problem
\[ \text{minimize } x^T x, \quad \text{subject to } Ax = b. \] (7.14)
with $A \in \mathbb{R}^{p \times n}$. The Lagrange function is the
\[ L(x, v) = x^T x + v^T (Ax - b), \] (7.15)
with $\text{dom } L = \mathbb{R}^n \times \mathbb{R}^p$. The binary function is defined as
\[ g(v) := \inf_{x \in \mathbb{R}^n} L(x, v). \] (7.16)
$L(x, v)$ is a quadratic convex function of $x$. Thus, the point $x$ which minimizes the Lagrange function is given by the solution of the equation
\[ \nabla_x L(x, v) = 0 \Rightarrow 2x + A^T v = 0 \Rightarrow x = -\frac{1}{2} A^T v. \] (7.17)
Therefore,
\[ g(v) = L \left( -\frac{1}{2} A^T v, v \right) = -\frac{1}{4} v^T A A^T v - b^T v, \] (7.18)
which is a quadratic concave function of $v \in \mathbb{R}^p$.

From weak duality, we have that, for every $v \in \mathbb{R}^p$,
\[ -\frac{1}{4} v^T A A^T v - b^T v \leq \inf_{x \in \mathbb{R}^n} \{ x^T x \mid Ax = b \}. \] (7.19)

**Example 7.2.2.** (Linear problem in standard form) Consider the linear programming problem in standard form
\[ \begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax = b, \\
& \quad x \geq 0.
\end{align*} \] (7.20)
The inequality \( x \geq 0 \) is equivalent to the inequality \(-x \leq 0\), that is, \(-x_i \leq 0\), for \( i = 1, \ldots, n \).

The Lagrange function is as follows:

\[
L(x, \lambda, v) = c^T x - \sum_{i=1}^{n} \lambda_i x_i + v^T (Ax - b) = -b^T v + (c + A^T v - \lambda)^T x. \tag{7.21}
\]

The dual function is given by the relation

\[
g(\lambda, v) := \inf_{x \in \mathbb{R}^n} L(x, \lambda, v) = -b^T v + \inf_{x \in \mathbb{R}^n} \{ (c + A^T v - \lambda)^T x \}. \tag{7.22}
\]

If there is no bound on \( x \), then a linear function of \( x \) is bounded below only if it is identically zero.

Therefore,

\[
g(\lambda, v) = \begin{cases} 
-b^T v, & c + A^T v - \lambda = 0, \\
-\infty, & \text{otherwise.} 
\end{cases} \tag{7.23}
\]

The inequality (7.13) is non-trivial only for pairs \((\lambda, v) \in \mathbb{R}^m \times \mathbb{R}^p\) such that \(c + A^T v - \lambda = 0\).

These pairs are the solution of the system of linear equations

\[
\begin{bmatrix} -I & A^T \end{bmatrix} \begin{bmatrix} \lambda \\ v \end{bmatrix} = -c. \tag{7.24}
\]

Therefore, they define an affine set.

For these pairs of points, the function \(-b^T v\) provides a non-trivial lower bound for the optimal value of the problem (7.20).

\[\square\]

### 7.3 The Lagrangian dual problem

We have seen that, for each pair \((\lambda, v)\), with \( \lambda \geq 0 \), the function \( g(\lambda, v) \) is a lower bound for the optimal value of the cost function \( p_* \).

An extremely important question is this: \textit{What is the largest lower bound that we can extract from the dual function \( g(\lambda, v) \)?}
To answer this question, we define the problem

\[
\begin{align*}
\text{maximize} & \quad g(\lambda, v) \\
\text{subject to} & \quad \lambda \geq 0.
\end{align*}
\] (7.25)

The problem (7.1) is called **primal** problem, while (7.25) is called **Lagrange dual** problem, problem.

If the problem (7.25) has a solution, then the optimal point, \((\lambda_*, v_*)\), is called **dual optimal** or **optimal Lagrange multipliers**.

The problem (7.25) is **always** a convex optimization problem because the constraint \(\lambda \geq 0\) is convex and the function \(g(\lambda, v)\) is concave.

**Example 7.3.1.** *(Linear problem in standard form)* Consider the linear programming problem in standard form

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax = b, \\
& \quad x \geq 0.
\end{align*}
\] (7.26)

As we have seen, the dual function is given by the relation

\[
g(\lambda, v) = \begin{cases} 
-b^T v, & c + A^T v - \lambda = 0, \\
-\infty, & \text{otherwise.}
\end{cases}
\] (7.27)

The dual problem is as follows

\[
\begin{align*}
\text{maximize} & \quad g(\lambda, v) \\
\text{subject to} & \quad \lambda \geq 0.
\end{align*}
\] (7.28)

The function \(g\) takes values other than \(-\infty\) only if \(c + A^T v - \lambda = 0\).

Therefore, an equivalent expression for the dual problem is as follows

\[
\begin{align*}
\text{maximize} & \quad -b^T v \\
\text{subject to} & \quad c + A^T v - \lambda = 0, \\
& \quad \lambda \geq 0.
\end{align*}
\] (7.29)

or, equivalently,

\[
\begin{align*}
\text{maximize} & \quad -b^T v \\
\text{subject to} & \quad c + A^T v \geq 0,
\end{align*}
\] (7.30)

which is a linear programming problem in inequality form. □

More examples in the book by Boyd and Vandenberghe.
7.4 Weak duality

We know that \( g(\lambda, v) \leq p^* \), for every \( \lambda \geq 0 \) and \( v \). If we define

\[
d^* := \max_{\lambda \geq 0, v} g(\lambda, v),
\]

then we get that

\[
d^* \leq p^*.
\]

The inequality (7.32) is called weak duality. The relation (7.32) holds even when the values of \( d^* \) and \( p^* \) are \( \pm \infty \). For example, if \( p^* = -\infty \), that is, if the primal problem is not bounded from below, then \( d^* = -\infty \), that is, the binary problem is not feasible, and vice versa.

The quantity \( p^* - d^* \) is called optimal duality gap and is always non-negative.

7.5 Strong Duality

If \( d^* = p^* \), then we say we have strong duality. Strong duality does not always hold. But, if the original problem is convex, then it very often does. For example, if the primal problem is convex and there exists \( x \in X \) which satisfies the Slater condition, then strong duality can be shown to hold (see section 5.3.2 of B&V’s book).

7.6 Optimality Conditions

Next, we assume that strong duality holds and derive optimality conditions for primary problem.\(^1\)

\(^1\)This approach is complementary to that of the previous chapter.
CHAPTER 7. DUALITY

7.6.1 Complementary slackness

Let \( x^* \) be the primal optimal point and \((\lambda^*, v^*)\), with \( \lambda^* \geq 0 \), the dual optimal point. Then

\[
\begin{align*}
  f_0(x^*) = g(\lambda^*, v^*) &= \inf_{x \in \mathbb{D}} \left( f_0(x) + \sum_{i=1}^{m} \lambda^*_{i} f_i(x) + \sum_{i=1}^{p} v^*_{i} h_i(x) \right) \\
  &\leq f_0(x^*) + \sum_{i=1}^{m} \lambda^*_{i} f_i(x^*) + \sum_{i=1}^{p} v^*_{i} h_i(x^*) \\
  &\leq f_0(x^*),
\end{align*}
\]

where

1. the inequality \((a)\) holds because the infimum of \( L(x, \lambda^*, v^*) \), over all \( x \in \mathbb{D} \), is less than or equal to \( L(x^*, \lambda^*, v^*) \),

2. the inequality \((b)\) holds because \( \lambda^*_{i} f_i(x^*) \leq 0 \), for \( i = 1, \ldots, m \), and \( v^*_{i} h_i(x^*) = 0 \), for \( i = 1, \ldots, p \).

With some thought, we can find out that the last two inequalities of (7.33) hold as equalities. Therefore, it should be true that

\[
\sum_{i=1}^{m} \lambda^*_{i} f_i(x^*) = 0, \quad (7.34)
\]

from which it follows that

\[
\lambda^*_{i} f_i(x^*) = 0, \quad \text{for } i = 1, \ldots, m. \quad (7.35)
\]

The relation (7.35) is called complementary slackness.

7.6.2 Conditions KKT

Since the inequality \((a)\) of (7.33) holds as equality, we conclude that \( x^* \) minimizes the function \( L(x, \lambda^*, v^*) \).

Assuming that the functions \( f_i \) for \( i = 0, \ldots, m \), and \( h_i \) for \( i = 1, \ldots, p \), are differentiable, then their domains \( \text{dom} f_i \), for \( i = 0, \ldots, m \) and \( \text{dom} h_i \), for \( i = 1, \ldots, p \) are open sets, therefore, \( \mathbb{D} \) is also an open set.
Notice that the function \( L(x, \lambda_*, v_*) \) is a differentiable function of \( x \). Therefore, at the point \( x_* \) which minimizes it, the gradient of \( L(x, \lambda_*, v_*) \) should be equal to the zero vector, that is,

\[
\nabla f_0(x_*) + \sum_{i=1}^{m} \lambda_{*,i} \nabla f_i(x_*) + \sum_{i=1}^{p} v_{*,i} \nabla h_i(x_*) = 0.
\]

(7.36)

If we collect all the relations which must be satisfied by \( x_*, \lambda_*, \) and \( v_* \), we get

\[
\nabla f_0(x_*) + \sum_{i=1}^{m} \lambda_{*,i} \nabla f_i(x_*) + \sum_{i=1}^{p} v_{*,i} \nabla h_i(x_*) = 0,
\]

\[
f_i(x_*) \leq 0, \quad i = 1, \ldots, m,
\]

\[
h_i(x_*) = 0, \quad i = 1, \ldots, p,
\]

\[
\lambda_{*,i} \geq 0, \quad i = 1, \ldots, m,
\]

\[
\lambda_{*,i} f_i(x_*) = 0, \quad i = 1, \ldots, m.
\]

(7.37)

That is, we get the KKT conditions.

Therefore, for any optimization problem with differentiable functions for which strong duality holds, every pair of primal and dual optimal points must satisfy the KKT conditions.

We will not expand further because we have already seen in the previous chapter when the KKT conditions are sufficient and necessary optimality conditions of convex optimization problems (see the section 5.3.3 of the book B&V).

### 7.7 Usefulness of duality

The dual problem is very useful for many reasons. Below, we briefly present two of them.

#### 7.7.1 Suboptimality Guarantee and Algorithm Termination Criteria

If \((\bar{\lambda}, \bar{v})\) is a feasible point for the dual problem and \(\bar{x}\) is a feasible point for the primal problem, then

\[
g(\bar{\lambda}, \bar{v}) \leq p_* \leq f(\bar{x}).
\]

(7.38)

Assuming that

\[
f(\bar{x}) - g(\bar{\lambda}, \bar{v}) = \bar{\epsilon},
\]

(7.39)
then
\[ f(\bar{x}) - p_* \leq f(\bar{x}) - g(\bar{\lambda}, \bar{v}) = \bar{\epsilon}. \] (7.40)

That is, if we know a feasible point \( \bar{x} \) and feasible dual point \((\bar{\lambda}, \bar{v})\) such that the relation (7.39) holds, then we can conclude that \( \bar{x} \) is (at most) \( \bar{\epsilon} \)-suboptimal. This conclusion can be used as a termination criterion for some optimization algorithms.

Others cases for which this conclusion is extremely useful are cases where the primary problem is extremely difficult to solve, therefore, we can be satisfied with some "good" suboptimal solutions.

### 7.7.2 Solving a primal problem through the dual

In some cases (1) solving the dual problem is easier than solving the primal one and (2) we can solve the primal problem using the dual solution.

**Example 7.7.1.** We consider the problem
\[
\begin{align*}
\text{minimize} & \quad f(x) = x^T x \\
\text{subject to} & \quad Ax = b,
\end{align*}
\] (7.41)

with \( A \in \mathbb{R}^{p \times n} \) and \( \text{rank}(A) = p \).

As we have seen, the dual function is given by the relation
\[
g(v) = -\frac{1}{4}v^T AA^T v - b^Tv. \] (7.42)

The dual problem is the quadratic problem, without constraints,
\[
\begin{align*}
\text{minimize} & \quad -g(v) = \frac{1}{4}v^T AA^T v + b^Tv.
\end{align*}
\] (7.43)

Since we have assumed that the rows of \( A \) are linearly independent, the \((p \times p)\) matrix \( AA^T \) is invertible. The solution of the dual problem is equal to
\[
v_* = -2(AA^T)^{-1}b. \] (7.44)

The KKT conditions are necessary and sufficient for this problem (why?) and are given by the relation
\[
\nabla f(x_*) + A^Tv_* = 0 \\
\implies 2x_* - 2A^T(AA^T)^{-1}b = 0. \] (7.45)
Therefore, the optimal solution is given by the relation

\[ x_\ast = A^T (AA^T)^{-1} b. \]  

That is, we solved the primal problem, with affine constraints, by using the solution of the unconstrained dual problem, which is the point of the affine set \( \{ x \in \mathbb{R}^n \mid Ax = b \} \) with minimum Euclidean norm.

In Figure 7.1, we draw

1. the affine set \( a^T x = b \) in \( \mathbb{R}^2 \), for \( a^T = [1, 1] \) and \( b = 1 \),
2. the optimal point \( x_\ast = (0.5, 0.5) \), with \( f(x_\ast) = x_\ast^T x_\ast = 0.5 \),
3. the binary function \( g(v) \), with maximum value \( g(v_\ast) = g(-1) = 0.5 \).

Figure 7.1: Affine set, binary function and affine set point of minimum Euclidean measure.
Chapter 8

Convex optimization with affine equality constraints

In this chapter, we will study problems of the form

\[
\begin{align*}
& \text{minimize} & & f(x) \\
& \text{subject to} & & Ax = b,
\end{align*}
\]

(8.1)

with \( f : \text{dom} f \subseteq \mathbb{R}^n \to \mathbb{R} \) being a convex doubly differentiable function, \( A \in \mathbb{R}^{p \times n} \), with \( \text{rank}(A) = p \), and \( b \in \mathbb{R}^p \).

We assume that the problem has a solution, say, \( x_* \), and define \( p_* := f(x_*) = \inf \{ f(x) | Ax = b \} \). We know that the point \( x_* \) is optimal if, and only if, the KKT conditions are satisfied, which, in this case, are as follows:

\[
\nabla f(x_*) + A^T v_* = 0, \quad Ax_* = b,
\]

(8.2)

with \( v_* \in \mathbb{R}^p \). The relations (8.2) constitute a system of \( (n + p) \) (usually) nonlinear equations.

The linear system of equations \( Ax_* = b \) is called the system of \textbf{primal feasibility}. The system of equations \( \nabla f(x_*) + A^T v_* = 0 \) is called \textbf{dual feasibility} equations, and is, in general, a system of nonlinear equations.

There are several approaches to solving the problem (8.2). We mention the following.

1. We can remove the equality constraints and then work on problems without constraints, as in the previous chapter.
2. If the dual function is differentiable, then we can solve the dual problem, which is an unconstrained problem, and then, somehow, calculate the primal optimal solution.

3. We can extend the Newton method from unconstrained problems to problems with equality constraints. This will be the basic approach we will follow.

### 8.1 Solving a primal problem via the dual

We assume that strong duality holds. The dual function for problem (8.1) is defined as:

$$g(v) = \inf_{x \in \text{dom } f} \left\{ f(x) + v^T (Ax - b) \right\}$$

$$= -b^T v + \inf_{x \in \text{dom } f} \left\{ f(x) + bv^T Ax \right\}.$$  \hspace{1cm} (8.3)

If the function $g$ is doubly differentiable, then we can use any algorithm for the solution of unconstrained optimization problems and to solve the dual problem

$$\max_v g(v).$$  \hspace{1cm} (8.4)

Then, the optimal dual variable, $v^*$, can be used to calculate the optimal primal variable, $x^*$.

**Example 8.1.1. (Analytical center with equality constraints)** Consider the problem

$$\min_x f(x) = -\sum_{i=1}^{n} \log x_i$$

subject to $Ax = b$,  \hspace{1cm} (8.5)

with $\text{dom } f = \mathbb{R}^n_{++}$ and $A \in \mathbb{R}^{p \times n}$ with $\text{rank}(A) = p$.

The Lagrange function is as follows

$$L(x, v) = f(x) + v^T (Ax - b) = -\sum_{i=1}^{n} \log x_i + v^T (Ax - b).$$  \hspace{1cm} (8.6)

For a given $v$, $L(x, v)$ is a convex function of $x$ (prove it).

The derivative of $L$, with respect to $x$, is as follows

$$\nabla_x L(x, v) = -\left[ x_1^{-1} \ldots x_n^{-1} \right]^T + A^T v.$$  \hspace{1cm} (8.7)

Solving the equation $\nabla_x L(x, v) = 0$, we obtain

$$x_i = \frac{1}{(A^T v)_i}, \text{ for } i = 1, \ldots, n.$$  \hspace{1cm} (8.8)
8.2. CONVEX QUADRATIC PROBLEMS

The dual function is as follows
\[ g(v) = \inf_x L(x, v) = L(1./(A^T v), v) = \sum_{i=1}^{n} \log (A^T v)_i + n - b^T v. \] (8.9)

The dual problem is
\[ \max_v g(v) = -b^T v + n + \sum_{i=1}^{n} \log (A^T v)_i, \] (8.10)

which is a problem with implicit constraints \( A^T v > 0 \). Suppose that the dual problem has a solution \( v_* \). Then, from (8.8), we get that
\[ x_{*,i} = \frac{1}{(A^T v)_i}, \quad \text{for } i = 1, \ldots, n. \] (8.11)

Therefore, if we solve the dual problem, then we can very easily solve the primal problem.

\[ \square \]

8.2 Convex quadratic problems with affine equality constraints

Before proceeding to the extension of the Newton method for the solution of convex optimization problems with affine equality constraints, we will study a simple but extremely important problem, which has a closed form solution.

The solution of this problem will be used as a building block for the development of algorithms for convex optimization problems, with affine equality constraints.

Consider the quadratic problem with affine equality constraints
\[ \min_x f(x) = \frac{1}{2} x^T P x + q^T x + r \]
subject to \( A x = b \),

with \( P = P^T \in \mathbb{R}^{n \times n}, \ P \succeq 0, \ q \in \mathbb{R}^n, \) and \( r \in \mathbb{R} \).

The KKT conditions for this problem are
\[ Px_* + q + A^T v_* = 0, \quad A x_* = b, \] (8.13)

or, equivalently,
\[ \begin{bmatrix} P & A^T \\ A & O \end{bmatrix} \begin{bmatrix} x_* \\ v_* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}, \] (8.14)
that is, they are a system of \((n + p)\) linear equations.

If the coefficient matrix is invertible, then there is a \textit{unique} solution for the (primal–dual) optimal pair \((\mathbf{x}_*, \mathbf{v}_*)\). It can be shown that, if \(\mathbf{P} \succ \mathbf{O}\), then the coefficient matrix is invertible. (For more details, see B&V, pages 522, 523).

### 8.3 Newton’s method starting from a feasible point

In this subsection, we extend the Newton method from unconstrained problems to problems with affine equality constraints.

The approaches are very similar. However, there are the following differences:

1. We must start from a feasible point, say \(\mathbf{x}_0\), that is, we must have \(\mathbf{x}_0 \in \text{dom} f\) and \(A\mathbf{x}_0 = \mathbf{b}\). Often, finding an initial feasible point is quite difficult and, perhaps, requires the solution of an optimization problem, in the form of \textit{feasibility check}.

2. The definition of the Newton step should be done so that the equality constraints are taken into account, that is, we should have that \(A\mathbf{\Delta x}_N = 0\). As we shall see below, this is relatively simple to attain.

### 8.4 Step Newton

As in the case of unconstrained optimization problems, the Newton step can be derived through various approaches.

#### 8.4.1 Definition via second-order approximation minimization

The first approach for the computation of the Newton step for the problem

\[
\begin{align*}
\text{minimize} & \quad f(\mathbf{x}) \\
\text{subject to} & \quad A\mathbf{x} = \mathbf{b},
\end{align*}
\]

at a feasible point \(\mathbf{x}\), that is, \(\mathbf{x} \in \text{dom} f\) and \(A\mathbf{x} = \mathbf{b}\), is as follows.

We replace the cost function \(f\) by its second-order approximation, at the point \(\mathbf{x}\), constructing the quadratic problem, with variable \(\mathbf{z}\),

\[
\begin{align*}
\min_{\mathbf{z}} & \quad \hat{f}(\mathbf{x} + \mathbf{z}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^T \mathbf{z} + \frac{1}{2} \mathbf{z}^T \nabla^2 f(\mathbf{x}) \mathbf{z} \\
\text{subject to} & \quad A(\mathbf{x} + \mathbf{z}) = \mathbf{b}.
\end{align*}
\]
We define the Newton step for the problem (8.15), at the feasible point \( x \), the solution of the problem (8.16), which we call \( \Delta x_{Nt} \). Essentially, the Newton step, \( \Delta x_{Nt} \), is the vector that must be added to \( x \) so that the second-order approximation of \( f \), at \( x \), be minimized at the point \( x + \Delta x_{Nt} \).

The KKT conditions for the quadratic problem (8.16) are as follows
\[
\begin{bmatrix}
\nabla^2 f(x) & A^T \\
A & O
\end{bmatrix}
\begin{bmatrix}
\Delta x_{Nt} \\
w
\end{bmatrix}
= 
\begin{bmatrix}
-\nabla f(x) \\
0
\end{bmatrix},
\]
(8.17)
where \( w \) is the optimal dual variable for the given quadratic problem. If \( \nabla^2 f(x) \succ O \), then the system (8.17) has a unique solution.

As in unconstrained optimization problems, we observe that, if the function \( f \) is quadratic, then the Newton algorithm solves the optimization problem in one step. Therefore, we expect that the Newton algorithm will behave (very) well when \( f \) is (very) close to a quadratic function.

### 8.4.2 Definition by solution of linearized optimality conditions

Another interpretation of the Newton step, at the point \( x \), is as follows. We have seen that the optimality conditions for the problem (8.1) are given by the relations
\[
\nabla f(x^*) + A^T v^* = 0, \quad Ax^* = b.
\]
(8.18)

Instead of the system (8.18), which, usually, is nonlinear, we solve its following approximation.

We replace \( x^* \) with \( x + \Delta x_{Nt} \) and \( v^* \) with \( w \). If, moreover, we replace the gradient \( \nabla f(x + \Delta x_{Nt}) \) by its first-order approximation, at the point \( x \), then we get
\[
\nabla f(x) = \nabla f(x + \Delta x_{Nt}) \approx \nabla f(x) + \nabla^2 f(x) \Delta x_{Nt},
\]
(8.19)
and an approximation of (8.18) is as follows:
\[
\nabla f(x) + \nabla^2 f(x) \Delta x_{Nt} + A^T w = 0
\]
\[
A(x + \Delta x_{Nt}) = b.
\]
(8.20)

Considering that \( Ax = b \), the above relations are expressed as
\[
\nabla^2 f(x) \Delta x_{Nt} + A^T w = -\nabla f(x), \quad A \Delta x_{Nt} = 0,
\]
(8.21)
which are the same expressions as those in (8.17).

\(^1\)We assume there is a solution.
8.4.3 Analytic expression for the Newton step

As we prove in the Appendix, the analytic expressions for $\Delta x_{Nt}$ and $w$ are as follows:

$$
\Delta x_{Nt} = - (\nabla^2 f(x))^{-1} \nabla f(x) + (\nabla^2 f(x))^{-1} A (\nabla^2 f(x))^{-1} \nabla f(x),
$$

(8.22)

$$
w = - \left( A (\nabla^2 f(x))^{-1} A^T \right)^{-1} A (\nabla^2 f(x))^{-1} \nabla f(x).
$$

(8.23)

Notice that these two quantities satisfy (8.21) and, in particular,

$$
\Delta x_{Nt} = - (\nabla^2 f(x))^{-1} (\nabla f(x) + A^T w).
$$

(8.24)

8.4.4 Newton decrement

As in unconstrained optimization problems, we define the Newton decrement, at the feasible point $x$, as follows:

$$
\lambda(x) := \frac{(\Delta x_{Nt}^T \nabla^2 f(x) \Delta x_{Nt})^{\frac{1}{2}}}{2}.
$$

(8.25)

The quantity $\lambda(x)$ provides an approximation of the difference $f(x) - p^*$, based on the second-order approximation of $f$, at the point $x$. More specifically,

$$
f(x) - \inf \left\{ \hat{f}(x + z) \mid Az = 0 \right\} = \frac{\lambda^2(x)}{2}.
$$

(8.26)

The proof is as follows:

$$
\inf \left\{ \hat{f}(x + z) \mid Az = 0 \right\} = \hat{f}(x + \Delta x_{Nt})
\quad = f(x) + \nabla f(x)^T \Delta x_{Nt} + \frac{1}{2} \Delta x_{Nt} \nabla^2 f(x) \Delta x_{Nt}.
$$

(8.27)

From (8.21), we get that

$$
\nabla f(x)^T \Delta x_{Nt} = - (\Delta x_{Nt}^T \nabla^2 f(x) + w^T A) \Delta x_{Nt}
\quad = -\Delta x_{Nt}^T \nabla^2 f(x) \Delta x_{Nt}
\quad = -\lambda^2(x).
$$

(8.28)

Combining (8.27), (8.28), and (8.25), we get (8.26).

Finally, the Newton decrement appears in the line search because

$$
\left. \frac{d}{dt} f(x + t \Delta x_{Nt}) \right|_{t=0} = \nabla f(x)^T \Delta x_{Nt} = -\lambda^2(x).
$$

(8.29)
8.4.5 Feasible descent direction

Suppose we are at the feasible point $x$. Then, $z \in \mathbb{R}^n$ is a feasible direction of movement if $Az = 0$, and descent direction if, for sufficiently small $t$, $f(x + tz) < f(x)$.

The Newton step $\Delta x_{Nt}$ is always a feasible descent step, because

1. from the second system of equations of (8.17), we have that $A \Delta x_{Nt} = 0$,
2. due to (8.29), we have that, for sufficiently small $t$, $f(x + t \Delta x_{Nt}) \leq f(x)$.

8.4.6 Newton’s algorithm starting from a feasible point

The Newton algorithm starting from a feasible point is presented in Table 8.1.

As we see in Fig. 8.1, if we start from a feasible starting point $x_0$, then, for $k \geq 0$, the points $x_k$ will be feasible and the inequality $f(x_k) < f(x_{k-1})$ will hold true, for every $k$, unless $x_{k-1} = x^*_*$. 

8.5 Convergence Analysis

The Newton method for problems with linear equality constraints, starting from a feasible point, behaves like the Newton method for unconstrained optimization problems.

For details of the convergence analysis of the method see Section 10.2.4 of the book of Boyd and Vandenberghe.
CHAPTER 8. AFFINE EQUALITY CONSTRAINTS

\[ x \in \text{dom } f, \ Ax = b, \text{tolerance } \epsilon. \]

while (TRUE)
1. compute Newton step and decrement : \( \Delta x_{Nt}, \lambda(x) \).
2. quit if \( \lambda^2(x) \leq \epsilon \).
3. Perform backtracking line search and choose \( t \).
4. \( x := x + t \Delta x_{Nt} \).

Table 8.1: The Newton algorithm for problems with equality constraints starting from a feasible point.

8.6 Newton algorithm starting from an infeasible point

Next, we will study the important case where we start the Newton algorithm from an infeasible \( x \in \text{dom } f \), that is, \( Ax \neq b \).

This may happen either because the computation of a feasible point is difficult, for example, when \( \text{dom } f \subset \mathbb{R}^n \), or for any other reason.

8.7 Newton step at infeasible points

We repeat that the optimality conditions for the problem (8.1) are as follows:

\[ \nabla f(x^*) + A^T v^* = 0, \quad Ax^* = b. \]

Let \( x \in \text{dom } f \) be a point which not is feasible, that is,

\[ Ax \neq b. \]

Our goal is to find a step \( \Delta x \) such that the point \( x + \Delta x \) is feasible and approximately optimal, that is,

\[ A(x + \Delta x) = b, \quad x + \Delta x \approx x^*. \]

Replacing \( x^* \) in the optimality conditions by \( x + \Delta x \) and \( v^* \) by \( w \), and using the first-order approximation

\[ \nabla f(x + \Delta x) \approx \nabla f(x) + \nabla^2 f(x) \Delta x, \]

we obtain

\[ \nabla f(x) + \nabla^2 f(x) \Delta x + A^T w = 0, \quad A(x + \Delta x) = b, \]

(8.30)
8.7. NEWTON STEP AT INFEASIBLE POINTS

or, equivalently,

\[
\begin{bmatrix}
\nabla^2 f(x) & A^T \\
A & O
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta w
\end{bmatrix} = -\begin{bmatrix}
\nabla f(x) \\
Ax - b
\end{bmatrix}.
\]  
(8.32)

The only difference of the relation (8.32), which defines the Newton step at an infeasible point, from the relation (8.17), which defines the Newton step at a feasible point, is in the second term of the right-hand side, which is \(Ax - b\) in (8.32), while it is zero in (8.17).

If \(x\) is feasible, then these two terms are identical. If, however, \(x\) is not feasible, then the two terms are different.

We note that if \(x\) is not feasible and we move along \(\Delta x\), then the new point \(x + \Delta x\) will be feasible.

An extremely interesting case is the one in which \(x\) is not a feasible point but it is not possible to move by \(\Delta x\) because \(\text{dom } f \subset \mathbb{R}^n\) and \(x + \Delta x \notin \text{dom } f\). We will consider this case later.

8.7.1 Interpretation as primal–dual Newton step

Next, we give another interpretation of the relation (8.32) in the context of the algorithm primal–dual, which updates both the primal variable \(x\) and the dual variable \(v\) with the goal of solving the KKT conditions.

Let the function \(r : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n \times \mathbb{R}^p\) be defined as

\[
r(x, v) := (r_{\text{dual}}(x, v), r_{\text{primal}}(x, v)),
\]  
(8.33)

with

\[
r_{\text{dual}}(x, v) := \nabla f(x) + A^T v, \quad r_{\text{primal}}(x, v) := Ax - b.
\]  
(8.34)

The optimality conditions (8.2) of the problem (8.1) can be expressed as

\[
r(x_*, v_*) = 0.
\]  
(8.35)

Therefore, the quantities \(r_{\text{dual}}(x, v)\) and \(r_{\text{primal}}(x, v)\), which indicate “how far the point \((x, v)\) is from the optimal point,” can be characterized, respectively, as dual residual and primal residual, at the point \((x, v)\).

The first-order approximation of the function \(r\), at the point \(y = (x, v)\), is given by the relation

\[
r(y + z) \approx r_y(z) = r(y) + Dr(y)z,
\]  
(8.36)
where \( Dr(y) \) is the derivative of \( r \) at \( y \). Function \( r_y(z) \) is an affine function of \( z \).

We define as **primal-dual Newton step**

\[
\Delta y_{pd} := (\Delta x_{pd}, \Delta v_{pd}),
\]

at the point \( y \), the \( z \) for which the first-order approximation \( r_y(z) \) becomes zero, i.e.,

\[
Dr(y)\Delta y_{pd} = -r(y).
\]  
(8.37)

That is, the primal–dual Newton step is defined in terms of \( x \) and \( v \).

The derivative \( Dr(y) = Dr(x, v) \) is as follows

\[
Dr(x, y) = \begin{bmatrix}
Dr_{\text{dual}}(x, v) \\
Dr_{\text{primal}}(x, v)
\end{bmatrix} = \begin{bmatrix}
\nabla^2 f(x) & A^T \\
A & O
\end{bmatrix},
\]  
(8.38)

therefore, (8.37) can be expressed as

\[
\begin{bmatrix}
\nabla^2 f(x) & A^T \\
A & O
\end{bmatrix} \begin{bmatrix}
\Delta x_{pd} \\
\Delta v_{pd}
\end{bmatrix} = - \begin{bmatrix}
r_{\text{dual}}(x, v) \\
r_{\text{primal}}(x, v)
\end{bmatrix} = - \begin{bmatrix}
\nabla f(x) + A^T v \\
Ax - b
\end{bmatrix}.
\]  
(8.39)

Setting \( v_+ = v + \Delta v_{pd} \), we obtain the relation

\[
\begin{bmatrix}
\nabla^2 f(x) & A^T \\
A & O
\end{bmatrix} \begin{bmatrix}
\Delta x_{pd} \\
v_+
\end{bmatrix} = - \begin{bmatrix}
\nabla f(x) \\
Ax - b
\end{bmatrix},
\]  
(8.40)

which is identical to (8.32) for

\[
\Delta x = \Delta x_{pd}, \quad w = v_+ = v + \Delta v_{pd}.
\]  
(8.41)

### 8.7.2 Merit function

We recall that the Newton direction, \( \Delta x \), at a **infeasible point** \( x \) is given by the relation (8.31). The direction \( \Delta x \) is not necessarily a descent direction for \( f \), at the point \( x \), because quantity

\[
\frac{d}{dt} f(x + t\Delta x) \bigg|_{t=0} = \nabla f(x)^T \Delta x
\]

(8.31) \( = -\Delta x^T (\nabla^2 f(x) \Delta x + bfA^T w) \) \( = -\Delta x^T \nabla^2 f(x) \Delta x + (Ax - b)^T w \)  
(8.42)
is not necessarily negative (unless $Ax = b$). This fact has several consequences. For example, we cannot speak of “descent Newton step” with all this implies.

However, the interpretation of the quantity $\Delta y_{pd} = (\Delta x_{pd}, \Delta v_{pd})$ as primal–dual Newton step allows the step $\Delta y_{pd}$ to be related to another function.

The directional derivative of function $\|r\|_2^2$, at the point $y$ along the direction $\Delta y_{pd}$, is\footnote{From the chain rule, we have}
\begin{align*}
\frac{d}{dt} \|r(y + t\Delta y_{pd})\|_2^2 \bigg|_{t=0} &= 2 r(y)^T Dr(y) \Delta y_{pd} \\
&= -2 r(y)^T r(y) \\
&= -2 \|r(y)\|_2^2 \\
&\leq 0.
\end{align*}
(8.43)

So, the direction $\Delta y_{pd}$ is a descent direction for the merit function $\|r\|_2^2$.

This fact allows the use of the quantity $\|r\|_2$ as a measure of progress of the Newton method, starting from a non-feasible point, and can be used, for example, in a line search and/or in algorithm termination criteria.

### 8.7.3 Full-Step Feasibility Property

As we have seen, if we are at an infeasible point and move by $\Delta x_{pd}$, then we will reach a feasible point. Then, we will remain at feasible points whether we take a full step (i.e., $t = 1$) or not.

If, however, we move by $t \Delta x_{pd}$, with $t \in (0, 1)$, then the new point will not be feasible. If we define $x_+ = x + t \Delta x_{pd}$, and $v_+ = v + t \Delta v_{pd}$, then
\begin{align*}
r_{primal}(x_+, v_+) &= A(x + t \Delta x_{pd}) - b \\
&\overset{(a)}{=} (1 - t)(Ax - b) \\
&\overset{}{=} (1 - t)r_{primal}(x, v),
\end{align*}
where at point $(a)$ we used the fact that, because of the relation (8.39), we have
\begin{equation*}
A \Delta x_{pd} = -(Ax - b).
\end{equation*}
8.7.4 Newton algorithm starting from an infeasible point

The Newton algorithm starting from an infeasible point is listed in Table 8.2 (see also Fig 8.2).

An obvious variation is as follows. If, at some step, we choose $t = 1$, then the next point will be feasible and we can continue with the Newton algorithm starting from a feasible point.

The main difference between these two approaches lies in the line search. However, this difference does not seem to significantly differentiate the behavior of the methods.

8.7.5 Simplify initialization via Newton by starting from a non-feasible point

As we have seen, if $\text{dom } f \subset \mathbb{R}^n$, then finding a feasible point may be difficult. In this case, one approach is to solve the problem in two phases.
\( x \in \text{dom} \, f, \, v \in \mathbb{R}^p, \text{tolerance } \epsilon. \)

Repeat

1. compute primal and dual Newton steps : \( \Delta x_{pd}, \Delta v_{pd}. \)

2. Backtracking line search

\[
\begin{align*}
& t := 1 \\
& \text{while } \| r(x + t\Delta x_{pd}, v + t\Delta v_{pd}) \|_2 > (1 - \alpha t) \| r(x, v) \|_2 \\
& \quad t := \beta t.
\end{align*}
\]

3. \( x := x + t\Delta x_{pd}, \, v := v + t\Delta v_{pd}. \)

until \( \| r(x, v) \|_2 \leq \epsilon. \)

Table 8.2: The Newton algorithm for problems with equality constraints starting from an infeasible point (primal–dual).

In Phase I, we solve a **feasibility problem**. If this problem has a solution, then we can proceed to Phase II, using the Newton algorithm starting from a feasible point.

If \( \text{dom} \, f \) is relatively simple and we know that the problem is feasible, then it may be preferable to use the Newton algorithm starting from an infeasible point.

**Example 8.7.1.** Consider the problem

\[
\begin{align*}
\text{minimize} \quad & f(x) = -\sum_{i=1}^n \log x_i \\
\text{subject to} \quad & Ax = b, \tag{8.46}
\end{align*}
\]

with \( \text{dom} \, f = \mathbb{R}^n_{++}. \) We can follow the following approaches.

1. Find a point \( x_0 \), if it exists, with \( x_0 \in \text{dom} \, f \) and \( Ax_0 = b \), and then use the Newton algorithm starting from a feasible point.

Finding a feasible starting point requires the solution of the **feasibility problem**

\[
\begin{align*}
\text{minimize} \quad & f(x) = 0 \\
\text{subject to} \quad & Ax = b \tag{8.47} \\
& x > 0.
\end{align*}
\]

We note that, so far, we do not know how to solve the problem (8.47), because it contains inequalities (we will find out in the next chapter).

2. Start from an infeasible point \( x \in \text{dom} \, f \), for example, \( x = 1 \), and use the primal–dual algorithm.
3. Solve the dual problem (if it has a solution)

\[
\text{maximize } g(v) = -b^T v + \sum_{i=1}^{n} \log (A^T v)_i + n, \tag{8.48}
\]

and compute the solution of the primal problem via the solution of the dual. More specifically, if \( v_* \) is the solution of the dual problem, then

\[
x_{*,i} = 1 / (A^T v_*)_i, \quad \text{for } i = 1, \ldots, n. \tag{8.49}
\]

We understand that, in general, there are many approaches for the solution of an optimization problem. It is extremely important to know the pros and cons of each of them so that we be able to choose the most suitable.
Chapter 9
Convex Optimization Problems

In this chapter, we study convex optimization problems in their general form, that is, convex minimization problems with convex inequality constraints and affine equality constraints.

There are several approaches for the solution of these important problems. We will focus on the interior point method, stating that it is one of the most widespread methods for solving convex optimization problems.

9.1 Convex optimization problems

Consider the convex optimization problem

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0, \ i = 1, \ldots, m, \\
& \quad Ax = b,
\end{align*}
\]

where \( f_i : \text{dom} f_i \subseteq \mathbb{R}^n \to \mathbb{R}, \) for \( i = 0, \ldots, m, \) are convex, doubly differentiable functions (thus, \( \text{dom} f_i, \) for \( i = 0, \ldots, m, \) are open, convex sets), \( A \in \mathbb{R}^{p \times n}, \) with \( \text{rank}(A) = p, \) and \( b \in \mathbb{R}^p. \)

The set for which the problem is defined is the \( \mathbb{D} := \bigcap_{i=0}^m \text{dom} f_i, \) while the feasible set is

\[
\mathbb{X} := \{ x \in \mathbb{D} \mid f_i(x) \leq 0, \ i = 1, \ldots, m, \ Ax = b \}. 
\]

We assume that the problem has a solution, say, \( x_\ast, \) and define \( p_\ast := f_0(x_\ast). \) We assume that the problem (9.1) is strictly feasible. That is, there exists \( x \in \mathbb{D} \) such that \( Ax = b \)
and \( f_i(x) < 0 \), for \( i = 1, \ldots, m \) (i.e., the Slater condition is satisfied). In this case, the KKT conditions are necessary and sufficient optimality conditions.

If we define \( \lambda_* := [\lambda_{*,1} \cdots \lambda_{*,m}]^T \), then the KKT conditions are expressed as:

\[
\nabla f_0(x_*) + \sum_{i=1}^{m} \lambda_{*,i} \nabla f_i(x_*) + A^T v_* = 0
\]

\[
Ax_* = b
\]

\[
f_i(x_*) \leq 0, \ i = 1, \ldots, m
\]

\[
\lambda_* \geq 0
\]

\[
\lambda_{*,i} f_i(x_*) = 0, \ i = 1, \ldots, m.
\]

### 9.2 Interior Point Methods

**Interior point methods** solve problem \((9.1)\) by solving a sequence of problems with affine equality constraints. These problems are “approximations” of problem \((9.1)\) and their solutions are strictly feasible for the problem \((9.1)\). Next, we describe the process.

### 9.2.1 Logarithmic barrier function

The problem \((9.1)\) can be expressed as

\[
\text{minimize} \quad f_0(x) + \sum_{i=1}^{m} I_-(f_i(x))
\]

subject to \( Ax = b \),

with

\[
I_-(u) := \left\{ \begin{array}{ll}
0, & u \leq 0, \\
\infty, & u > 0.
\end{array} \right.
\]

The cost function of the problem \((9.4)\) is, in general, non-differentiable, therefore, gradient or Newton methods cannot be used for its solution.

An approximation of the function \( I_-(u) \) is \( \hat{I}_-(u) \), which is defined as

\[
\hat{I}_-(u) := -\frac{1}{t} \log(-u),
\]

with \( \text{dom} \hat{I}_- = -\mathbb{R}_{++} \) and \( t > 0 \). The parameter \( t \) determines the approximation quality (see Figure 9.1). The function \( \hat{I}_- \) is convex and closed (it tends to infinity when \( u \) tends to 0) and it will prove extremely useful later.
9.2. INTERIOR POINT METHODS

Figure 9.1: Approximations of the function \( I_u(x) \), for \( t = 1, 10, 30 \).

First, we define the **logarithmic barrier function** for the problem (9.1)

\[
\phi(x) := -\sum_{i=1}^{m} \log(-f_i(x)),
\]

(9.7)

with \( \text{dom} \phi = \{ x \in \bigcap_{i=1}^{m} \text{dom} f_i \mid f_i(x) < 0, \ i = 1, \cdots , m \} \). As we will prove next, the function \( \phi \) is convex. Moreover, it is doubly differentiable.

Next, we define the optimization problem

\[
\text{minimize} \quad f_0(x) + \frac{1}{t} \phi(x)
\]

subject to \( Ax = b \),

(9.8)

for \( x \in \text{dom} f_0 \cap \text{dom} \phi \)\(^1\) and \( t > 0 \).

Adding the function \( \frac{1}{t} \phi(t) \) to the cost function \( f_0(x) \), intuitively, “raises a barrier” on the boundary of the feasible set \( \mathbb{X} \), essentially “trapping” the solution of the problem (9.8) inside \( \mathbb{X} \) (but allowing it to reach “very close” to its boundary, for “big” \( t \)).

The problem (9.8) is convex and the Newton algorithm for problems with affine equality constraints can be used for its solution. Obviously, other algorithms can be used as well. We will focus on the Newton algorithm.

\(^1\)Note that \( \text{dom} f_0 \cap \text{dom} \phi \subset \mathbb{X} \). In fact, the set \( \mathbb{X} \) is the closed hull of \( \text{dom} f_0 \cap \text{dom} \phi \).
Next, we study the problem

$$
\begin{align*}
\text{minimize} & \quad tf_0(x) + \phi(x) \\
\text{subject to} & \quad Ax = b
\end{align*}
$$

(9.9)

which is equivalent to the problem (9.8).

**Example 9.2.1.** Consider the simple optimization problem with inequality constraints

$$
\begin{align*}
\text{minimize} & \quad f_0(x) = \frac{1}{x} \\
\text{subject to} & \quad 0.5 \leq x \leq 2.
\end{align*}
$$

(9.10)

The function $\phi(x)$ in this case is the

$$
\phi(x) = -\log(-(x - 2)) - \log(-(0.5 - x)).
$$

(9.11)

In Figure 9.2, we plot the function $f(x) + \frac{1}{t} \phi(x)$, for $t = 1, 10, 100$. We observe that, as $t$ increases, the function $f(x) + \frac{1}{t} \phi(x)$ approaches $f(x)$ except at the edges of the interval $[0.5, 2]$ where it increases sharply, forming a “barrier.”

The first and second derivatives of $\phi(x)$ are as follows:

$$
\begin{align*}
\nabla \phi(x) &= \sum_{i=1}^{m} \frac{1}{f_i(x)} \nabla f_i(x), \\
\nabla^2 \phi(x) &= \sum_{i=1}^{m} \frac{1}{f_i^2(x)} \nabla f_i(x) \nabla f_i(x)^T - \sum_{i=1}^{m} \frac{1}{f_i(x)} \nabla^2 f_i(x).
\end{align*}
$$

(9.12)
From the form of $\nabla^2\phi(x)$ and from the fact that $f_i(x) < 0$ for $x \in \text{dom} \phi$, we easily see that $\phi(x)$ is a convex function (because the Hessian is non-negative definite).

### 9.3 Central Path

We assume that, for $t > 0$, there is a solution to the problem (9.9), say, $x_*(t)$ (hence, $f_0(x_*(t)) > -\infty$). The function $x_*(t)$ is called central path, while each point of the central path is called central point.

If $x_*(t)$ is a central point, then there exists $\hat{v} \in \mathbb{R}^p$ that satisfies the KKT conditions for the problem (9.9), that is,

$$t \nabla f_0(x_*(t)) + \nabla \phi(x_*(t)) + A^T \hat{v} = 0, \quad Ax_*(t) = b,$$

or, equivalently,

$$t \nabla f_0(x_*(t)) + \sum_{i=1}^{m} \frac{1}{-f_i(x_*(t))} \nabla f_i(x_*(t)) + A^T \hat{v} = 0, \quad Ax_*(t) = b. \quad (9.14)$$

Moreover, $x_*(t) \in \text{dom} f_0 \cap \text{dom} \phi$, therefore,

$$f_i(x_*(t)) < 0, \quad \text{for } i = 1, \ldots, m. \quad (9.15)$$

#### 9.3.1 Feasible dual points from central path points

From (9.14), we can draw the following important conclusion about the central path. Each central point, $x_*(t)$, defines a dual feasible point for the problem (9.1) and, thus, a lower bound for $p_*$. More specifically, define

$$\lambda_{*,i}(t) := \frac{1}{tf_i(x_*(t))}, \quad \text{for } i = 1, \ldots, m, \quad v_*(t) = \frac{\hat{v}}{t}. \quad (9.16)$$

It can be shown that the point $(\lambda_*, v_*)$, with $\lambda_* := [\lambda_{*,1}(t) \cdots \lambda_{*,m}(t)]^T$, is a feasible dual point for the problem (9.1).\(^2\)

The proof is as follows. By definition, we have that $\lambda_*(t) > 0$, because $t > 0$ and $f_i(x_*(t)) < 0$, for $i = 1 \ldots, m$.

\(^2\)We recall that a point $(\lambda, v)$ is binary feasible if $\lambda \geq 0$ and $g(\lambda, v) > -\infty$. 
The relation (9.14) can be expressed as

$$\nabla f_0(x_\ast(t)) + \sum_{i=1}^{m} \lambda_{\ast,i}(t) \nabla f_i(x_\ast(t)) + A^T v_\ast(t) = 0.$$  \hfill (9.17)

The Lagrange function for the problem (9.1) is

$$L(x, \lambda, v) = f_0(x) + \sum_{i=1}^{n} \lambda_i f_i(x) + v^T (Ax - b).$$ \hfill (9.18)

Let $\lambda = \lambda_\ast(t)$ and $v = v_\ast(t)$. Then, we know that

$$g(\lambda_\ast(t), v_\ast(t)) = \inf_{x \in \mathbb{D}} L(x, \lambda_\ast(t), v_\ast(t)).$$ \hfill (9.19)

The function $L(x, \lambda_\ast(t), v_\ast(t))$ is a convex function of $x$ (why?). Therefore, a necessary and sufficient condition for its minimization over the open set $\mathbb{D}$ is

$$\nabla_x L(x, \lambda_\ast(t), v_\ast(t)) = 0.$$ \hfill (9.20)

We easily see that, due to (9.17), $x_\ast(t)$ satisfies (9.20) and, therefore, minimizes $L(x, \lambda_\ast(t), v_\ast(t))$.

The dual function at point $(\lambda_\ast(t), v_\ast(t))$ is

$$g(\lambda_\ast(t), v_\ast(t)) = L(x_\ast(t), \lambda_\ast(t), v_\ast(t))$$

$$= f_0(x_\ast(t)) + \sum_{i=1}^{m} \lambda_{\ast,i}(t) f_i(x_\ast(t)) + v_\ast(t)^T (Ax_\ast(t) - b)$$ \hfill (9.21)

$$\overset{(9.16)}{=} f_0(x_\ast(t)) - \frac{m}{t}.$$  \hfill (9.21)

Since $f_0(x_\ast(t)) > -\infty$, we have that $g(\lambda_\ast(t), v_\ast(t)) > -\infty$. So, the pair $(\lambda_\ast(t), v_\ast(t))$ is a feasible dual point for the problem (9.1). Therefore,

$$g(\lambda_\ast(t), v_\ast(t)) \leq p_\ast.$$ \hfill (9.22)

Combining the relations (9.21) and (9.22), we obtain the very important inequality

$$f_0(x_\ast(t)) - p_\ast \leq \frac{m}{t}.$$ \hfill (9.23)

That is,

1. for a given $t$, the point $x_\ast(t)$ is at most $\frac{m}{t}$ suboptimal for the problem (9.1),
2. for $t \to \infty$, $x_\ast(t)$ converges to an optimal point for the problem (9.1).
9.4. THE BARRIER METHOD

9.3.2 Interpretation via KKT

We have seen that relations

\[ Ax = b, \quad f_i(x) \leq 0, \quad i = 1, \ldots, m \]
\[ \lambda \geq 0 \]
\[ \nabla f_0(x) + \sum_{i=1}^m \lambda_i \nabla f_i(x) + A^T v = 0 \]

\[ \lambda_i f_i(x) = -\frac{1}{t}, \quad i = 1, \ldots, m \]

are satisfied by \( x_*(t), \lambda_*(t), \) and \( v_*(t) \).

The relations (9.24) are the KKT conditions for the problem (9.1), with the only difference being that, instead of zero in the right-hand member of the fourth line, there is the term \(-\frac{1}{t}\).

Therefore, for “large” \( t \), \( x_*(t) \), \( \lambda_*(t) \), and \( v_*(t) \) “almost” satisfy the KKT conditions for the problem (9.1).

9.4 The Barrier Method

We have seen that the point \( x_*(t) \) is at most \( \frac{m}{t} \)-suboptimal for the problem (9.1). An approach to solving (9.1), with tolerance \( \epsilon \), is to solve the problem

\[
\begin{align*}
\text{minimize} & \quad \frac{m}{t} f_0(x) + \phi(x) \\
\text{subject to} & \quad Ax = b.
\end{align*}
\]

In general, this problem is hard to solve for “small” \( \epsilon \). An efficient approach to solving the problem (9.25) is the iterative solution of problems of the form (9.9), with increasing \( t \). The solution of the problem for a particular \( t \) is the starting point for the next problem (with a larger \( t \)).

This method, which is called the \textbf{barrier method}, is described in Table 9.4. It was originally proposed by Fiacco and McCormick in the 1960s. But, during the next decade, it was not widely accepted mainly because there were arguments against the numerical stability of the method. For example, for large \( t \), the condition number of the second derivative of the function \( t f_0 + \phi \) is very large, which was expected to lead to difficulties in the numerical implementation of the method.
The method came back during the 1980s, when it was associated with Karmakar’s algorithm for solving linear programming problems. Recently, it has been one of the most popular approaches for the solution of (not very large) convex optimization problems.

\[ x \in \text{dom } f_0 \cap \text{dom } \phi, \quad Ax = b, \quad t > 0, \quad \mu > 1, \quad \text{tolerance } \epsilon > 0. \]

repeat
1. Compute \( x_\ast(t) \), by minimizing \( tf_0 + \phi \) subject to \( Ax = b \), starting at \( x \).
2. \( x := x_\ast(t) \).
3. quit if \( \frac{m_t}{t} < \epsilon \).
3. \( t := \mu t \).

Table 9.1: Barrier method for convex optimization problems (starting from a feasible point).

### 9.5 Feasibility and Phase I

One approach towards finding an initial feasible point for the problem (9.1) is the solution of the feasibility problem

\[
\begin{align*}
\text{minimize} & \quad s \\
\text{subject to} & \quad f_i(x) \leq s, \quad i = 1, \ldots, m \\
& \quad Ax = b,
\end{align*}
\]

(9.26)

with optimization variables \( x \in \mathbb{R}^n, \ s \in \mathbb{R} \) (we recall that \( A \in \mathbb{R}^{p \times n} \) and \( b \in \mathbb{R}^p \)).

If there exists \( x \) such that \( Ax = b \), say, \( \bar{x} \), then the problem (9.26) is always strictly feasible, because we can put as initial point the vector \( (\bar{x}, \bar{s}) \) with

\[
\bar{s} \geq \max\{f_1(\bar{x}), \ldots, f_m(\bar{x})\}. \tag{9.27}
\]

If \( p < n \), then the system \( Ax = b \) has infinite solutions (recall that we have assumed that the rows of \( A \) are linearly independent). One solution is given by the relationship \( \bar{x} = A^\dagger b \), where \( A^\dagger \) is the pseudoinverse of \( A \).

Let \( \bar{x}_\ast \) be the optimal point and \( \bar{p}_\ast \) the solution of the problem (9.26). Then

1. if \( \bar{p}_\ast > 0 \), then the problem (9.1) is infeasible.
2. if \( \bar{p}_* < 0 \), then the problem (9.1) is feasible and a feasible point for problem (9.1) is \( \bar{x}_* \).

3. if \( \bar{p}_* = 0 \), then the problem is not strictly feasible.

In practice, it may happen that \( |\bar{p}_*| < \epsilon \), for a very small \( \epsilon > 0 \).

Then, the inequalities \( f_i(x) \leq -\epsilon \), for \( i = 1, \ldots, m \), are not feasible, while the inequalities \( f_i(x) \leq \epsilon \), for \( i = 1, \ldots, m \), are feasible.

We note that it is not necessary to solve the problem (9.26). It suffices to find a \( x \) which leads to \( s < 0 \).

9.6 Primal–Dual Interior Point Method

9.6.1 Primal–dual direction

We define

\[
    f(x) := \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}, \quad Df(x) := \begin{bmatrix} \nabla f_1(x)^T \\ \vdots \\ \nabla f_m(x)^T \end{bmatrix}. \tag{9.28}
\]

We define the vector \( r_t(x, \lambda, v) \) as follows:

\[
    r_t(x, \lambda, v) := \begin{bmatrix} \nabla f_0(x) + Df(x)^T \lambda + A^T v \\ -\text{diag}(\lambda)f(x) - \frac{1}{\tau} 1 \\ A x - b \end{bmatrix}. \tag{9.29}
\]

If \( x \in \text{dom } f_0 \cap \text{dom } \phi, \lambda > 0 \), and \( v \) satisfy the equation \( r_t(x, \lambda, v) = 0 \), then \( x = x_*(t) \), \( \lambda = \lambda_*(t) \), and \( v = v_*(t) \).

In particular, \( x_*(t) \) is a primal feasible point and \((\lambda_*(t), v_*(t))\) is a dual feasible point, with duality gap \( \frac{m}{\tau} \).

Suppose we are at the point \( y = (x, \lambda, v) \), with \( x \in \text{dom } f_0 \cap \text{dom } \phi \) and \( \lambda > 0 \).

The vector

\[
    r_{t,d}(y) := \nabla f_0(x) + Df(x)^T \lambda + A^T v \tag{9.30}
\]

is called dual residual, the vector

\[
    r_{t,p}(y) := Ax - b \tag{9.31}
\]
is called **primal residual**, while the vector

\[ r_{t,c}(y) := -\text{diag}(\lambda)f(x) - \frac{1}{t} 1, \]

(9.32)

where \( 1 \) is the \( m \)-dimensional vector with unit elements, is called **centrality residual**.

The first-order approximation of \( r_t \), at the point \( y \), is as follows:

\[ r_t(y + z) \approx r_t(y) + Dr_t(y)z. \]

(9.33)

We define as the Newton step \( \Delta y = (\Delta x, \Delta \lambda, \Delta v) \), at the point \( y \), the vector \( z \) for which \( r_t(y) \) vanishes, i.e.

\[ Dr_t(y)\Delta y = -r_t(y). \]

(9.34)

Calculating the derivative \( Dr_t \), at the point \( y \), we obtain

\[
\begin{bmatrix}
\nabla^2 f_0(x) + \sum_{i=1}^m \lambda_i \nabla^2 f_i(x) & Df(x)^T & A^T & \Delta x \\
-\text{diag}(\lambda)Df(x) & -\text{diag}(f(x)) & O & \Delta \lambda \\
A & O & O & \Delta v
\end{bmatrix}
= -
\begin{bmatrix}
\nabla^2 f_0(x) + \sum_{i=1}^m \lambda_i \nabla^2 f_i(x) & Df(x)^T & A^T & \Delta x \\
-\text{diag}(\lambda)Df(x) & -\text{diag}(f(x)) & O & \Delta \lambda \\
A & O & O & \Delta v
\end{bmatrix}
\begin{bmatrix}
r_{t,d}(y) \\
r_{t,c}(y) \\
r_{t,p}(y)
\end{bmatrix}.
\]

(9.35)

We define as **primal-dual search direction**, \( \Delta y_{pd} := (\Delta x_{pd}, \Delta \lambda_{pd}, \Delta v_{pd}) \), the Newton step \( \Delta y \).

The variables \( x_k \) and \((\lambda_k, v_k)\) of the primal–dual interior point algorithm are not necessarily feasible points of the problem (9.1), but only asymptotically (i.e., at the convergence of the algorithm). This means that we have no obvious way to compute the duality gap during algorithm execution.

For \( x \in \text{dom } f_0 \cap \text{dom } \phi \) and \( \lambda > 0 \), we define the **surrogate duality gap** as

\[ \hat{\eta}(x, \lambda) := -f(x)^T \lambda. \]

(9.36)

The quantity \( \hat{\eta}(x, \lambda) \) would be equal to the duality gap \( \frac{m}{t} \) if \( x \) and \( \lambda \) were feasible points for the problem (9.1).

The value of \( t \) which corresponds to the surrogate duality gap \( \hat{\eta} \) equals \( \frac{m}{\hat{\eta}} \).

### 9.6.2 Line search

Let

\[ x_+ := x + s\Delta x_{pd}, \lambda_+ := \lambda + s\Delta \lambda_{pd}, v_+ := v + s\Delta v_{pd}. \]

(9.37)

A line search method should guarantee that \( \lambda_+ > 0, f(x_+) < 0 \) and the function \( \| r(\cdot) \|_2 \) decreases quite a bit (see page 613 Boyd & Vandenberghe). We proceed as follows:
1. Compute the quantity $s_{\text{max}} := \sup\{s \in [0, 1] | \lambda + s\Delta\lambda > 0\}$.

2. Backtrack by $\beta \in (0, 1)$ until $f(x_+) < 0$.

3. Backtrack by $\beta \in (0, 1)$ until

$$\|r_t(x_+, \lambda_+, v_+)/(2) \leq (1 - \alpha s)\|r_t(x, \lambda, v)/(2$$

for $a \in (0, 0.5)$.

The relation (9.38) is compatible with backtracking line search we have encountered in previous algorithms. Thus, it can be shown that backtracking line search will always terminate after a finite number of steps.

### 9.6.3 Primal-dual algorithm for convex optimization problems

The primal-dual algorithm is presented next. We note that we should start from a point that strictly satisfies all inequality constraints.

\[
\begin{align*}
\text{x} & \in \text{dom} f_0 \cap \text{dom} \phi, \ \lambda > 0, \ t > 0, \ \mu > 1, \ e_{\text{feas}} > 0, \ e > 0. \\
\text{repeat} & \\
1. & t := \mu \frac{m}{\eta}. \\
2. & \text{Compute } \Delta y_{\text{pd}}. \\
3. & \text{perform line search and choose step } s. \\
4. & y := y + s\Delta y_{\text{pd}}. \\
\text{until } & (\|r_p\|_2 < e_{\text{feas}}, \ \|r_d\|_2 < e_{\text{feas}}, \ \hat{\eta} < \epsilon). 
\end{align*}
\]