2 Theoretical Basics

In this chapter, we introduce the basic concepts of multiobjective optimization. We introduce basic definitions and derive a concept of optimality for multiobjective optimization problems. Based on this, we formulate the central optimization problem that we study throughout this book and introduce a relaxed optimization problem that we use in order to solve the central optimization problem. Throughout this book $\mathbb{N}$ denotes the set of natural numbers, $\mathbb{Z}$ the set of integers, $\mathbb{R}$ the set of real numbers and $\mathbb{R}_+$ the set of nonnegative real numbers.

2.1 Basics of multiobjective optimization

In this section, we set the basis for comparing vectors in $\mathbb{R}^p$. Therefore, we use the pointed convex cone $K = \mathbb{R}^p_+$, which induces a partial order relation $\leqslant$ on $\mathbb{R}^p$ defined by

$$z^1 \leqslant z^2 :\Leftrightarrow z^2 - z^1 \in \mathbb{R}^p_+$$

for elements $z^1, z^2 \in \mathbb{R}^p$, also called the componentwise ordering in $\mathbb{R}^p$. Obviously, it holds

$$z^1 \leqslant z^2 \Leftrightarrow z^1_i \leqslant z^2_i \text{ for all } i \in \{1, ..., p\}.$$

S. Rocktäschel, *A Branch-and-Bound Algorithm for Multiobjective Mixed-integer Convex Optimization*, BestMasters, https://doi.org/10.1007/978-3-658-29149-5_2

Note that the partial order $\leqslant$ on $\mathbb{R}^p$ is not a total order relation for $p \geqslant 2$. Additionally, we use the following notations:

- $z^1 \lneq z^2 :\Leftrightarrow z^2 - z^1 \in \mathbb{R}^p_+ \backslash \{0_p\}$ and
- $z^1 < z^2 :\Leftrightarrow z^2 - z^1 \in \mathrm{int}(\mathbb{R}^p_+)$,

where $\mathrm{int}(\mathbb{R}^p_+)$ is *the interior* of the set $\mathbb{R}^p_+$.
The following definition allows us to consider projections of subsets $Z \subseteq \mathbb{R}^p$.

Definition 2.1. *Let the set $Z \subseteq \mathbb{R}^p$ be nonempty and $j \in \{1, ..., p\}$. We define the j-projection of Z as*

$$Z_j := \{z_j \in \mathbb{R} \mid \exists \tilde{z} \in Z \colon \tilde{z}_j = z_j\}.$$

Remark 2.2. *We recall that for nonempty and compact set $Z \subseteq \mathbb{R}^p$ all projections $Z_1, ..., Z_p$ of Z are nonempty and compact.*

The following definition generalizes the concepts of minimality and maximality from scalar optimization to multiobjective optimization and introduces the ideal- and the anti-ideal point of a set $Z \subseteq \mathbb{R}^p$.

Definition 2.3. *Let $Z \subseteq \mathbb{R}^p$ be a given set. An element $z^* \in Z$ is called*

- *a minimal element of Z, if $Z \cap (\{z^*\} - \mathbb{R}^p_+) = \{z^*\}$ and*
- *a maximal element of Z, if $Z \cap (\{z^*\} + \mathbb{R}^p_+) = \{z^*\}$.*

Additionally,

- *the set of minimal elements of Z is given by*

$$\min(Z) := \{z^* \in Z \mid Z \cap (\{z^*\} - \mathbb{R}^p_+) = \{z^*\}\}$$

and

- *the set of maximal elements of Z is given by*

$$\max(Z) := \{z^* \in Z \mid Z \cap (\{z^*\} + \mathbb{R}^p_+) = \{z^*\}\}.$$

Furthermore, if Z is nonempty and compact, we define

- *the ideal point $\underline{\min}(Z)$ of Z as the element $z^* \in \mathbb{R}^p$ with $z^*_j = \min(Z_j) = \min\{z_j \in \mathbb{R} \mid z \in Z\}$ for all $j \in \{1, ..., p\}$ and*

- *the anti-ideal point $\overline{\max}(Z)$ of Z as the element $z^* \in \mathbb{R}^p$ with $z^*_j = \max(Z_j) = \max\{z_j \in \mathbb{R} \mid z \in Z\}$ for all $j \in \{1, ..., p\}$.*

Thereby, $\min(Z_j)$ and $\max(Z_j)$ for $j \in \{1, ..., p\}$ refer to the scalar minimum and maximum of the set Z_j.

Remark 2.4. *The above definitions of $\underline{\min}(Z)$ and $\overline{\max}(Z)$ are well defined, because the respective minima and maxima exist due to the compactness and non emptiness of the set Z and hence, of the projections.*
Furthermore, it is easy to see that $\{z \in Z \mid z \leqslant z^*\} = \{z^*\}$ holds for all $z^* \in \min(Z)$. Analogously, $\{z \in Z \mid z^* \leqslant z\} = \{z^*\}$ holds for all $z^* \in \max(Z)$.
Additionally,

$$\begin{aligned}
\min(-Z) &= \{z \in \mathbb{R}^p \mid (-Z) \cap (\{z\} - \mathbb{R}^p_+) = \{z\}\} \\
&= \{-z \in \mathbb{R}^p \mid (-Z) \cap (\{-z\} - \mathbb{R}^p_+) = \{-z\}\} \\
&= \{-z \in \mathbb{R}^p \mid Z \cap (\{z\} + \mathbb{R}^p_+) = \{z\}\} \\
&= -\{z \in \mathbb{R}^p \mid Z \cap (\{z\} + \mathbb{R}^p_+) = \{z\}\} \\
&= -\max(Z).
\end{aligned}$$

This equation can be used to reformulate maximizing problems as minimizing problems and vice versa.

The following example illustrates the above definitions.

Example 2.5. *We consider the set* $Z = \{z \in \mathbb{R}^2 \mid ||z||_2 \leqslant 1\}$*.*
The minimum of Z *is given by* $\min(Z) = \{z \in \mathbb{R}^2 \mid ||z||_2 = 1, z_1 \leqslant 0, z_2 \leqslant 0\}$ *and the maximum of* Z *is given by* $\max(Z) = \{z \in \mathbb{R}^2 \mid ||z||_2 = 1, z_1 \geqslant 0, z_2 \geqslant 0\}$*. Furthermore, it holds* $\underline{\min}(Z) = (-1, -1)^\top$ *and* $\overline{\max}(Z) = (1, 1)^\top$*. The sets and points are illustrated in Figure 2.1.*

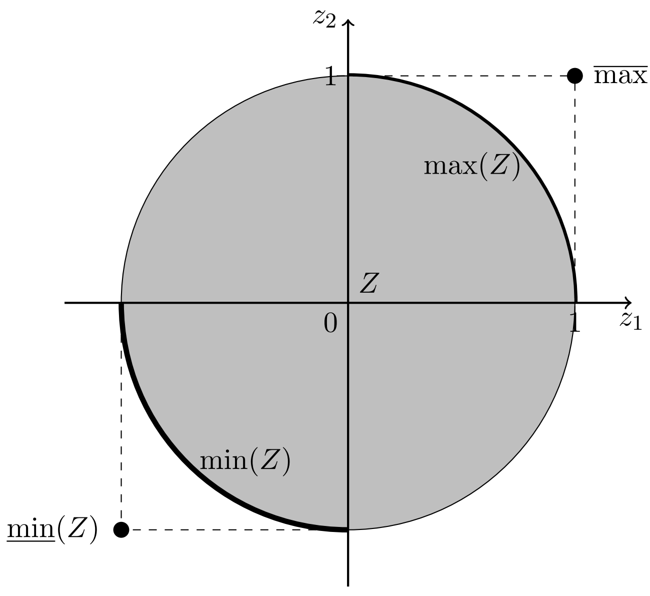

Figure 2.1: the sets and points from Example 2.5

2.2 The central multiobjective mixed-integer optimization problem

In this section, we introduce basic notations and definitions that allow us to formulate the central optimization problem of this book. At first, we recall the definition of intervals and generalize this concept for the multidimensional case.
At first, we introduce intervals and boxes [15].

Definition 2.6.

(i) For $\underline{b}$, $\overline{b} \in \mathbb{R}$ with $\underline{b} \leqslant \overline{b}$ we define the nonempty, compact interval B by

$$B = [\underline{b}, \overline{b}] := \{b \in \mathbb{R} \mid \underline{b} \leqslant b \leqslant \overline{b}\}.$$

(ii) We denote the set of all of these intervals by $\mathbb{IR} := \{[\underline{b}, \overline{b}] \subseteq \mathbb{R} \mid \underline{b}, \overline{b} \in \mathbb{R}, \underline{b} \leqslant \overline{b}\}$.

Now, we introduce the definition of boxes that are a generalization of intervals for the multidimensional case.

Definition 2.7.

(i) For $\underline{b}, \overline{b} \in \mathbb{R}^r$ with $\underline{b} \leqslant \overline{b}$ we define the nonempty, compact, r-dimensional Box B by

$$B = [\underline{b}, \overline{b}] := \{b \in \mathbb{R}^r \mid \underline{b} \leqslant b \leqslant \overline{b}\} = \bigtimes_{i=1}^{r} [\underline{b}_i, \overline{b}_i].$$

(ii) We denote the set of all of these r-dimensional boxes in $\mathbb{R}^r$ by $\mathbb{IR}^r$.

(iii) *Let $B, \tilde{B} \in \mathbb{R}^r$ be boxes with $\tilde{B} \subseteq B$. Then we call $\tilde{B}$ a subbox of B.*

We introduce notations for certain parameters of boxes.

Definition 2.8. *For $B = [\underline{b}, \overline{b}] \in \mathbb{R}^r$ we define*

(i) *the infimum of B as $\inf(B) := \underline{b}$,*

(ii) *the supremum of B as $\sup(B) := \overline{b}$,*

(iii) *the midpoint of B as $\text{mid}(B) := \frac{\underline{b}+\overline{b}}{2}$ and*

(iv) *the width of B as $\text{wid}(B) := \max_{i \in \{1,...,r\}} (\overline{b}_i - \underline{b}_i)$.*

Furthermore, we call elements $b \in B$ with $b_i \in \{\underline{b}_i, \overline{b}_i\}$ for all $i \in \{1, ..., r\}$ vertices of B.

Obviously, boxes are convex sets.
Moreover, we will use the following notations. Let $B \subseteq \mathbb{R}^r$ be a set and $f : B \to \mathbb{R}$ a function that is twice continuously differentiable on an open superset $\overline{B}$ of B. We define $f(\tilde{B}) := \{f(b) \mid b \in \tilde{B}\}$ for $\tilde{B} \subseteq B$. Furthermore, for an element $b \in \overline{B}$ we denote the gradient of f in b as $\nabla f(b)$ and the Hessian of f in b as $H_f(b)$. In addition, $\lambda_{\min_f}(b)$ denotes the smallest eigenvalue of $H_f(b)$. Since f is twice continuously differentiable $\lambda_{\min_f}(b)$ is well defined, because $H_f(b)$ is symmetric for all $b \in \overline{B}$ and hence, every eigenvalue of $H_f(b)$ is real.
The following assumptions will be used to define the central optimization problem of this book and to prove associated theoretical results.

Assumption 2.9. *Let $X = [\underline{x}, \overline{x}] \in \mathbb{R}^m$, $Y = [\underline{y}, \overline{y}] \in \mathbb{R}^n$ be boxes for $m, n \in \mathbb{N}$ and $B := X \times Y \in \mathbb{R}^r$ with $r := m + n$. Additionally, we assume $\underline{y}, \overline{y} \in \mathbb{Z}^n$.*

Furthermore, let the functions $f\colon B \to \mathbb{R}^p$ *and* $g\colon B \to \mathbb{R}^q$ *for* $p, q \in \mathbb{N}$, $p \geqslant 2$ *be convex and twice continuously differentiable on an open superset* $\overline{B}$ *of* B*.*

Remark 2.10. *For our purposes, the assumption* $\underline{y}, \overline{y} \in \mathbb{Z}^n$ *can be made without loss of generality, since we could simply round these vectors up or down respectively and would obtain equivalent optimization problems.*
We will also make an additional assumption on regularity of certain optimization problems in Assumption 3.9 later on.

Using the notation of Assumption 2.9, we introduce notations for subsets of B that will be used later on.

Definition 2.11. *Let Assumption 2.9 be fulfilled and let* $\tilde{B} = \tilde{X} \times \tilde{Y} \subseteq B$ *with* $\tilde{X} \in \mathbb{R}^m$ *and* $\tilde{Y} \in \mathbb{R}^n$ *be a given subbox of* B*. We define*

$$\tilde{B}^g := \{(x, y) \in \tilde{X} \times \tilde{Y} \mid g(x, y) \leqslant 0_q\} \text{ and}$$

$$\tilde{B}^{\mathbb{Z}} := \{(x, y) \in \tilde{X} \times \tilde{Y} \mid y \in \mathbb{Z}^n\}, \text{ as well as}$$

$$\tilde{B}^{g,\mathbb{Z}} := \tilde{B}^g \cap \tilde{B}^{\mathbb{Z}}.$$

Now, we are able to formulate the central optimization problem that we are going to study in this book. Under Assumption 2.9 we are going to study the multiobjective mixed-integer convex problem

$$\begin{aligned} \min\ & f(x, y) \\ \text{s.t.}\ & g(x, y) \leqslant 0_q \\ & (x, y) \in X \times Y = B \\ & y \in \mathbb{Z}^n. \end{aligned} \tag{MOMICP}$$

Using the notations from above, we can also write (MOMICP) as

$$\begin{aligned} \min\ & f(b) \\ \text{s.t.}\ & b \in B^{g,\mathbb{Z}}. \end{aligned}$$

Remark 2.12. *Although (MOMICP) is a generalization of multi-objective convex optimization problems, this approach cannot be used in order to solve the interesting class of optimization problems defined by*

$$\begin{aligned} \min\ & y^\top \tilde{f}(x) \\ \text{s.t.}\ & y^\top \tilde{g}(x) \leqslant 0_q \\ & x \in X \\ & y \in \{0,1\}^n \end{aligned} \tag{P_y}$$

with $X \in \mathbb{R}^m$, $\tilde{f}\colon X \to \mathbb{R}^{n\times p}$ *and* $\tilde{g}\colon X \to \mathbb{R}^{n\times q}$, *where* $\tilde{f}_{i,j}$ *and* $\tilde{g}_{i,k}$ *are convex and twice continuously differentiable on an open superset* $\overline{X}$ *of* X *for all* $i \in \{1,...,n\}$, $j \in \{1,...,p\}$, $k \in \{1,...,q\}$.
This is due to the fact that we demand convexity of f *and* g *in Assumption 2.9. With* $Y := [0,1]^n$, $B := X \times Y$, $f\colon B \to \mathbb{R}^p$ *defined by* $f(x,y) := y^\top \tilde{f}(x)$ *and* $g\colon B \to \mathbb{R}^q$ *defined by* $g(x,y) := y^\top \tilde{g}(x)$ *we can rewrite* (P_y) *in the form of (MOMICP). However, the functions* f *and* g *are not necessarily convex in this case, even not if* $\tilde{f}$ *and* $\tilde{g}$ *are linear. We illustrate this in Example 2.13.*

Example 2.13. *Consider* $X = [-1,1]$, $n = 2$, $q = 1$ *and* $\tilde{g}\colon X \to \mathbb{R}^2$ *with* $\tilde{g}(x) := (x,-x)^\top$ *in* (P_y). *Note that* $\tilde{g}$ *is linear and hence convex. Then we define* $g\colon X \times [0,1]^2 \to \mathbb{R}$ *by* $g(x,y) := y^\top \tilde{g}(x)$.
For $(x,y) = (-1,(0,0)^\top) \in X \times [0,1]^2$, $(x',y') = (1,(0,1)^\top) \in X \times$

$[0,1]^2$ *and* $\lambda = \frac{1}{2} \in [0,1]$ *we obtain*

$$\begin{aligned} g(\lambda(x,y) + (1-\lambda)(x',y')) &= g((-\frac{1}{2},(0,0)^\top) + (\frac{1}{2},(0,\frac{1}{2})^\top)) \\ &= g(0,(0,\frac{1}{2})^\top) \\ &= (0,\frac{1}{2}) \cdot \tilde{g}(0) \\ &= (0,\frac{1}{2}) \cdot (0,0)^\top = 0, \end{aligned}$$

but

$$\begin{aligned} \lambda g(x,y) + (1-\lambda)g(x',y') &= \lambda(0,0) \cdot \tilde{g}(-1) + (1-\lambda)(0,1) \cdot \tilde{g}(1) \\ &= \frac{1}{2}(0,0) \cdot (-1,1)^\top + \frac{1}{2}(0,1) \cdot (1,-1)^\top \\ &= 0 - \frac{1}{2} = -\frac{1}{2} < 0. \end{aligned}$$

Therefore, g *is not convex.*

Note that we could also find an example where the objective functions f of (P_y) are nonconvex in an analog way.
At least, there are possibilities that allow us to neglect the assumption of convexity of f for (MOMICP). We will outline one approach for this in Section 6.
Now, in order to derive a concept of optimality for the central multiobjective optimization problem (MOMICP), we introduce the concept efficiency for multiobjective optimization problems

$$\begin{aligned} &\min\ f(b) \\ &\text{s.t.}\ b \in B, \end{aligned} \qquad \text{(MOP)}$$

where $B \subseteq \mathbb{R}^r$ is a given set with $r \in \mathbb{N}$ and $f\colon B \to \mathbb{R}^p$ is a given function with $p \in \mathbb{N}$. In the following definitions we introduce the concepts of efficiency and nondominated points [11], [8].

Definition 2.14.

(i) A point $b^* \in B$ is called efficient for (MOP), if there is no $b \in B$ with

$$f(b) \lneq f(b^*).$$

(ii) Let $b, b^* \in B$ with

$$f(b) \lneq f(b^*).$$

Then we say b dominates b^*.

(iii) The set of efficient points for (MOP) is called the efficient set of (MOP).

(iv) A set $\mathcal{L} \subseteq \mathcal{P}(\mathbb{R}^r)$ is called a cover of the efficient set E of (MOP), if $E \subseteq \bigcup \mathcal{L} := \bigcup_{\tilde{B} \in \mathcal{L}} \tilde{B}$ holds.

Remark 2.15. *Referring to Definition 2.3, we observe that a feasible point $b^* \in B$ is efficient for (MOP) if, and only if, $f(b^*)$ is a minimal element of $f(B)$ or in other words $f(b^*) \in \min(f(B))$.*

The above definitions consider points in the pre-image space of f. Additionally, we introduce definitions for points in the image space of f.

Definition 2.16.

(i) A point $z^* = f(b^*)$ is called nondominated for (MOP), if $b^* \in B$ is efficient for (MOP).

(ii) The set of all nondominated points of (MOP) is called the nondominated set of (MOP).

(iii) *For elements* $z^1, z^2 \in \mathbb{R}^p$ *with* $z^1 \lneq z^2$ *we say* z^1 *dominates* z^2*.*

(iv) *A set* $Z \subseteq \mathbb{R}^p$ *is called stable, if there are no* $z^1, z^2 \in Z$ *with* $z^1 \lneq z^2$*.*

Remark 2.17. *Referring to Definition 2.3, we observe that a point* $z^* \in \mathbb{R}^p$ *is nondominated if, and only if,* $z^* \in \min(f(B))$*. Hence, the nondominated set of (MOP) is equal to* $\min(f(B))$*.*

2.3 A relaxation of (MOMICP)

In this section, deriving from the optimization problem (MOMICP), we introduce a relaxed optimization problem (ROP($\tilde{B}$)) for a given subbox $\tilde{B}$ of B. We use (ROP($\tilde{B}$)) to determine lower bounds for f on $\tilde{B}^{g,\mathbb{Z}}$. These lower bounds will be sets $L \subseteq \mathbb{R}^p$ of a 'simple' structure that fulfill $f(\tilde{B}^{g,\mathbb{Z}}) \subseteq L + \mathbb{R}^p_+$.
Under Assumption 2.9, the relaxed optimization problem we are interested in is given by

$$\begin{aligned} &\min\ f(b) \\ &\text{s.t.}\ \ b \in \tilde{B}^g. \end{aligned} \qquad \text{(ROP}(\tilde{B}))$$

Obviously, the set of feasible points for (MOMICP) $B^{g,\mathbb{Z}}$ is a subset of the set of feasible points for (ROP(B)), which is B^g. Because of this, every feasible point for (MOMICP) is feasible for (ROP(B)) and we call the optimization problem (ROP(B)) a *relaxation* of (MOMICP). Since there are no integer constraints tied to the relaxation (ROP(B)), it is a multi-objective (continuous) convex optimization problem. This type of optimization problems has already been studied extensively.