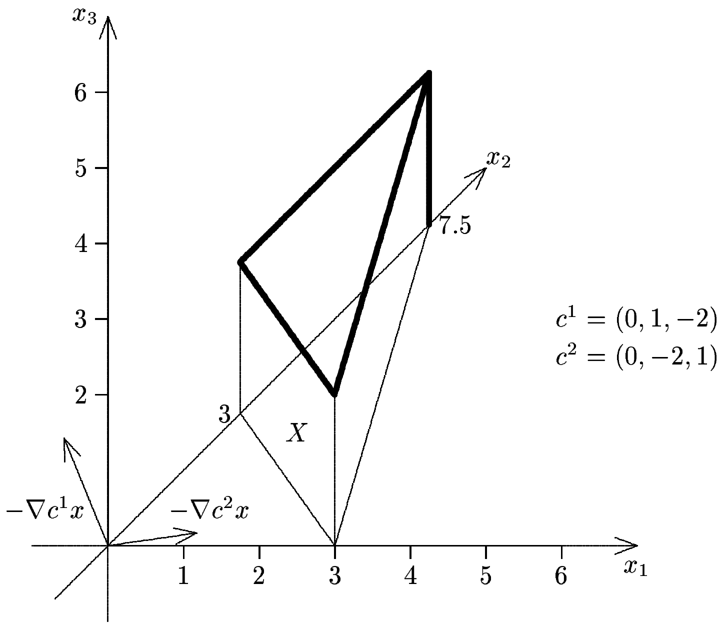
UNIVERSITY OF KAISERSLAUTERN
Department of Mathematics



Multicriteria Optimization

Dr. Matthias Ehrgott

Preface

Life is about decisions. Decisions, no matter if taken by a group or an individual, involve several conflicting objectives. The observation that real world problems have to be solved optimally according to criteria, which prohibit an „ideal“ solution – optimal for each decisionmaker under each of the criteria considered –, has led to the development of multicriteria optimization.

From its first roots, which where laid by Pareto at the end of the 19th century the discilpine has prospered and grown, especially during the last three decades. Today, many decision support systems incorporate methods to deal with conflicting objectives. The foundation for such systems is a mathematical theory of optimaztion under multiple objectives.

With this manuscript, which is based on lectures I taught in the winter semester 1998/99 at the University of Kaiserslautern, I intend to give an introduction to and overview of this fascinating field of mathematics. I tried to present theoretical questions such as existence of solutions as well as methodological issues and hope the reader finds the balance not too heavily on one side. The interested reader should be able to find classical results as well as up to date research. The text is accompanied by exercises, which hopefully help to deepen students' understanding of the topic.

I am indebted to the many researchers in the field, on whose work the lectures and manuscripts are based. Also, I would like to thank the students who followed my class and to my colleagues of the working group. They contributed with their questions and comments. Last but not least my gratitude goes to Stefan Zimmermann, whose diligence and aptitude in preparing the manuscript was enormous.

Matthias Ehrgott
April 1999

Contents

Preface i

Contents iii

1 Introduction 1

1.1 Optimization with Multiple Criteria 1

1.2 Decision Spaces and Objective (Criterion) Space 2

1.3 Notions of Optimality 4

1.4 Orderings and Cones 5

1.5 Classification 9

1.6 Exercises to Chapter 1 11

2 Pareto Optimality and Efficiency 13

2.1 Pareto Optimal and Efficient Points 13

2.2 Weak and Strict Pareto Optimal Points 20

2.3 Properly Pareto Optimal / Efficient Points 27

2.4 Exercises to Chapter 2 36

3 Weighted Sum Scalarization 39

3.1 Scalarization and Efficiency 40

3.2 Scalarization and Weak / Proper Efficiency 41

3.3 Connectedness of Y_{eff} and X_{Par} 49

3.4 Exercises to Chapter 3 52

4 Methodology 55

4.1 Bounds on the Efficient Set 55

4.2 The ε-Constraint Method 58

4.3 Benson's Method 60

4.4 Compromise Solutions — Approximation of the Ideal Point 64

4.5 Exercises to Chapter 4 70

5 Multicriteria Linear Programming **73**

5.1 Introduction 73

5.2 Theory of MCLP 78

5.3 A Multicriteria Simplex Algorithm 83

5.4 Identifying Scalarization Vectors and Pareto Faces 91

5.5 Exercises to Chapter 5 94

6 Other Optimality Concepts **97**

6.1 Lexicographic Optimization 97

6.2 Max-Ordering Optimization 99

6.3 Lexicographic Max-Ordering 101

6.4 Exercises to Chapter 6 108

Bibliography **109**

List of Figures **113**

Index **115**

Chapter 1

Introduction

1.1 Optimization with Multiple Criteria

An optimization problem is to choose among a set of „alternatives" an „optimal one". Optimality refers to certain criteria, according to which the quality of the alternatives is measured.

Example 1.1. To decide which new car to buy you consider a VW Golf, an Opel Astra, a Ford Mondeo and a Toyota Avensis. The decision will be made according to price ($\rightarrow$ cheap), engine efficiency (i.e. $\frac{\mathrm{l}}{100\mathrm{km}}$ $\rightarrow$ low) and horsepower ($\rightarrow$ high). Here you have 4 alternatives and 3 criteria.

		Alternatives			
		VW	Opel	Ford	Toyota
	Price (TDM)	31	29	30	27
Criteria	$\frac{l}{100km}$	7.2	7.0	7.5	7.8
	horsepower (HP)	90	75	80	75

Which is the best alternative ?

Note that with each one of the 3 criteria a decision is easy.

Example 1.2. For the construction of a water dam an electrical power plant is interested in maximizing storage capacity while at the same time minimizing water loss due to evaporation and construction cost. The decision has to take into account man months devoted to the construction, the mean radius of the lake, and respect certain constraints such as minimal strength of the dam. Here, the set of alternatives (possible dams) is a whole continuum and the criteria are functions of the decision variables to be maximized or minimized. The criteria are conflicting: the minima of each criterion are not optimal for others.

Optimization problems with a countable number of alternatives are called **discrete**, others **continuous**.

Example 1.3. Two criteria and one decision variable

$$f_1(x) = \sqrt{x+1}, \qquad f_2(x) = x^2 - 4x + 5 \tag{1.1}$$

$$\text{„}\min_{x \geq 0}\text{“} \; (f_1(x), f_2(x)) \tag{1.2}$$

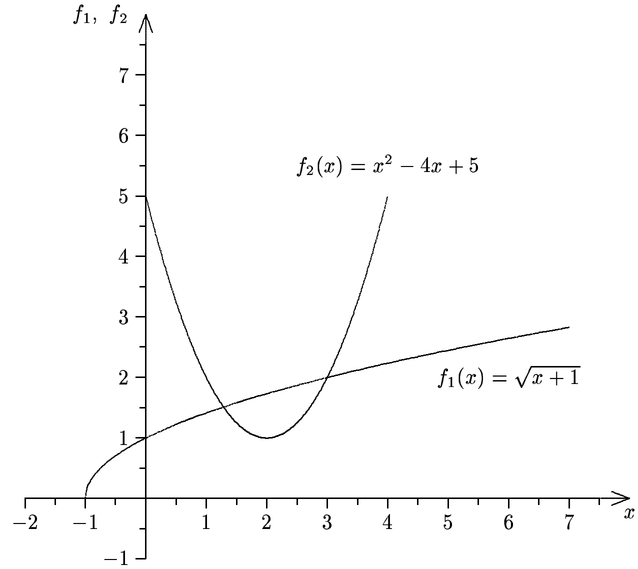


Figure 1.1: Objective Functions of Example 1.3

What are the „minima“ ?

(Again for each function individually the problem is easy: $x_1 = 0$ for f_1 and $x_2 = 2$ for f_2 are the minimizers.)

Pareto, 1906:

„We will say that the members of a collectivity enjoy **maximum** ophelimity in a certain position when it is **impossible to find a way of moving from that position very slightly in such a manner that the ophelimity enjoyed by each of the individuals of that collectivity increases or decreases**. That is to say, any small displacement in departing from that position necessarily has the effect of increasing the ophelimity which certain individuals enjoy, anf decreasing that which others enjoy, of being agreeable to some and disagreeable to others.“

Consequence: In Example 1.1 all alternatives enjoy „maximum ophelimity“, in Example 1.3 all points in $[0, 2]$ (in $[0, 2]$ one of the functions is increasing, the other decreasing). These are called today **Pareto optimal solutions** of a multiple criteria optimization problem.

1.2 Decision Spaces and Objective (Criterion) Space

Let us consider Example 1.1 again with price and efficiency only. We can illustrate this in a two-dimensional coordinate system:

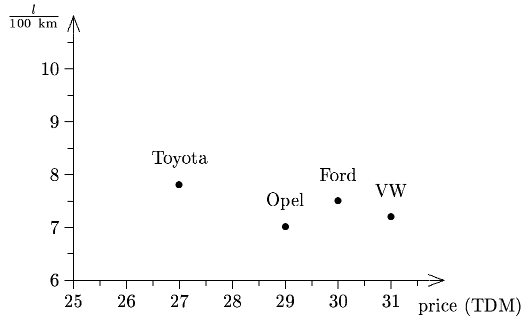


Figure 1.2: Criterion Space in Example 1.1

Here, it is easy to see that Opel and Toyota are Pareto optimal choices. (Both Ford and VW are more expensive and less efficient than Opel.)

We call $X = \{\text{VW, Opel, Ford, Toyota}\}$ the **feasible set** (set of alternatives) of the optimization problem. Denote the price by f_1, the efficiency by f_2 then $f_i : X \to \mathbb{R}$ are criteria or objective functions and the optimization problem is

$$„\min_{x \in X}“ \ (f_1(x), f_2(x)) \ . \tag{1.3}$$

The image of X under $f = (f_1, f_2)$ is $f(X)$.
For Example 1.3 we have

$$X = \{x \in \mathbb{R} : x \geq 0\} \qquad \text{as feasible set} \tag{1.4}$$

$$f_1(x) = \sqrt{1+x}, \qquad f_2(x) = x^2 - 4x + 5 \qquad \text{as objective functions.} \tag{1.5}$$

So we can use $x = (f_1)^2 - 1$ to get $f_2 = ((f_1)^2 - 1)^2 + 4 - 4 \cdot (f_1)^2 + 5 = (f_1)^4 - 6 \cdot (f_1)^2 + 10$ to obtain a picture similar to that for Example 1.1:

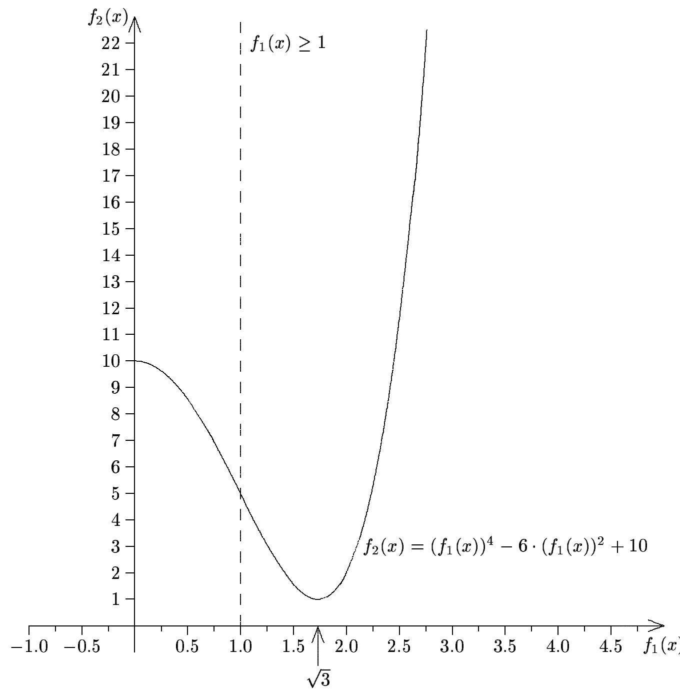


Figure 1.3: Criterion Space in Example 1.3

Pareto optimal solutions $[0, 2]$ correspond to values of f_1 in $[1, \sqrt{3}]$.

In this problem the feasible set $X \subset \mathbb{R}$, the decision space, and $f(X) \subset \mathbb{R}^2$ the objective (criterion) space. Our first drawing for Example 1.3 is in decision space, the second in criterion space.

The image of Pareto optimal points:

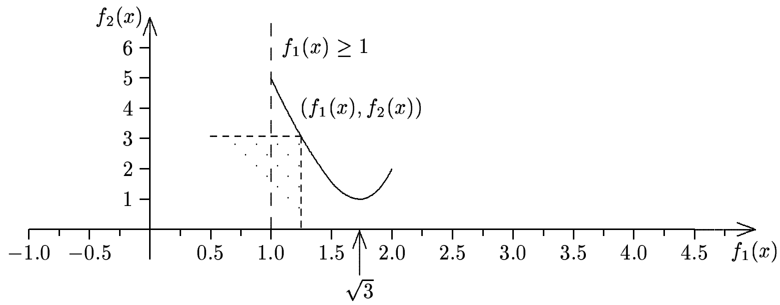


Figure 1.4: Efficient Points in Example 1.3

There is no <u>other</u> point $y \in f(X)$ such that $y^1 \leq f_1(x)$ and $y^2 \leq f_2(x)$ for any $x \in [0, 2]$: $(f_1(x), f_2(x))$ is called an **efficient point**.

The set of all efficient points is the image of the set of Pareto optimal points under the objective function.

The objective space is very useful in multicriteria optimization. However, figures like above are usually not available.

In the examples we have many Pareto optimal solutions. What is their use for finding an „optimal decision“ ?

1.3 Notions of Optimality

A multicriteria optimization problem can be written as

$$\begin{aligned} &\text{„ min “ } (f_1(x), \ldots, f_Q(x)) \\ &\text{subject to } x \in X \end{aligned} \tag{1.6}$$

But what does „minimize“ really mean ?

We have discovered Pareto optimality before. Any x which is not Pareto optimal cannot represent an optimal decision, because $\exists\ \overline{x} \in X \quad f_i(\overline{x}) \leq f_i(x) \quad \forall\ i$ and strict inequality at least once.

In some cases there will be a ranking among the objectives. E.g. for Example 1.1, the price might be more important than engine efficiency, this more than horsepower. Then the criterion vectors $(f_1(x), f_2(x), f_3(x))$ are compared lexicographically and one would want to solve

$$\operatorname*{lexmin}_{x \in X}(f(x)) \tag{1.7}$$

Result: $\quad x^* =$ Toyota is the unique optimal solution

When in Example 1.3 the objectives measure some negative impacts of a decision (to be minimized) one might not want to accept a high value of one for a low one of the other.

It is then more useful to minimize the worst of the two, e.g. in Example 1.3

$$\min_{x \geq 0} \max_{i=1,2} f_i(x) \tag{1.8}$$

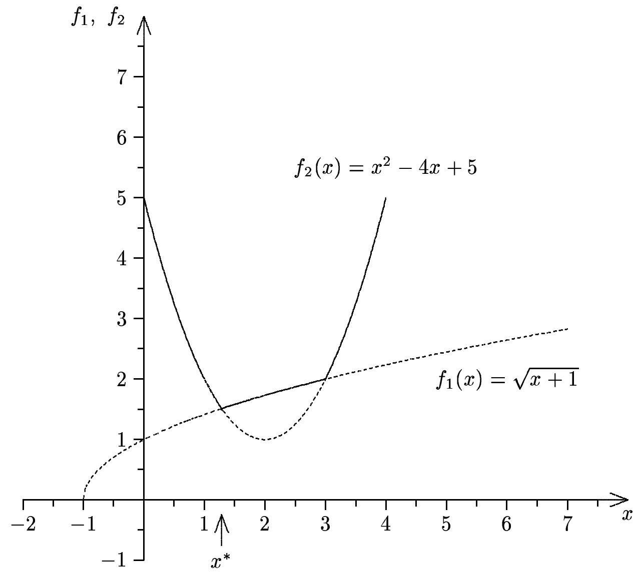


Figure 1.5: Min Max Solution for Example 1.3

Result: $x^* \approx 1.285$.

The meaning of „min" is defined, if we fix an ordering on the objective space. The different possibilities arise from the fact that for $n \geq 2$ there is no relation satisfying the axioms of order on $\mathbb{R}^n$. Therefore weaker definitions of orderings have to be used.

1.4 Orderings and Cones

A **binary relation** $\sim$ on a set A is a subset of $M \times M$.

Definition 1.1. A binary relation $\sim$ on A is called

- **reflexive** if $\forall\, a \in A \quad a \sim a$
- **irreflexive** if $\forall\, a \in A \quad a \not\sim a$
- **symmetric** if $\forall\, a, b \in A \quad a \sim b \implies b \sim a$
- **asymmetric** if $\forall\, a, b \in A \quad a \sim b \implies b \not\sim a$
- **antisymmetric** if $\forall\, a, b \in A \quad a \sim b$ and $b \sim a \implies a = b$
- **transitive** if $\forall\, a, b \in A \quad a \sim b$ and $b \sim c \implies a \sim c$
- **negatively transitive** if $\forall\, a, b \in A \quad a \not\sim b$ and $b \not\sim c \implies a \not\sim c$
- **connected** if $\forall\, a, b \in A\ :\ a \neq b \implies a \sim b$ or $b \sim a$
- **strongly connected** (total) if $\forall\, a, b \in A \quad a \sim b$ or $b \sim a$

Definition 1.2. A binary relation $\sim$ on a set A is

- an **equivalence relation** if it is reflexive, symmetric and transitive.
- a **preorder (quasiorder)** if it is reflexive and transitive; $(A, \preceq)$ is called a **preordered set**.

  Two relations are associated with $\preceq$:
  - $x \prec y \iff x \preceq y$ and $y \not\preceq x$
  - $x \sim y \iff x \preceq y$ and $y \preceq x$

Proposition 1.1. *Let $\preceq$ be a preorder on A. Then $\prec$ is irreflexive and transitive and $\sim$ is an equivalence relation.*

Proof: $\sim$ is reflexive because $\preceq$ is. $\sim$ is symmetric by definition.

Let $x \sim y$ and $y \sim z$.

$$\left.\begin{array}{l} \Longrightarrow\ x \preceq y \preceq z \ \Longrightarrow\ x \preceq z \\ \Longrightarrow\ z \preceq y \preceq x \ \Longrightarrow\ z \preceq x \end{array}\right\} \Longrightarrow\ x \sim z$$

$\prec$ is irreflexive by definition.

Suppose $x \prec y,\ y \prec z \implies x \preceq y \preceq z \implies x \preceq z$. To show that $x \prec z$ suppose $z \preceq x$. But $x \preceq y \implies z \preceq y$ (transitivity) ↯ Contradiction ! $\implies z \not\preceq x \implies x \prec z$.

□

Proposition 1.2. *An asymmetric relation is irreflexive. A transitive, irreflexive relation is asymmetric.*

Proof: Exercise 3.

□

Definition 1.3. A binary relation $\preceq$ on A is

- a total preorder if it is reflexive, transitive and connected
- a total order if it is an antisymmetric total preorder
- a strict weak order if it is asymmetric and negatively transitive

Proposition 1.3. *If $\preceq$ is a total preorder an A, then the associated relation $\prec$ is a strict weak order.*

If $\prec$ is a strict weak order on A, then $\preceq$ defined by

$$x \preceq y \iff \text{either } x \prec y \text{ or } (x \not\prec y \quad \text{and} \quad y \not\prec x)$$

is s total preorder.

Proof: Let $\preceq$ be a total preorder. Then $\prec$ is irreflexive and transitive (Proposition 1.1) and hence asymmetric (Proposition 1.2).

For negative transitivity show $x \not\prec y,\ y \not\prec z \implies x \not\prec z$.

So take $x, y, z \quad x \prec z$ and show $x \prec y$ or $y \prec z$. Suppose $x \not\prec y \implies y \prec x$ or $y \preceq x$ because $\preceq$ is total. In both cases $\implies y \prec z$. ✓

Let $\prec$ be a strict weak order on A. $\preceq$ is reflexive by definition.

For transitivity consider the following cases:

1) $x \prec y,\ y \not\prec z$ and $z \not\prec y$. Then $x \prec z$.

Otherwise $x \not\prec z$ and $z \not\prec y \implies x \not\prec y$ ↯ contradiction ! $\implies x \prec z \implies x \preceq z$

2) $x \not\prec y,\ y \not\prec x$ and $y \prec z$. Then $x \prec z$.

Otherwise $x \not\prec z$ and $y \not\prec x \implies y \not\prec z$ ↯ contradiction ! $\implies x \prec z \implies x \preceq z$

3) $x \not\prec y,\ y \not\prec x,\ y \not\prec z,\ z \not\prec y \implies x \not\prec z$ and $z \not\prec x \implies x \preceq z$

4) $x \prec y$ and $y \prec z$.

If $x \not\prec z$: by saymmetry and $x \prec y \implies y \not\prec x \implies y \not\prec z$ ↯ contradiction ! $\implies$ $x \prec z \implies x \preceq z$.

Connectedness: $x, y \in A \quad x \neq y$. Then $x \prec y$ or $y \prec x$ or ($x \not\prec y$ and $y \not\prec x$) $\implies x \preceq y$ or $y \succeq x$.

□

Definition 1.4. A binary relation $\preceq$ is called

- **partial order** if it is reflexive, transitive and antisymmetric.
- **strict partial order** if it is asymmetric and transitive (irreflexive and transitive).

Some orderings in $\mathbb{R}^n$:

Let $x, y \in \mathbb{R}^n$. We say

$x \leq y$ if $x_i \leq y_i \quad i = 1, \dots, n$	weak componentwise order	(1.9)
$x < y$ if $x_i \leq y_i \quad i = 1, \dots, n \ x \neq y$	componentwise order	(1.10)
$x \ll y$ if $x_i < y_i \quad i = 1, \dots, n$	strict componentwise order	(1.11)

The properties of orderings (especially in $\mathbb{R}^n$ ($\mathbb{R}^2$)) can be interpreted geometrically using cones.

Definition 1.5. A subset $K \subseteq \mathbb{R}^n$ is called **cone**, if

$$\forall\, x \in K \quad \text{and} \quad \forall\, \lambda \in \mathbb{R},\ \lambda > 0 \quad \lambda x \in K. \tag{1.12}$$

Example 1.4. $K = \{x \in \mathbb{R}^2 : x_i \geq 0\}$

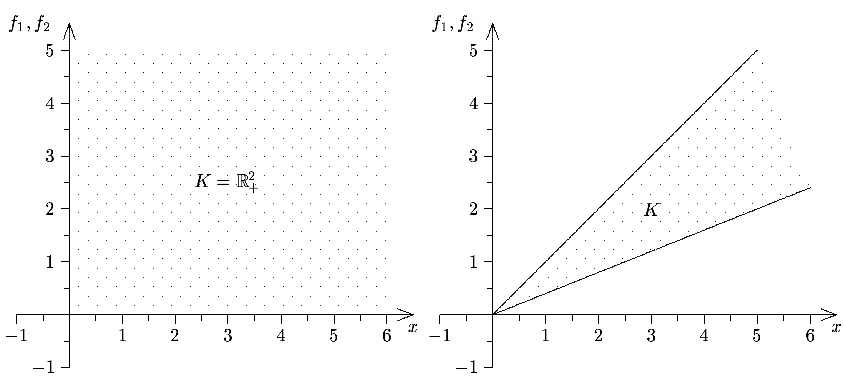


Figure 1.6: Illustration of Two Cones

For any set $M \subset \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ we denote by $\lambda M := \{\lambda x : x \in M\}$, especially $-M = \{-x : x \in M\}$.

Definition 1.6. A cone K is called

- **nontrivial**, if $0 \in K$, $K \neq \{0\}$ and $K \neq \mathbb{R}^n$.
- **convex**, if $\lambda x_1 + (1-\lambda)x_2 \in K \quad \forall\, x_1, x_2 \in K,\ \forall\, 0 < \lambda < 1$.
- **pointed**, if $x \in K \implies -x \notin K$, i.e. $K \cap (-K) \subset \{0\}$.

Remark. If K is a cone then K is convex if $\forall\, x_1, x_2 \in K \quad x_1 + x_2 \in K$. (Note that $\lambda x_1 \in K,\ (1-\lambda)x_2 \in K$, therefore closedness of K under addition is sufficient).

Given a binary relation $\preceq$ on $\mathbb{R}^n$, we can define a set $K_{\preceq} = \{y - x : x \preceq y\}$, loosely speaking the „set of nonnegative elements“.

If $\preceq$ is a relation compatible with scalar multiplication, i.e. for $x \preceq y$ and $\lambda > 0 \quad \lambda x \preceq \lambda y$, we have the following result.

Proposition 1.4. *$K_{\preceq}$ is a cone.*

Proof: Let $u \in K_{\preceq} \implies u = y - x$ for some $x, y \in \mathbb{R}^n \implies x \preceq y \implies \lambda x \preceq \lambda y \implies \lambda(y - x) = \lambda u \in K_{\preceq} \quad \forall\, \lambda > 0$.

□

Example 1.5. Weak componentwise order in $\mathbb{R}^n$.

$x \leq y \iff x_i \leq y_i \quad \forall\, i \iff y_i - x_i \geq 0 \quad \forall\, i \implies K_{\leq} = \{x \in \mathbb{R}^n : x_i \geq 0\} = \mathbb{R}^n_+$.

We know that $\leq$ is a partial order. How do the properties of orderings translate to properties of K ?

Theorem 1.5. *Let $\preceq$ be a relation on $\mathbb{R}^n$ that is compatible with scalar multiplication for $\lambda > 0$. Then*

a) $\preceq$ *is reflexive* $\implies 0 \in K_{\preceq}$

b) $\preceq$ *is transitive* $\implies K_{\preceq}$ *is convex*

c) $\preceq$ *is antisymmetric* $\implies K_{\preceq}$ *is pointed*

Proof:

a) $\preceq$ is reflexive $\implies x \preceq x\ \forall\, x \in \mathbb{R}^n \implies x - x = 0 \in K_{\preceq}$

b) Let $u, v \in K \implies u - 0 \in K_{\preceq},\ 0 - v \in -K_{\preceq} \implies 0 \preceq u,\ -v \preceq 0$.
transitivity $\implies -v \preceq u \implies u - (-v) = u + v \in K \implies K$ convex.

c) Suppose $u = y - x \in K_{\preceq}$ and $-u = x - y \in K_{\preceq},\ u \neq 0 \implies x \preceq y$ and $y \preceq x$ but $x \neq y$
↯ Contradiction !

□

On the other hand, we can use a cone to define an ordering, which is compatible with scalar multiplication.

Let K be a cone. Define $\preceq_K$ by $x \preceq_K y \iff y - x \in K$.

Proposition 1.6. *Let K be a cone. Then $\preceq_K$ is compatible with scalar multiplication and addition in $\mathbb{R}^n$. Furthermore:*

a) $0 \in K \implies \preceq_K$ *is reflexive*

b) K *convex* $\implies \preceq_K$ *is transitive*

c) K *pointed* $\implies \preceq_K$ *is antisymmetric*

Proof: Let $x, y, z \in \mathbb{R}^n$ and $\lambda > 0 \in \mathbb{R}$.

$$x \preceq_K y \implies y - x \in K \implies \lambda(y-x) \in K \implies \lambda x \preceq_K \lambda y$$
$$\implies y - x = (y+z) - (x+z) \in K \implies x + z \preceq_K y + z.$$

a) Let $x \in \mathbb{R}^n \implies x - x = 0 \in K \implies x \preceq_K x$

b) Let $x \preceq_K y,\ y \preceq_K z \implies y - x \in K,\ z - y \in K$ convexity $\implies y - x + z - y = z - x \in K \implies x \preceq_K z$

c) Let $x, y \in \mathbb{R}^n,\ x \preceq_K y,\ y \preceq_K x \implies y - x \in K,\ x - y \in K \implies y - x \in K \cap (-K) = \{0\} \implies y = x$

□

Note. We will only consider orderings which are compatible with scalar multiplication and addition.

1.5 Classification

By the choice of an ordering $\preceq$ on $\mathbb{R}^n$, we can define the meaning of „min“ in

$$„\min_{x \in X}“\ f(x) = „\min_{x \in X}“\ (f_1(x), \dots, f_Q(x)) \tag{1.13}$$

We have seen that objective vectors $y = f(x),\ x \in X$ are not always compared in objective space (i.e. $\mathbb{R}^Q$) directly.
In Example 1.3 we have also considered

$$\min_{x \in X} \max_{i=1,2} f_i(x) \tag{1.14}$$

We have used a mapping $\theta : \mathbb{R}^2 \to \mathbb{R}$ from objective space $\mathbb{R}^2$ to $\mathbb{R}$, where the min in equation (1.14) is actually defined. This mapping is called the **model map**.

The **elements of a multicriteria optimization problem (MCOP)** are:

– the feasible set X (data)

– the objective functions $(f_1, \dots, f_Q)$ (data)

– the objective space $\mathbb{R}^Q$ (data)

– an ordered set $(\mathbb{R}^P, \preceq)$

– a model map θ providing the link between objective space and ordered set

Thus $(X, f, \mathbb{R}^Q)/\theta/(\mathbb{R}^p, \preceq)$ completely describes a multicriteria optimization problem.

Example 1.6.

$$\text{Pareto} - \min_{x \geq 0}(\sqrt{x+1},\ x^2 - 4x + 1) \tag{1.15}$$

Here $X = \{x : x \geq 0\} = \mathbb{R}_+$ is the feasible set

$f = (f_1, f_2) = (\sqrt{x+1},\ x^2 - 4x + 1)$ is the objective function

$\mathbb{R}^Q = \mathbb{R}^2$ is the objective space

$\theta(y) = y$ model map, denoted id, the identity mapping

$(\mathbb{R}^P, \preceq) = (\mathbb{R}^2, <)$ ordered set

Thus (1.15) is

$$(\mathbb{R}_+, f, \mathbb{R}^2)/\mathrm{id}/(\mathbb{R}^2, <) \tag{1.16}$$

Example 1.7. If we have a ranking of objectives as described earlier we compare objective vectors lexicographically.

Let $x, y \in \mathbb{R}^Q$. Then $x <_{\text{lex}} y$ if $\exists\, k,\ 1 \leq k \leq Q$ s.t. $x_i = y_i \quad i = 1, \dots, k-1$ and $x_k < y_k$.

If $X = \{VW, Opel, Ford, Toyota\}$ is the set of alternatives, f_1 is price, f_2 is engine efficiency, f_3 is horsepower, we define $\theta(y) = (y_1, y_2, -y_3)$. Note that higher horsepower is prefered to lower.

The problem is

$$(X, f, \mathbb{R}^3)/\theta/(\mathbb{R}^3, <_{\text{lex}}) \tag{1.17}$$

Definition 1.7. $x^* \in X$ is called an **optimal solution** of an MCOP $(X, f, \mathbb{R}^Q)/\theta/(\mathbb{R}^P, \preceq)$ if there is no $x \neq x^*$ such that

$$\theta(f(x)) \preceq \theta(f(x^*)) \tag{1.18}$$

For an optimal solution x^*, $f(x^*)$ is called an **optimal value** for the MCOP.

Remark.

1) Since we are often dealing with orderings which are not total, a positive definition of optimality, like $\theta(f(x^*)) \preceq \theta(f(x))\ \forall\, x \in X$ is not possible.

2) For special choices of θ and $(\mathbb{R}^P, \preceq)$ specific names for optimal solutions and values are commonly used.

Example 1.8. With the choices $(\mathbb{R}_+, f, \mathbb{R}^2)/\mathrm{id}/(\mathbb{R}^2, <)$ the optimality definition reads: $\nexists\, x \neq x^*$ such that $f(x) < f(x^*)$, i.e. $f_i(x) \leq f_i(x^*)$, and $f(x) \neq f(x^*)$. This is Pareto optimality as introduced before.

Example 1.9. For $(X, f, \mathbb{R}^3)/(y_1, y_2, -y_3)/(\mathbb{R}^3, <_{\text{lex}})$ $x^* \in X$ is an optimal solution if

$$\nexists\, x \in X, x \neq x^* \text{ s.t. } (f_1(x), f_2(x), -f_3(x)) <_{\text{lex}} (f_1(x^*), f_2(x^*), -f_3(x^*)) \quad .$$

We will often speak generally of MCOP in the sense of Pareto or lexicographic optimality, not using any information on problem data.

Definition 1.8. A **multicriteria optimization class (MCO class)** is the set of all MCOP with the same model map and ordered set, and denoted by

$$\bullet/\theta/(\mathbb{R}^P, \preceq). \tag{1.19}$$

So, $\bullet/\mathrm{id}/(\mathbb{R}^Q, <)$ will denote the class of all MCOP, where optimality is understood as Pareto optimality.

1.6 Exercises to Chapter 1

1. Consider the problem

$$\text{„}\min\text{“}\ (f_1(x),\ f_2(x)) \quad \text{subject to} \quad x \in [-1, 1]$$

where

$$f_1(x) = \sqrt{5 - x^2}, \quad f_2(x) = \frac{x}{2}.$$

Illustrate the problem in decision and objective space and determine the Pareto set and the efficient set.

2. Consider the following relations on $\mathbb{R}^n$:

$$\begin{aligned} x \leq y &\iff x_i \leq y_i \quad i = 1, \ldots, n \\ x < y &\iff x_i \leq y_i \quad i = 1, \ldots, n \quad \text{and} \quad x \neq y \\ x \ll y &\iff x_i < y_i \quad i = 1, \ldots, n \quad . \end{aligned}$$

Which of the properties listed in Definition 1.1 do these relations have?

3. Prove the following statements

   a) An asymmetric relation is irreflexive.

   b) A transitive and irreflexive relation is asymmetric.

   c) A negatively transitive and asymmetric relation is transitive.

   d) A transitive, irreflexive and connected relation is negatively transitive.

4. a) Determine the cones related to the (strict, weak) component-wise order, the lexicographic and the max-order on $\mathbb{R}^2$.

   b) Give an example of a non-convex cone and list the properties of the related order.

   c) A cone K is called acute, if there exists an open halfspace $H_a = \{x \in \mathbb{R}^n : \langle x, a \rangle > 0\}$ such that $clK \subset H_a$. Is a pointed cone always acute ? What about a convex cone ?

Chapter 2

Pareto Optimality and Efficiency

Much of the material in this and the following chapter is based on the two books [GN90] and [SNT85].

2.1 Pareto Optimal and Efficient Points

We consider problems of the class $\bullet/\mathrm{id}/(\mathbb{R}^Q,<)$ here:

$$\begin{gathered}\text{Pareto}-\min(f_1(x),\dots,f_Q(x))\\ \text{subject to } x\in X\end{gathered} \tag{2.1}$$

Definition 2.1. A point $x^*\in X$ is called **Pareto optimal**, if there is no $x\in X$ such that $f(x)<f(x^*)$. If x^* is Pareto optimal $f(x^*)$ is called **efficient**. Both x^* and $f(x^*)$ are also called **nondominated**.
If $x^1,x^2\in X$ and $f(x^1)<f(x^2)$ we say x^1 **dominates** x^2 and $f(x^1)$ **dominates** $f(x^2)$.
The **set of all Pareto optimal** $x^*\in X$ is X_{Par}. Let $Y=f(X)$. The **set of all efficient points** $y=f(x^*)\in Y$ is Y_{eff}.
These names are not unique in literature !

For two sets A,B we denote $A+B=\{a+b: a\in A,\ b\in B\}$.

Remark. Equivalent Definitions: x^* is **Pareto optimal** if

1) $\nexists\ x\in X \quad f_i(x)\le f_i(x^*),\ i=1,\dots,Q$ and

$$f_j(x)<f_j(x^*) \text{ for some } j\in\{1,\dots,Q\}$$

2) $\nexists\ x\in X$ s.t. $f(x)-f(x^*)\in -\mathbb{R}^Q_+\setminus\{0\}$

3) $f(x)-f(x^*)\in\mathbb{R}^Q\setminus\{-\mathbb{R}^Q_+\setminus\{0\}\}\ \forall\ x\in X$

4) $f(X)\cap(f(x^*)-\mathbb{R}^Q_+)=\{f(x^*)\}$

5) $\nexists\ f(x)\in f(X)\setminus\{f(x^*)\}$ s.t. $f(x)\in f(x^*)-\mathbb{R}^Q_+$

6) $f(x)\le f(x^*)$ for some $x\in X \implies f(x)=f(x^*)$

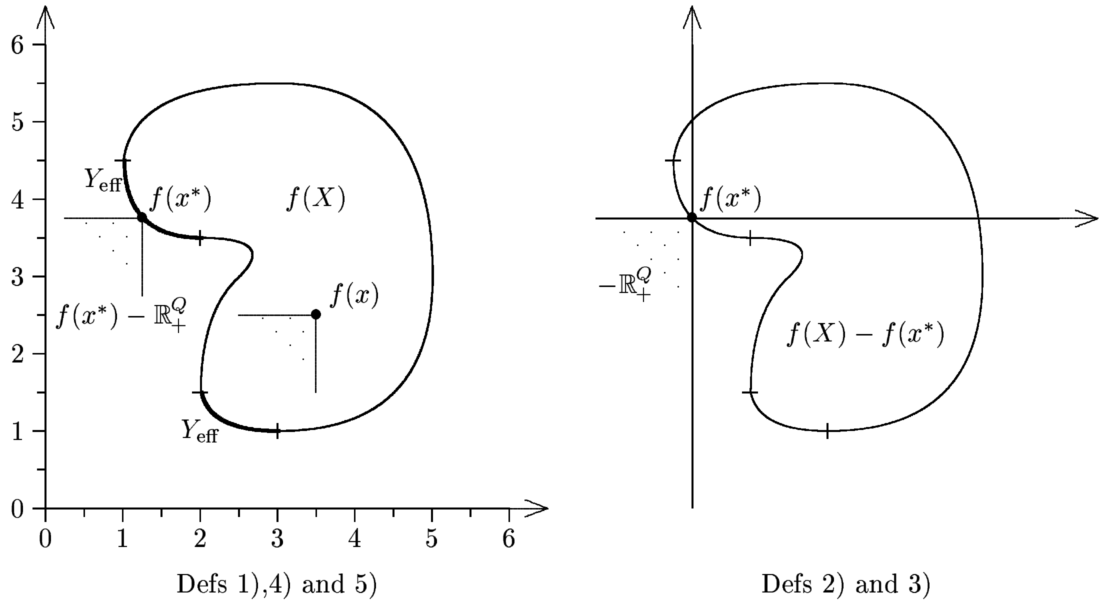


Figure 2.1: Illustration of Definition 2.1

The first questions we discuss are the existence and the properties of the sets X_{Par} and Y_{eff}.

Example 2.1.

$$X = \{\ (x_1, x_2) \in \mathbb{R}^2 : -1 \leq x \leq 1, \quad \begin{array}{lll} -\sqrt{-x_1^2+1} < x_2 \leq 0 & \text{for} & -1 \leq x_1 \leq 0 \\ -\sqrt{-x_1^2+1} \leq x_2 \leq 0 & \text{for} & 0 < x_1 \leq 1\ \} \end{array}$$

$$f(x_1, x_2) = (x_1, x_2) \to \min \tag{2.2}$$

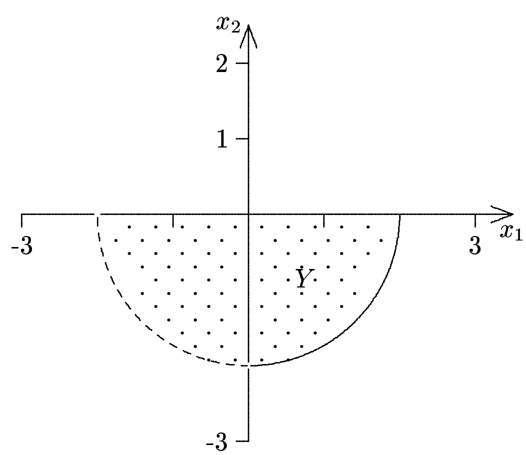


Figure 2.2: Feasible Set of Example 2.1

$Y_{\text{eff}} = \emptyset$, even though f is continuous.

If we take

$$X = \{\ (x_1, x_2) \in \mathbb{R}^2 : -1 \leq x_1 \leq 1, \quad \begin{array}{lll} x_2 = 0 & \text{for} & x_1 = -1 \\ -\sqrt{-x_1^2+1} < x_2 \leq 0 & \text{for} & -1 < x_1 < 0 \\ -\sqrt{-x_1^2+1} \leq x_2 \leq 0 & \text{for} & 0 \leq x_1 \leq 1\ \} \end{array}$$

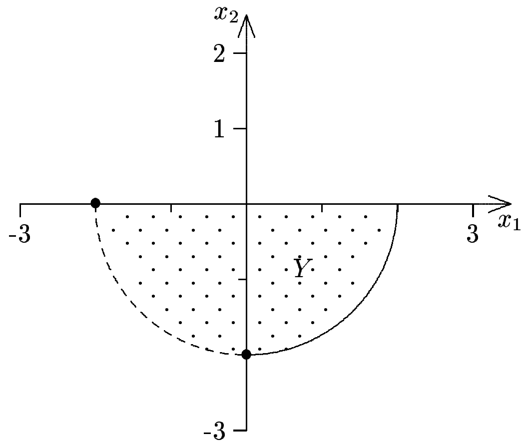


Figure 2.3: Feasible Set of Example 2.1

Now $Y_{\text{eff}} = \left\{ \binom{-1}{0},\ \binom{0}{-1} \right\}$, a disconnected set.

We shall first discuss Y_{eff}. For the following discussion Y may just be a subset of $\mathbb{R}^Q$. For a multicriteria optimization problem $Y = f(X)$.

Let $Y \subset \mathbb{R}^Q$. Let $Y_{\text{eff}} = \{y \in Y : \nexists\, y' \in Y,\ y' < y\}$, in particular $Y_{\text{eff}} \subset Y$.

Proposition 2.1. $Y_{\text{eff}} = (Y + \mathbb{R}^Q_+)_{\text{eff}}$

Proof: Trivial if $Y = \emptyset$.

Assume $Y \neq \emptyset$. First, assume $y \in (Y + \mathbb{R}^Q_+)_{\text{eff}}$, but $y \notin Y_{\text{eff}}$.

If $y \notin Y \implies \exists\, y' \in Y$ and $0 \neq d \in \mathbb{R}^Q_+$ s.t. $y = y' + d$; since $y' \in Y + \mathbb{R}^Q_+ \implies y \notin (Y + \mathbb{R}^Q_+)_{\text{eff}}$ $\lightning$ Contradiction !

If $y \in Y \implies \exists\, y' \in Y_{\text{eff}}$ s.t. $y' < y$, let $d = y - y' (\in \mathbb{R}^Q_+ \setminus \{0\}) \implies y = y' + d \implies y \notin (Y + \mathbb{R}^Q_+)_{\text{eff}}$ $\lightning$ Contradiction !

Hence in either case $y \in Y_{\text{eff}}$.

Second, assume $y \in Y_{\text{eff}}$ but $y \notin (Y + \mathbb{R}^Q_+)_{\text{eff}}$

$\implies \exists\, y' \in Y + \mathbb{R}^Q_+$ with $y - y' = d' \in \mathbb{R}^Q_+ \setminus \{0\}$

$\implies y' = y'' + d''$ with $y'' \in Y,\ d'' \in \mathbb{R}^Q_+$

$\implies y = y' + d' = y'' + \underbrace{(d' + d'')}_{\in \mathbb{R}^Q_+ \setminus \{0\}} = y'' + d$ with $d \neq 0 \implies y \notin Y_{\text{eff}}$ $\lightning$ Contradiction !

Hence $y \in (Y + \mathbb{R}^Q_+)_{\text{eff}}$.

$\square$

Interpretation: We only need to look at the „lower left sector“ of Y.

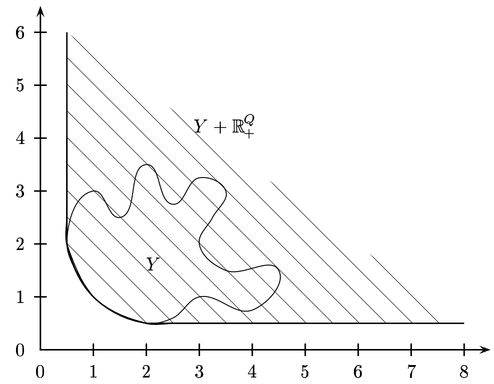


Figure 2.4: „lower left sector“ of Y

Furthermore, efficient points cannot lie everywhere in Y:

Proposition 2.2. $Y_{\text{eff}} \subset \delta Y$ *(boundary of Y)*

Proof: Let $y \in Y_{\text{eff}}$. Suppose $y \notin \delta Y$.

$\Longrightarrow\ y \in \operatorname{int} Y \implies \exists\ \varepsilon -$ neighbourhood $U(y,\varepsilon)$ of y s.t. $U(y,\varepsilon) = y + U(0,\varepsilon) \subset Y$.

Let $d \neq 0,\ d \in \mathbb{R}^Q_+ \implies \exists\ \lambda > 0$ s.t. $\lambda d \in U(0,\varepsilon) \implies y + \lambda d \in Y$ with $\lambda d \in \mathbb{R}^Q_+ \setminus \{0\}$

$\Longrightarrow\ y \notin Y_{\text{eff}}$ ↯ Contradiction !

□

Corollary 2.3. *If Y is open $\implies Y_{\text{eff}} = \emptyset$. If $Y + \mathbb{R}^Q_+$ is open $\implies Y_{\text{eff}} = \emptyset$.*

Some algebra of Y_{eff}:

Proposition 2.4. $(Y_1 + Y_2)_{\text{eff}} \subset Y_{1_{\text{eff}}} + Y_{2_{\text{eff}}}$

Proof: Let $y \in (Y_1 + Y_2)_{\text{eff}}$

$\Longrightarrow\ y = y_1 + y_2$ for some $y_1 \in Y_1,\ y_2 \in Y_2$

suppose $y_1 \notin Y_{1_{\text{eff}}} \implies \exists\ y' \in Y_1,\ d \in \mathbb{R}^Q_+ \setminus \{0\}$ s.t. $y_1 = y' + d \implies y = y' + y_2 + d$ with $y' + y_2 \in Y_1 + Y_2 \implies y \notin (Y_1 + Y_2)_{\text{eff}}$ ↯ Contradiction !

Analogously $y_2 \in Y_{2_{\text{eff}}} \implies y_1 + y_2 \in Y_{1_{\text{eff}}} + Y_{2_{\text{eff}}}$

□

Proposition 2.5. $(\alpha \cdot Y)_{\text{eff}} = \alpha \cdot Y_{\text{eff}}$, *where* $\alpha \in \mathbb{R},\ \alpha > 0$.

Proof: Exercise 8.

□

In order to prove an existence result for efficient points we need Zorn's Lemma.

Definition 2.2. A set M is **inductively ordered**, if every totally ordered subset of M (a chain) has a lower bound. The ordering on M is reflexive and transitive.

Lemma 2.6 (Zorn's Lemma). *Let M be a set on which a reflexive, transitive relation $\preceq$ is given, and such that M is inductively ordered, then M contains a minimal element $\overline{m}$, i.e.*

$$m \in M, \qquad m \preceq \overline{m} \quad \Longrightarrow \quad \overline{m} \preceq m \tag{2.3}$$

Theorem 2.7. *Suppose $Y \neq \emptyset$ and $\exists\, y^0 \in Y$ s.t. the section $Y^0 = \{y \in Y : y \leq y^0\} = (y^0 - \mathbb{R}^Q_+) \cap Y$ is compact. Then $Y_{\text{eff}} \neq \emptyset$.*

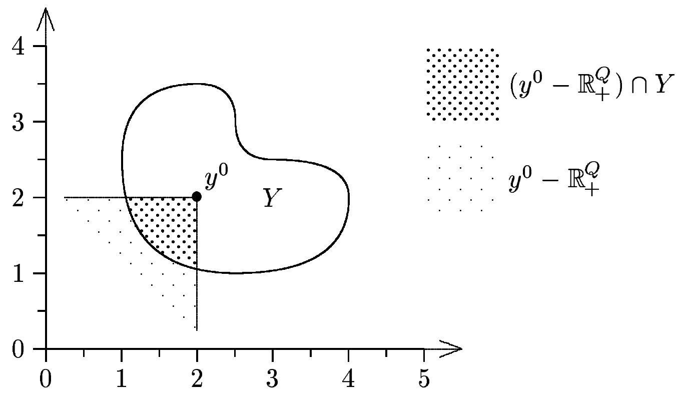


Figure 2.5: $(y^0 - \mathbb{R}^Q_+) \cap Y$, $y^0 - \mathbb{R}^Q_+$

Proof: Let Y^0 be a „compact section". Let $\{y^\alpha,\ \alpha \in A\}$ be a chain in Y^0.

(A **chain** is a totally ordered subset of a partially ordered set.)

We prove that $\{y^\alpha\}$ has a lower bound.

Let $B := \{a \subset A : |a| < \infty\}$. Suppose $a \in B \implies y^a = \inf\{y^\alpha : \alpha \in a\}$ exists and $y^a \in Y^0$ because $\{y^\alpha\}$ is a chain and a is finite.

Consider all sets $Y_\alpha := (y^\alpha - \mathbb{R}^Q_+) \cap Y$. Obviously $Y_\alpha \subset Y^0$ and Y_α are compact ($\mathbb{R}^Q_+$ is closed). Furthermore, if $a \in B$, i.e. finite $\bigcap_{\alpha \in a} Y_\alpha \neq \emptyset$ because it contains y^a.

By compactness of Y^0 it follows $\bigcap_{\alpha \in A} Y_\alpha \neq \emptyset$

$\implies \exists\, y' \in \bigcap_{\alpha \in A} (y^\alpha - \mathbb{R}^Q_+) \cap Y^0$

$\implies y' \leq y^\alpha \quad \forall\, \alpha \in A$

$\implies y' \in Y^0$ is a lower bound of $\{y^\alpha : \alpha \in A\}$, which is therefore inductively ordered.

Hence Y^0 contains a minimal element y^*. We show that $y^* \in Y_{\text{eff}}$.

Otherwise there would exist $\overline{y} \in Y$, $\overline{y} \neq y^*$

$\overline{y} \in (y^* - \mathbb{R}^Q_+) \cap Y \subset (y^0 - \mathbb{R}^Q_+ - \mathbb{R}^Q_+) \cap Y = (y^0 - \mathbb{R}^Q_+) \cap Y - \mathbb{R}^Q_+$

contradicting minimality of y^* for Y^0.

$\square$

Another existence result does not use a compact section but a condition on Y which is similar to the finite subcover property of compact sets.

Definition 2.3. $Y \subset \mathbb{R}^Q$ is called $\mathbb{R}^Q_+$**-semicompact** if every open cover of Y of the form $\{(y^\alpha - \mathbb{R}^Q_+)^c : y^\alpha \in Y,\ \alpha \in A\}$ has a finite subcover. This means: $Y \subset \bigcup_{\alpha \in A} (y^\alpha - \mathbb{R}^Q_+)^c \implies$
$\exists\ m \in \mathbb{N}\quad \alpha_1, \ldots, \alpha_m$ s.t. $Y \subset \bigcup_{i=1}^{m} (y^{\alpha_i} - \mathbb{R}^Q_+)$.

Note that $(y^\alpha - \mathbb{R}^Q_+)^c$ is open, and the complement of $y^\alpha - \mathbb{R}^Q_+$.

Theorem 2.8. *If $Y \neq \emptyset$ is $\mathbb{R}^Q_+$-semicompact then $Y_{\text{eff}} \neq \emptyset$.*

Proof: We show that Y is inductively ordered and apply Zorn's Lemma.

Assume Y is not inductively ordered

$\implies \exists$ totally ordered subset (chain) of Y, $\overline{Y} = \{y^\alpha : \alpha \in A\}$ which has no lower bound.

$\implies \bigcap_{\alpha \in A} ((y^\alpha - \mathbb{R}^Q_+) \cap Y) = \emptyset$

(As in the proof of Theorem 2.7, any element in this intersection would be a lower bound of $\overline{Y}$.)

$\implies \forall\ y \in Y \quad \exists\ y^\alpha \in \overline{Y}$ s.t. $y \notin y^\alpha - \mathbb{R}^Q_+$

Since $y^\alpha - \mathbb{R}^Q_+$ is closed $\implies \{(y^\alpha - \mathbb{R}^Q_+)^c : \alpha \in A\}$ is an open cover of Y.

Also: $y^\alpha - \mathbb{R}^Q_+ \subset y^{\alpha'} - \mathbb{R}^Q_+ \iff y^\alpha \leq y^{\alpha'}$

$\implies$ The sets of the cover are totally ordered by inclusion because $\overline{Y}$ is a chain. (1)

Y is $\mathbb{R}^Q_+$-semicompact $\implies \exists$ finite subcover of $\{(y^\alpha - \mathbb{R}^Q_+)^c : \alpha \in A\}$. (2)

(1), (2) $\implies \exists$ single $y^{\overline{\alpha}} \in \overline{Y}$ such that $Y \subset (y^{\overline{\alpha}} - \mathbb{R}^Q_+)^c$

This implies $y^{\overline{\alpha}} \leq y^\alpha \quad \forall\ \alpha \in A \implies y^{\overline{\alpha}} \notin Y$ ↯ Contradiction !

$\implies Y$ is inductively ordered. As in the proof of Theorem 2.7 we conclude $Y_{\text{eff}} \neq \emptyset$.

□

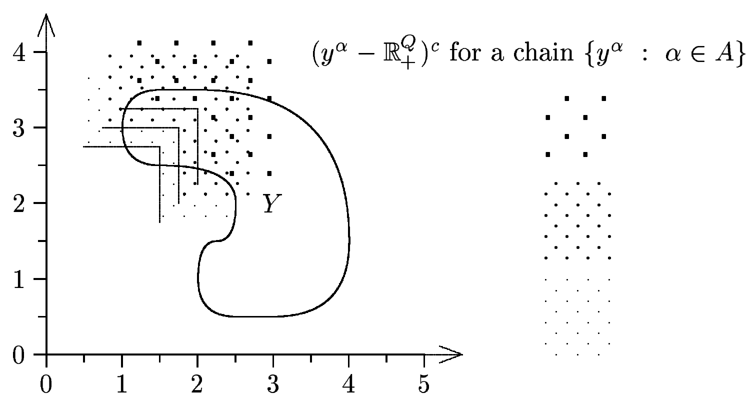


Figure 2.6: $(y^\alpha - \mathbb{R}^Q_+)^c$

It is usually not easy to check $\mathbb{R}^Q_+$-semicompactness. A weaker result is obtained if we use the stronger assumption of $\mathbb{R}^Q_+$-compactness.

Definition 2.4. $Y \subset \mathbb{R}^Q$ is called $\mathbb{R}^Q_+$**-compact**, if $\forall\ y \in Y \quad (y - \mathbb{R}^Q_+) \cap Y$ is compact.

Proposition 2.9. *If Y is $\mathbb{R}^Q_+$-compact then Y is $\mathbb{R}^Q_+$-semicompact.*

Proof: Let $\{(y^\alpha - \mathbb{R}^Q_+)^c : y^\alpha \in Y,\ \alpha \in A\}$ be an open cover of Y. For arbitrary $y^{\alpha'} \in Y$ take

$$\{(y^\alpha - \mathbb{R}^Q_+)^c : y^\alpha \in Y,\ \alpha \in A,\ \alpha \neq \alpha'\} \tag{2.4}$$

This is an open cover of $(y^{\alpha'} - \mathbb{R}^Q_+) \cap Y$, a compact set (by definition).
$\Longrightarrow$ (2.4) must contain a finite subcover of $(y^{\alpha'} - \mathbb{R}^Q_+) \cap Y$, together with $(y^{\alpha'} - \mathbb{R}^Q_+)^c$ we have a finite cover of Y.

□

Corollary 2.10. *If $Y \subset \mathbb{R}^Q$ is nonempty and $\mathbb{R}^Q_+$-compact, then $Y_{\text{eff}} \neq \emptyset$.*

Proof: Theorem 2.8 and Proposition 2.9.

□

Note. The condition of $\mathbb{R}^Q_+$-compactness can be replaced by $\mathbb{R}^Q_+$-closedness and $\mathbb{R}^Q_+$-boundedness, which are generalizations of closedness and boundedness. For closed convex sets Y it can be shown that the conditions of Theorem 2.7, Corollary 2.10 and $\mathbb{R}^Q_+$-closedness and $\mathbb{R}^Q_+$-boundedness coincide.

We now consider existence of X_{Par}.

Definition 2.5. A function $f : \mathbb{R}^n \to \mathbb{R}^Q$ is said to be $\mathbb{R}^Q_+$**-semicontinuous** if

$$f^{-1}(y - \mathbb{R}^Q_+) = \{x \in \mathbb{R}^n : y - f(x) \in \mathbb{R}^Q_+\} \tag{2.5}$$

is closed for all $y \in \mathbb{R}^Q$.
(The preimage of translated negative orthants is closed).

Lemma 2.11. *$f : \mathbb{R}^n \to \mathbb{R}^Q$ is $\mathbb{R}^Q_+$-semicontinuous if and only if $f_i : \mathbb{R}^n \to \mathbb{R}$ are lower semicontinuous $\forall\ i = 1, \dots, Q$.*

Proof: Exercise 9.

□

Proposition 2.12. *Let $X \subset \mathbb{R}^n$ be nonempty and compact, $f : \mathbb{R}^n \to \mathbb{R}^Q$ be $\mathbb{R}^Q_+$-semicontinuous. Then $Y = f(X)$ is $\mathbb{R}^Q_+$-semicompact.*

Proof: Let $\{(y^\alpha - \mathbb{R}^Q_+)^c : y^\alpha \in Y,\ \alpha \in A\}$ be an open cover of Y. By $\mathbb{R}^Q_+$-semicontinuity of $f \implies \{f^{-1}((y^\alpha - \mathbb{R}^Q_+)^c) : y^\alpha \in Y,\ \alpha \in A\}$ is an open cover of X.
X is compact. $\implies \exists$ finite subcover of X
$\implies$ The image of this subcover is a finite subcover of $Y \implies Y$ is $\mathbb{R}^Q_+$ semicompact.

□

Theorem 2.13. *Let $X \subset \mathbb{R}^Q$ be nonempty, compact. Let f be $\mathbb{R}^Q_+$-semicontinuous. Then $X_{\text{Par}} \neq \emptyset$.*

Proof: Theorem 2.8, Proposition 2.12.

□

Remark. All results presented here are still valid, if $\mathbb{R}^Q_+$ is replaced by a convex, pointed, nontrivial, closed cone K. Closedness is not required if $(y - \operatorname{cl} K)$ is used instead of $(y - K)$ everywhere.

2.2 Weak and Strict Pareto Optimal Points

Definition 2.6. A point $x^* \in X$ is called **weakly Pareto optimal** if there is no $x \in X$ such that $f(x) \ll f(x^*)$, i.e. $f_i(x) < f_i(x^*)\ \forall\, i$. $y^* = f(x^*)$ is called **weakly efficient**.
A point $x^* \in X$ is called **strictly Pareto optimal** if there is no $x \in X,\ x \neq x^*$ such that $f(x) \leq f(x^*)$. $y^* = f(x^*)$ is called **strictly efficient**.
The weak (strict) Pareto optimal and efficient sets are denoted $X_{\text{w-Par}}$ ($X_{\text{s-Par}}$) and $Y_{\text{w-eff}}$ ($Y_{\text{s-eff}}$), respectively.

Remark.

- $Y_{\text{s-eff}} \subset Y_{\text{eff}} \subset Y_{\text{w-eff}}, \quad X_{\text{s-Par}} \subset X_{\text{Par}} \subset X_{\text{w-Par}}$
- Equivalent definitions:
  $$\begin{aligned} y^* \in Y_{\text{w-eff}} &\iff \nexists\, y \in Y : y^* - y \in \operatorname{int} \mathbb{R}^Q_+ \\ &\iff (y^* - \operatorname{int} \mathbb{R}^Q_+) \cap Y = \emptyset \end{aligned}$$
- $x^* \in X_{\text{s-Par}} \iff x^* \in X_{\text{Par}}$ and $|\{x : f(x) = f(x^*)\}| = 1$
  In fact, strictly efficient points can only be defined in the context of multicriteria optimization problems, not for Y alone.

Theorem 2.14 (Existence of weakly efficient points).
Let $\emptyset \neq Y \subset \mathbb{R}^Q$ be compact. Then $Y_{\text{w-eff}} \neq \emptyset$.

Proof: Suppose $Y_{\text{w-eff}} = \emptyset$. $\implies \forall\, y \in Y\ \exists\, y' \in Y$ s.t. $y \in y' + \operatorname{int} \mathbb{R}^Q_+$
$\implies Y \subset \bigcup_{y' \in M} (y' + \operatorname{int} \mathbb{R}^Q_+)$
Therefore we have an open cover of Y.
By compactness $\implies \exists$ finite subcover

$$Y \subset \bigcup_{i=1}^{k} (y^i + \operatorname{int} \mathbb{R}^Q_+) \tag{2.6}$$

$\implies \forall\, i = 1, \ldots, k \quad \exists\, 1 \leq j \leq k \quad y^i \in y^j + \operatorname{int} \mathbb{R}^Q_+$.
In other words $\forall\, i \quad \exists\, j : y^j \ll y^i$.
By transitivity $\implies \exists\, i^*$ and a chain of inequalities s.t. $y^{i^*} \ll y^{i_1} \ll \ldots \ll y^{i_m} \ll y^{i^*}$ ↯
Contradiction !

□

Remark. Note that Zorn's Lemma was not needed here. Compactness is enough ! The important difference is that in Theorems 2.7 and 2.8 we deal with sets $y - \mathbb{R}^Q_+$ which are closed. Here we have sets $y - \operatorname{int} \mathbb{R}^Q_+$ which are open. (Note that $y \notin y - \operatorname{int} \mathbb{R}^Q_+$).

Corollary 2.15. *Let $X \subset \mathbb{R}^n$ be compact and $f : \mathbb{R}^n \to \mathbb{R}^Q$ continuous, then $X_{\text{w-Par}} \neq \emptyset$.*

Proof: Follows from Theorem 2.13 and $X_{\text{Par}} \subset X_{\text{w-Par}}$ or from Theorem 2.14 and the fact that $f(X)$ is compact for compact X.

□

The inclusion $Y_{\text{eff}} \subset Y_{\text{w-eff}}$ is strict, in general the latter may be nonempty even if Y is not compact.

Example 2.2. $Y = \{(y_1, y_2) \in \mathbb{R}^2 \ : \ 0 < y_1 < 1,\ 0 \leq y_2 \leq 1\}$. Then $Y_{\text{eff}} = \emptyset$, $\quad Y_{\text{w-eff}} = (0,1) \times \{0\}$.

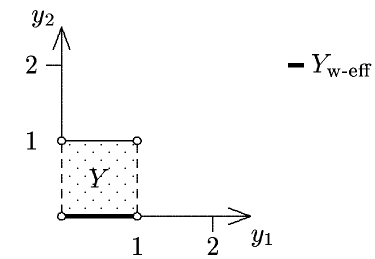


Figure 2.7: Empty Efficient Set

If we close the square $\overline{Y} = \{(y_1, y_2) \in \mathbb{R}^2 \ : \ 0 \leq y_i \leq 1\}$ we get $Y_{\text{eff}} = \{0\}$, $\quad Y_{\text{w-eff}} = \{(y_1, y_2) : y_1 = 0 \ \text{ or } \ y_2 = 0\}$.

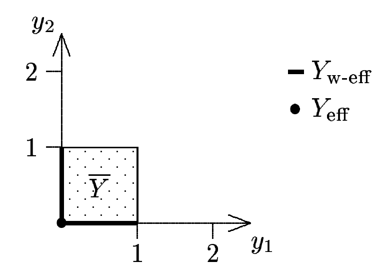


Figure 2.8: Efficient and Weakly Efficient Points

$X_{\text{Par}}, X_{\text{s-Par}}$ and $X_{\text{w-Par}}$ can be characterized geometrically. To do that we need **level sets** and **level curves** of functions.

Definition 2.7. Let $f : X \to \mathbb{R}$, $X \subset \mathbb{R}^n$ and $\overline{x} \in X$. Then

$$L_{\leq} f((\overline{x})) = \{x \in X \ : \ f(x) \leq f(\overline{x})\} \tag{2.7}$$

is called the **level set** of $\overline{x}$ for f.

$$L_{=} f((\overline{x})) = \{x \in X \ : \ f(x) = f(\overline{x})\} \tag{2.8}$$

is called the **level curve** of $\overline{x}$ for f.

Example 2.3. $f(x_1, x_2) = x_1^2 + x_2^2$. Let $\overline{x} = (3,4) \implies$
$L_{\leq} f((\overline{x})) = \{(x_1, x_2) \ : \ x_1^2 + x_2^2 \leq 25\}$, $\quad L_{=} f((\overline{x})) = \{(x_1, x_2) \ : \ x_1^2 + x_2^2 = 25\}$

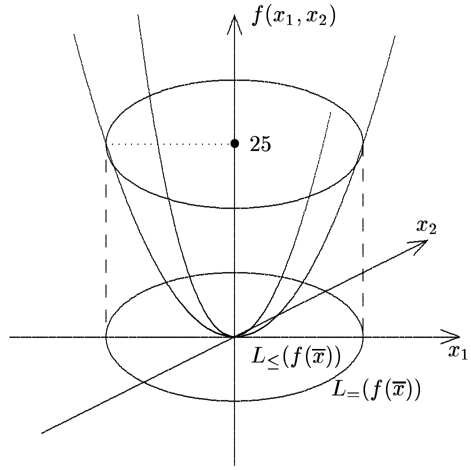


Figure 2.9: Level Set and Level Curve in Example 2.3

For a multicriteria problem we consider the level sets / curves of $\overline{x}$ for all $f_1, \ldots, f_Q$. Obviously $L_{=}(f_i(\overline{x})) \subset L_{\leq}(f_i(\overline{x}))$ and $x \in L_{=}(f_i(\overline{x}))$ $\forall\, i = 1, \ldots, Q$.

Strict level sets are $L_{<}(f(\overline{x})) = L_{\leq}(f(\overline{x})) \setminus L_{=}(f(\overline{x}))$.

Consider the situation:

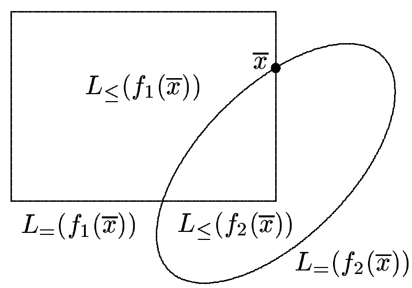


Figure 2.10: Level Sets

Can $\overline{x}$ be Pareto optimal ?

No ! We can move into the intersection of both level sets and find points which are better with respect to both f_1 and f_2. $\overline{x}$ is not even weakly Pareto optimal.

Formally we can state: [EHK+97]

Theorem 2.16. *Let $x^* \in X$, $y_q := f_q(x^*)$ then*

a) *x^* is* **strictly Pareto optimal** *if and only if*

$$\bigcap_{q=1}^{Q} L_{\leq}(y_q) = \{x^*\} \tag{2.9}$$

b) *x^* is* **Pareto optimal** *if and only if*

$$\bigcap_{q=1}^{Q} L_{\leq}(y_q) = \bigcap_{q=1}^{Q} L_{=}(y_q) \tag{2.10}$$

c) *x^* is* **weakly Pareto optimal** *if and only if*

$$\bigcap_{q=1}^{Q} L_{<}(y_q) = \emptyset \tag{2.11}$$

Proof:

a) x^* is strictly Pareto optimal

$$\iff \nexists\, x \in X,\ x \neq x^* \ \text{ s.t. } \ f(x) \leq f(x^*)$$
$$\iff \nexists\, x \in X,\ x \neq x^* \ \text{ s.t. } \ f_q(x) \leq f_q(x^*) \quad \forall\, q = 1, \ldots, Q$$
$$\iff \nexists\, x \in X,\ x \neq x^* \ \text{ s.t. } \ x \in \bigcap_{q=1}^{Q} L_{\leq}(y_q)$$
$$\iff \bigcap_{q=1}^{Q} L_{\leq}(y_q) = \{x^*\}$$

b) x^* is Pareto optimal

$$\iff \nexists\, x \in X, \ \text{ s.t. } \ (f_q(x) \leq f_q(x^*) \quad \forall\, q = 1, \ldots, Q \ \text{ and } \ f_j(x) < f_j(x^*) \ \text{ for some } \ j)$$
$$\iff \nexists\, x \in X, \ \text{ s.t. } \ (x \in \bigcap_{q=1}^{Q} L_{\leq}(y_q) \ \text{ and } \ \exists\, j : x \in L_{<}(y_j))$$
$$\iff \bigcap_{q=1}^{Q} L_{\leq}(y_q) = \bigcap_{q=1}^{Q} L_{=}(y_q)$$

c) x^* is weakly Pareto optimal

$$\iff \nexists\, x \in X : f_q(x) < f_q(x^*) \quad \forall\, q = 1, \ldots, Q$$
$$\iff \nexists\, x \in X : x \in \bigcap_{q=1}^{Q} L_{<}(y_q)$$
$$\iff \bigcap_{q=1}^{Q} L_{<}(y_q) = \emptyset.$$

□

Example 2.4. Consider the points $x^1 = \binom{1}{1}$, $x^2 = \binom{1}{4}$, $x^3 = \binom{4}{4}$. Find a point $x^* \in \mathbb{R}^2$ such that the sum of weighted squared distances of x^* to x^i is minimal.

Two weights for each x^i are given:

$$w_1^1 = 1, \quad w_1^2 = 1, \quad w_1^3 = 1$$
$$w_2^1 = 2, \quad w_2^2 = 1, \quad w_2^3 = 4$$

$$f_i(x) = \sum_{j=1}^{3} w_i^j((x_1^j - x_1)^2 + (x_2^j - x_2)^2) \tag{2.12}$$
$$\begin{aligned} f_1(x) &= (1-x_1)^2 + (1-x_1)^2 + (4-x_1)^2 + (1-x_2)^2 + (4-x_2)^2 + (4-x_2)^2 \\ &= 2\cdot(1-x_1)^2 + (4-x_1)^2 + (1-x_2)^2 + 2\cdot(4-x_2)^2 \\ &= 2\cdot(1-2x_1+x_1^2) + (16-8x_1+x_1^2) + (1-2x_2+x_2^2) + 2\cdot(16-8x_2+x_2^2) \\ &= 3\cdot(x_1^2 - 4x_1 + x_2^2 - 6x_2) + 51 \end{aligned}$$

$$\begin{aligned} f_2(x) &= 2\cdot((1-x_1)^2 + (1-x_2)^2) + 1\cdot((1-x_1)^2 + (4-x_2)^2) + 4\cdot((4-x_1)^2 + (4-x_2)^2) \\ &= 3\cdot(1-x_1)^2 + 4\cdot(4-x_1)^2 + 2\cdot(1-x_2)^2 + 5\cdot(4-x_2)^2 \\ &= 3\cdot(1-2x_1+x_1^2) + 4\cdot(16-8x_1+x_1^2) + 2\cdot(1-2x_2+x_2^2) + 5\cdot(16-8x_2+x_2^2) \\ &= 7\cdot(x_1^2 - \frac{38}{7}x_1 + x_2^2 - \frac{44}{7}x_2) + 149 \end{aligned}$$

We want to know if $x = (2, 2)$ is Pareto optimal. So we check the level sets and level curves:

$$\begin{aligned} f_1(2,2) &= 3\cdot(x_1^2 - 4x_1 + x_2^2 - 6x_2) + 51 = 3\cdot(4 - 8 + 4 - 12) + 51 = 15 \\ f_2(2,2) &= 7\cdot(x_1^2 - \frac{38}{7}x_1 + x_2^2 - \frac{44}{7}x_2) + 149 = 7\cdot(4 - \frac{76}{7} + 4 - \frac{88}{7}) + 149 = 41 \end{aligned}$$

$$L_=(f_1(2,2)) \quad = \quad \{x \in \mathbb{R}^2 \ : \ f_1(x) = 15\}$$

$$\begin{aligned} f_1(x) = 15 \quad &\Longleftrightarrow \quad 3 \cdot (x_1^2 - 4x_1 + x_2^2 - 6x_2) + 51 = 15 \\ &\Longleftrightarrow \quad (x_1^2 - 4x_1 + x_2^2 - 6x_2) + 17 = 5 \\ &\Longleftrightarrow \quad (x_1 - 2)^2 + (x_2 - 3)^2 + 4 = 5 \\ &\Longleftrightarrow \quad (x_1 - 2)^2 + (x_2 - 3)^2 = 1 \end{aligned}$$

$$\Longrightarrow \ L_=(f_1(2,2)) = \{x \in \mathbb{R}^2 \ : \ (x_1 - 2)^2 + (x_2 - 3)^2 = 1\}$$

$$\begin{aligned} f_2(x) = 41 \quad &\Longleftrightarrow \quad 7 \cdot (x_1^2 - \frac{38}{7}x_1 + x_2^2 - \frac{44}{7}x_2) + 149 = 41 \\ &\Longleftrightarrow \quad (x_1 - \frac{19}{7})^2 + (x_2 - \frac{22}{7})^2 = \frac{89}{49} \end{aligned}$$

$$\Longrightarrow \ L_=(f_2(2,2)) = \left\{x \in \mathbb{R}^2 \ : \ (x_1 - \frac{19}{7})^2 + (x_2 - \frac{22}{7})^2 = \frac{89}{49}\right\}$$

This is a circle around $(\frac{19}{7}, \frac{22}{7})$ with radius $\frac{\sqrt{89}}{7}$.

In Figure 2.11 we see that $\bigcap_{i=1}^{2} L_\leq(f_i(2,2)) \neq \bigcap_{i=1}^{2} L_=(f_i(2,2))$ because the disks intersect in a region.

Let us check $(2,3)$:

$$\begin{aligned} f_1(2,3) \quad &= \quad 3 \cdot (x_1^2 - 4x_1 + x_2^2 - 6x_2) + 51 = 3 \cdot (4 - 8 + 9 - 18) + 51 = 12 \\ f_2(2,3) \quad &= \quad 7 \cdot (x_1^2 - \frac{38}{7}x_1 + x_2^2 - \frac{44}{7}x_2) + 149 = 7 \cdot (4 - \frac{76}{7} + 9 - \frac{132}{7}) + 149 = 32 \end{aligned}$$

$$L_=(f_1(2,3)) \quad = \quad \{x \in \mathbb{R}^2 \ : \ f_1(x) = 12\}$$

$$\begin{aligned} f_1(x) = 12 \quad &\Longleftrightarrow \quad 3 \cdot (x_1^2 - 4x_1 + x_2^2 - 6x_2) + 51 = 12 \\ &\Longleftrightarrow \quad (x_1^2 - 4x_1 + x_2^2 - 6x_2) + 17 = 4 \\ &\Longleftrightarrow \quad (x_1 - 2)^2 + (x_2 - 3)^2 + 4 = 4 \\ &\Longleftrightarrow \quad (x_1 - 2)^2 + (x_2 - 3)^2 = 0 \end{aligned}$$

$$\Longrightarrow \ L_=(f_1(2,3)) = \{x \in \mathbb{R}^2 \ : \ (x_1 - 2)^2 + (x_2 - 3)^2 = 0\}$$

$$\begin{aligned} f_2(x) = 32 \quad &\Longleftrightarrow \quad 7 \cdot (x_1^2 - \frac{38}{7}x_1 + x_2^2 - \frac{44}{7}x_2) + 149 = 32 \\ &\Longleftrightarrow \quad (x_1 - \frac{19}{7})^2 + (x_2 - \frac{22}{7})^2 = \frac{26}{49} \end{aligned}$$

$$\Longrightarrow \ L_=(f_2(2,3)) = \left\{x \in \mathbb{R}^2 \ : \ (x_1 - \frac{19}{7})^2 + (x_2 - \frac{22}{7})^2 = \frac{26}{49}\right\}$$

This is a circle around $(\frac{19}{7}, \frac{22}{7})$ with radius $\frac{\sqrt{26}}{7}$.

We have to check if $L_=(f_1(2,3)) \cap L_=(f_2(2,3)) = L_\leq(f_1(2,3)) \cap L_\leq(f_2(2,3))$.

But for $x = (2,3)$ $\ L_=(f_1(2,3)) = \{(2,3)\}$. Then $L_\leq(f_1(2,3))$ is only one point. The radius of $L_\leq(f_2(2,3))$ is $\frac{\sqrt{26}}{7}$. Thus $\binom{2}{2}$ is not Pareto optimal, $\binom{2}{3}$ is Pareto optimal.

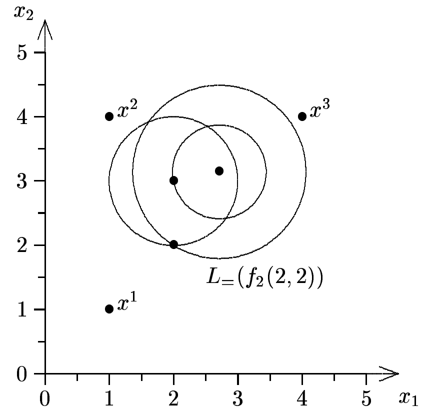


Figure 2.11: Location Problem of Example 2.4

$$L_=(f_1(2,2)) = L_=(15) \longrightarrow \text{circle around } (2,3) \text{ with radius } 1$$
$$L_=(f_1(2,3)) = L_=(12) \longrightarrow \text{circle around } (2,3) \text{ with radius } 0$$
$$L_=(f_2(2,2)) = L_=(41) \longrightarrow \text{circle around } (\frac{19}{7},\frac{22}{7}) \text{ with radius } \frac{\sqrt{89}}{7}$$
$$L_=(f_2(2,3)) = L_=(32) \longrightarrow \text{circle around } (\frac{19}{7},\frac{22}{7}) \text{ with radius } \frac{\sqrt{26}}{7}$$

Theorem 2.16 shows that sometimes not all the criteria are needed to see if a point x is weakly or strictly Pareto optimal: once $\bigcap_{i=1}^{j<Q} L_{\leq}(f_i(x^*))$ is empty it will remain so, if intersected with more level sets.

Let $P \subset \{1,\dots,Q\}$ and denote by f^P the objective function vector that only contains criteria $f_j,\ j \in P$.

Corollary 2.17. *Let $P \subset \{1,\dots,Q\}$. Then*

a) If x is weakly Pareto optimal for $(X, f^P, \mathbb{R}^{|P|})/\text{id}/(\mathbb{R}^{|P|}, \ll)$ it is also weakly Pareto optimal for $(X, f, \mathbb{R}^Q)/\text{id}/(\mathbb{R}^Q, \ll)$.

b) If x is strictly Pareto optimal for $(X, f^P, \mathbb{R}^{|P|})/\text{id}/(\mathbb{R}^{|P|}, \leq)$ it is also strictly Pareto optimal for $(X, f, \mathbb{R}^Q)/\text{id}/(\mathbb{R}^Q, \leq)$.

Stronger results can be obtained for convex functions. So suppose that $X \subset \mathbb{R}^n$ is convex and that $f_i : \mathbb{R}^n \to \mathbb{R}$ are convex. This implies that all level sets are convex. Convex analysis has an important theorem concerning the intersection of convex sets: [Hel23]

Theorem 2.18 (Helly, 1923). *Let $C_1,\dots,C_Q \subset \mathbb{R}^n$ be convex sets $(Q > n)$. Then $\bigcap_{i=1}^{Q} C_i \neq \emptyset$ if and only if for all collections of $n+1$ sets $C_{i_1},\dots,C_{i_{n+1}}$ holds $\bigcap_{i=1}^{n+1} C_{i_j} \neq \emptyset$.*

In other words: $\bigcap_{i=1}^{Q} C_i = \emptyset$ if and only if $\exists\ \{i_1,\dots,i_{n+1}\} \subset \{1,\dots,Q\}$ s.t. $\bigcap_{j=1}^{n+1} C_{i_j} = \emptyset$.
Putting this result and Corollary 2.17 together we get, if we take as $C_i = L_{\leq}(f_i(x))$:

Proposition 2.19. *Consider the problem $(X, f, \mathbb{R}^Q)/\text{id}/(\mathbb{R}^Q, \ll)$, where $X \subset \mathbb{R}^n$ is convex, $f_i : \mathbb{R}^n \to \mathbb{R}$ are convex and $Q > n$. Then $x^* \in X$ is weakly Pareto optimal if and only if $\exists\ P \subset \{1,\dots,Q\},\ 0 < |P| \leq n+1$ such that x^* is weakly Pareto optimal for $(X, f^P, \mathbb{R}^{|P|})/\text{id}/(\mathbb{R}^{|P|}, \ll)$.*

In other words: $X_{\text{w-Par}}(f) = \bigcup_{\substack{P\subset\{1,\dots,Q\} \\ |P|\le n+1}} X_{\text{w-Par}}(f^P)$.

It is even possible to describe $X_{\text{w-Par}}(f)$ in terms of Pareto optimal points of subproblems for f^P. These results are from [MB94].

Proposition 2.20. $X_{\text{w-Par}}(f) = \bigcup_{\substack{P\subset\{1,\dots,Q\} \\ P\neq\emptyset}} X_{\text{Par}}(f^P)$, *for f_i continuous, convex, X convex.*

Proof:

a) „ $\supseteq$ " Take $x \in X,\ x \notin X_{\text{w-Par}}(f) \implies \exists\, \overline{x} \in X \quad f_i(\overline{x}) < f_i(x) \quad \forall\ i = 1,\dots,Q \implies$ $x \notin X_{\text{Par}}(f^P) \quad \forall\ P \subset \{1,\dots,Q\}$

b) „ $\subseteq$ " Take $x \in X,\ x \notin \bigcup_{P\subset\{1,\dots,Q\}} X_{\text{Par}}(f^P)$

$\implies x \notin X_{\text{Par}}(f)$. Let $P = \{1,\dots,Q\}$

$\implies \exists\ i_1 \in P,\ x_1 \in X$ s.t. $f_{i_1}(x_1) < f_{i_1}(x),\ f_i(x_1) \le f_i(x),\ i \neq i_1$. Let $P_1 = P \setminus \{i_1\}$.

Now for $l \ge 1$ and $P_l = \{1,\dots,Q\} \setminus \{i_1,\dots,i_l\}$ suppose we found x_l s.t. $f_i(x_l) < f_i(x)\ \forall\ i \in \{i_1,\dots,i_l\}$ and $f_i(x_l) \le f_i(x)\ \forall\ i \in P_l$.

Since $x \notin X_{\text{Par}}(f^{P_l}) \implies \exists\ i_{l+1} \in P_l \quad \overline{x}_{l+1} \in X$ s.t. $f_{i_{l+1}}(\overline{x}_{l+1}) < f_{i_{l+1}}(x)$ and $f_i(x_{l+1}) \le f_i(x)\ \forall\ i \in P_l$. Then for $x_{l+1} = \lambda x_l + (1-\lambda)\overline{x}_{l+1},\ \lambda \in (0,1)$

- $f_i(x_{l+1}) < f_i(x)\ \forall\ i \in \{i_1,\dots,i_l\}$ for small $(1-\lambda)$ by continuity of f_i.
- $f_{i_{l+1}}(x_{l+1}) \le \lambda f_{i_{l+1}}(x_l) + (1-\lambda) f_{i_{l+1}}(\overline{x}_{l+1}) < \lambda f_{i_{l+1}}(x) + (1-\lambda) f_{i_{l+1}}(x) = f_{i_{l+1}}(x)$ by convexity and induction-hypothesis.
- $f_i(x_{l+1}) \le f_i(x)\ \forall\ i \in P_{l+1} = \{1,\dots,Q\} \setminus \{i_1,\dots,i_{l+1}\}$ by convexity.

After Q steps we have found x_Q such that $f_i(x_Q) < f_i(x)\ \forall\ i = 1,\dots,Q$ i.e. $x \notin X_{\text{w-Par}}(f)$.

$\square$

Proposition 2.20 can be combined with Helly's Theorem again, to obtain

Theorem 2.21. *For convex X, convex and continuous f_i:*

$$X_{\text{w-Par}}(f) = \bigcup_{\substack{P\subset\{1,\dots,Q\} \\ 1\le|P|\le n+1}} X_{\text{Par}}(f^P) \tag{2.13}$$

Proof: We need only consider $Q > n+1$ and only prove „ $\subseteq$ ".

Take $x \in X,\ x \notin \bigcup_{1\le|P|\le n+1} X_{\text{Par}}(f^P)$

Let $J \subset \{1,\dots,Q\},\ J \neq \emptyset,\ |J| \le n+1$

$$\implies x \notin \bigcup_{I\subset J} X_{\text{Par}}(f^I). \text{ By Proposition 2.20} \implies x \notin X_{\text{w-Par}}(f^J).$$

$$\implies \exists\ x_J \in X \quad \text{s.t.} \quad f_j(x_J) < f_j(x)\ \forall\ j \in J \tag{2.14}$$

Now for $i \in \{1,\dots,Q\}$ define

$$C_i = \text{conv}\{x_J \ :\ J \subset \{1,\dots,Q\},\ J \neq \emptyset,\ |J| \le n+1,\ i \in J\}$$

By (2.14) $f_i(x_J) < f_i(x)$ for each $J \subset \{1, \dots, Q\},\ 1 \le |J| \le n+1,\ i \in J$.

Furthermore by convexity

$$f_i(x') < f_i(x) \quad \forall\, x' \in C_i \tag{2.15}$$

Also for fixed J : $\bigcap_{i \in J} C_i \supset \{x_J\}$, i.e. $\bigcap_{i \in J} C_i \neq \emptyset$. By Helly's Theorem $\Longrightarrow$ $\exists\, x^* \in \bigcap_{i=1}^{Q} C_i$ and by (2.15) $f_i(x^*) < f_i(x)$, thus $x \notin X_{\text{w-Par}}(f)$.

□

2.3 Properly Pareto Optimal / Efficient Points

Within the set X_{Par}, it is possible to trade off improvements of one objective for worse values of another. However these trade-offs may be unbounded.

Example 2.5. $Y = X = \{(x_1, x_2) \in \mathbb{R}^2 \ :\ (x_1 - 1)^2 + (x_2 - 1^2) \le 1,\ 0 \le x_1 \le 1,\ 0 \le x_2 \le 1\}$

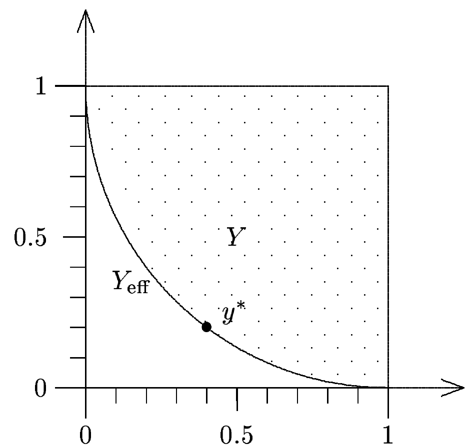


Figure 2.12: Properly Efficient Point

$Y_{\text{eff}} = \{(y_1, y_2) \in Y \ :\ (y_1 - 1)^2 + (y_2 - 1)^2 = 1\}$.

The closer you move from y^* to $(1, 0)$ the smaller is the decrease of y_2 per unit increase of y_1, it actually tends to infinity.

Definition 2.8 (Geoffrion, 1968, [Geo68]). $x^* \in X$ is called **properly Pareto optimal**, if it is Pareto optimal and if $\exists\, M > 0$ s.t. $\forall\, i$ and $x \in X$ satisfying $f_i(x) < f_i(x^*)$ $\Longrightarrow \exists\, j$ s.t. $f_j(x^*) < f_j(x)$ and

$$\frac{f_i(x^*) - f_i(x)}{f_j(x) - f_j(x^*)} \le M. \tag{2.16}$$

$y^* = f(x^*)$ is called **properly efficient**.

Thus, the trade-offs are bounded for properly efficient points.

Example 2.6. In the above Example 2.5 consider the point $y^* = (1, 0)$. To show that it is not properly efficient we have to prove

$\forall\, M > 0 \quad \exists\, i \in \{1, 2\} \quad \exists\, x \in X$ with $f_i(x) < f_i(x^*)$ s.t. $\frac{f_i(x^*) - f_i(x)}{f_j(x) - f_j(x^*)} > M \quad \forall\, j$ s.t. $f_j(x) >$

$f_j(x^*)$.

So choose x with $x_1 = 1 - \varepsilon$ and $x_2 = 1 - \sqrt{1-\varepsilon^2}$.

(Thus $(x_1 - 1)^2 + (x_2 - 1)^2 = 1$, i.e. $x \in X$ and $x_1 < x_1^*,\ x_2 > x_2^*$ so $i = 1,\ j = 2$).

Then

$$\frac{f_i(x^*) - f_i(x)}{f_j(x) - f_j(x^*)} = \frac{1 - (1-\varepsilon)}{1 - \sqrt{1-\varepsilon^2}} = \frac{\varepsilon}{1 - \sqrt{1-\varepsilon^2}} \overset{\varepsilon \to 0}{\longrightarrow} \infty. \tag{2.17}$$

Properly Pareto optimal points are related to the solution of a scalarized problem:
Let $\lambda_i,\ i = 1, \dots, Q$ be nonnegative weights for the objectives s.t. $\sum_{i=1}^{Q} \lambda_i = 1$. Then consider the problem

$$\min_{x \in X} \sum_{i=1}^{Q} \lambda_i f_i(x) \tag{2.18}$$

The following theorems are from [Geo68].

Theorem 2.22. *Let $\lambda_i > 0,\ i = 1, \dots, Q$. If x^* is an optimal solution of* (2.18) *then x^* is properly Pareto optimal.*

Proof: To show that x^* is Pareto optimal consider $x' \in X$ with $f(x') < f(x^*)$
$\Longrightarrow \sum_{i=1}^{Q} \lambda_i f_i(x') < \sum_{i=1}^{Q} \lambda_i f_i(x^*)$ (by positivity of λ_i's), a contradiction.
To show that x^* is properly Pareto optimal let $M := (Q-1) \max_{i,j} \frac{\lambda_j}{\lambda_i}$.
Suppose that x^* is not properly Pareto optimal

$$\begin{aligned}
\Longrightarrow\quad & \exists\, i,\ \exists\, x \in X \text{ s.t. } f_i(x) < f_i(x^*) \\
& f_i(x^*) - f_i(x) > M \cdot (f_j(x) - f_j(x^*)) \text{ for all } j \text{ such that } f_j(x^*) < f_j(x) \\
\Longrightarrow\quad & f_i(x^*) - f_i(x) > \frac{Q-1}{\lambda_i} \lambda_j (f_j(x) - f_j(x^*))\ \forall\, j \neq i \text{ by the choice of } M. \\
\overset{\cdot \frac{\lambda_i}{Q-1},\ \sum_{j \neq i}}{\Longrightarrow}\quad & \lambda_i (f_i(x^*) - f_i(x)) > \sum_{j \neq i} \lambda_j (f_j(x) - f_j(x^*)) \\
\Longrightarrow\quad & \lambda_i f_i(x^*) - \lambda_i f_i(x) > \sum_{j \neq i} \lambda_j f_j(x) - \sum_{j \neq i} \lambda_j f_j(x^*) \\
\Longrightarrow\quad & \lambda_i f_i(x^*) + \sum_{j \neq i} \lambda_j f_j(x^*) > \lambda_i f_i(x) + \sum_{j \neq i} \lambda_j f_j(x) \\
\Longrightarrow\quad & \sum_{i=1}^{Q} \lambda_i f_i(x^*) > \sum_{i=1}^{Q} \lambda_i f_i(x) \quad \text{↯ Contradiction.}
\end{aligned}$$

□

The natural question is, whether this condition is also necessary. This is not true in general.

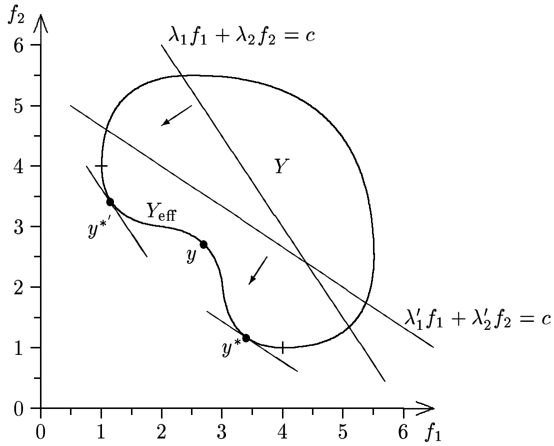


Figure 2.13: Properly Pareto Optimal $y \in Y_{\text{eff}}$

$y \in Y_{\text{eff}}$, properly Pareto optimal but not optimal for (2.18).

Theorem 2.23. *Let $X \subset \mathbb{R}^n$ be convex and $f_i : X \to \mathbb{R}$ be convex. Then $x^* \in X$ is properly Pareto optimal if and only if x^* is optimal for* (2.18).

Proof:

„ $\Longleftarrow$ " Theorem 2.22.

„ $\Longrightarrow$ " Let x^* be a properly Pareto optimal point

$\Longrightarrow \exists\ M > 0$ such that $\forall\ i$ the system

$$\begin{aligned} f_i(x) &< f_i(x^*) \\ f_i(x) + M \cdot f_j(x) &< f_i(x^*) + M \cdot f_j(x^*) \quad \forall\ j \neq i \end{aligned} \tag{2.19}$$

has no solution.

A property of convex functions implies that for the i-th such system there $\exists\ \lambda_j^i \geq 0,\ j = 1, \dots, Q,\ \sum\limits_{j=1}^{Q} \lambda_j^i = 1 \quad \forall\ i$ s.t. $\forall\ x \in X$ holds:

$$\begin{aligned} & \lambda_i^i f_i(x) + \sum_{j\neq i} \lambda_j^i (f_i(x) + M \cdot f_j(x)) \geq \lambda_i^i f_i(x^*) + \sum_{j\neq i} \lambda_j^i (f_i(x^*) + M \cdot f_j(x^*)) \\ \Longleftrightarrow\ & \lambda_i^i f_i(x) + \sum_{j\neq i} \lambda_j^i f_i(x) + M \cdot \sum_{j\neq i} \lambda_j^i f_j(x) \geq \lambda_i^i f_i(x^*) + \sum_{j\neq i} \lambda_j^i f_i(x^*) + M \cdot \sum_{j\neq i} \lambda_j^i f_j(x^*) \\ \Longleftrightarrow\ & \sum_{j=1}^{Q} \lambda_j^i f_i(x) + M \cdot \sum_{j\neq i} \lambda_j^i f_j(x) \geq \sum_{j=1}^{Q} \lambda_j^i f_i(x^*) + M \cdot \sum_{j\neq i} \lambda_j^i f_j(x^*) \\ \Longleftrightarrow\ & f_i(x) + M \cdot \sum_{j\neq i} \lambda_j^i f_j(x) \geq f_i(x^*) + M \cdot \sum_{j\neq i} \lambda_j^i f_j(x^*) \end{aligned}$$

Summing over i:

$$\begin{aligned} \Longrightarrow\ & \sum_{i=1}^{Q} f_i(x) + M \cdot \sum_{j=1}^{Q} \sum_{j\neq i} \lambda_j^i f_j(x) \geq \sum_{i=1}^{Q} f_i(x^*) + M \cdot \sum_{j=1}^{Q} \sum_{j\neq i} \lambda_j^i f_j(x^*) \quad \forall\ x \in X \\ \Longrightarrow\ & \sum_{i=1}^{Q} \underbrace{(1 + M \cdot \sum_{j\neq i} \lambda_j^i)}_{\lambda_i} f_i(x) \geq \sum_{i=1}^{Q} \underbrace{(1 + M \cdot \sum_{j\neq i} \lambda_j^i)}_{\lambda_i} f_j(x^*) \quad \forall\ x \in X \end{aligned}$$

Norming the λ_i to 1 yields the result.

□

Theorem 2.24. *$S \subset \mathbb{R}^n$ convex, $h_i : \mathbb{R}^n \to \mathbb{R}$ convex, $i = 1, \dots, m$. If there is no solution $x \in S$ s.t. $h_i(x) < 0 \quad \forall\, i = 1, \dots, m$ then $\exists\, \lambda_i \geq 0, \; \sum_{i=1}^{m} \lambda_i = 1$ s.t.*

$$\sum_{i=1}^{m} \lambda_i h_i(x) \geq 0 \quad \forall\, x \in S \tag{2.20}$$

See e.g. [Man69, p.65].

Geoffrion is not the only one who introduced properly Pareto optimal points. To look at other definitions, we have to introduce two more cones:

Definition 2.9. Let $Y \subset \mathbb{R}^Q$ and $y \in Y$.

a) The **tangent cone** of Y at y is

$$T_Y(y) := \{d \in \mathbb{R}^Q \;:\; \exists\, t_k \in \mathbb{R},\; y^k \in Y \text{ s.t. } y^k \to y,\; t_k \cdot (y^k - y) \to d\} \tag{2.21}$$

b) The **conical hull** of Y is

$$\text{cone}(Y) = \{\alpha \cdot y \;:\; \alpha \geq 0,\; y \in Y\} = \bigcup_{\alpha \geq 0} \alpha \cdot Y \tag{2.22}$$

Example 2.7.

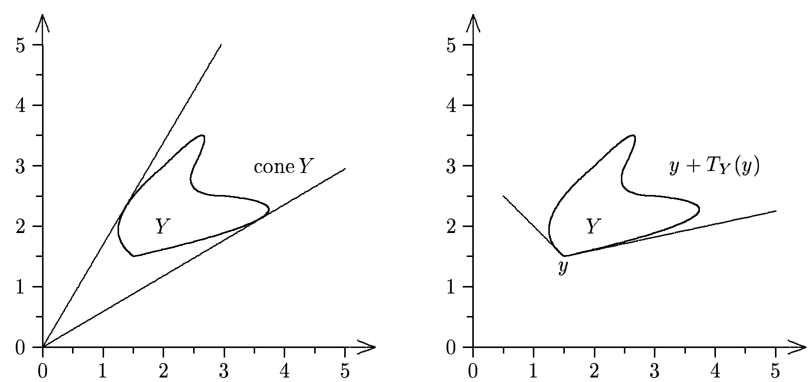


Figure 2.14: Conical Hull and Tangent Cone

Proposition 2.25.

a) $T_Y(y)$ is a closed cone.

b) If Y is convex then $T_Y(y) = \text{cl}(\text{cone}(Y - y))$, which is a closed convex cone.

Proof:

a) Note that $0 \in \text{int}\, T_Y(y)$ (take $y^k = y \;\forall\, k$) and $T_Y(y)$ is indeed a cone: For $\alpha > 0,\; d \in T_Y(y) \implies \alpha \cdot d \in T_Y(y)$. Just take $\alpha \cdot t_k$ instead of t_k.

To see that it is closed take a sequence $\{d_l\} \subset T_Y(y),\; y \in Y$, s.t. $d_l \to d$.

$\forall\, l\; \exists$ sequences $\{y^{l,k}\},\; \{t_{l,k}\}$ as in the definition.

For fixed $l\; \exists\, k_l$ s.t.

$$\|t_{l,k_l}(y^{l,k_l} - y) - d_l\| \leq \frac{1}{l} \tag{2.23}$$

Now if $l \to \infty$ the sequence $t_{l,k_l}(y^{l,k_l} - y) \to d$ i.e. $d \in T_Y(y)$.

b) Let Y be convex, $y \in Y$. By definition, it is obvious that $\mathrm{cl}(\mathrm{cone}(Y-y))$ is a closed convex cone.

„$T_Y(y) \subset \mathrm{cl}(\mathrm{cone}(Y-y))$"

Let $d \in T_Y(y) \implies \exists\, t_k, y^k \;:\; \underbrace{t_k \cdot (y^k - y)}_{\in \alpha\cdot(Y-y)} \to d \quad \checkmark$

„ $\mathrm{cl}(\mathrm{cone}(Y-y)) \subset T_Y(y)$"

$T_Y(y)$ is closed, so only show $\mathrm{cone}(Y-y) \subset T_Y(y)$.

Let $d \in \mathrm{cone}(Y-y) \implies d = \alpha \cdot (y' - y),\ \alpha \geq 0,\ y' \in Y$.

So define $y^k := (1-\frac{1}{k})y + \frac{1}{k}y' \in Y,\ t_k = \alpha \cdot k \geq 0$

$\implies t_k \cdot (y^k - y) = \alpha \cdot ((1-\frac{1}{k})y + \frac{1}{k}y') = \alpha \cdot k \cdot ((\frac{k-1}{k}y + \frac{1}{k}y') - y) = \alpha \cdot ((k-1)y + y' - k \cdot y) = \alpha \cdot (y' - y)$

So $y^k \overset{k\to\infty}{\longrightarrow} y,\ t_k \cdot (y^k - y) \to d \implies d \in T_Y(y)$.

□

Definition 2.10.

a) (Borwein, 1977, [Bor77])

$\hat{x} \in X$ is called **properly Pareto optimal** if

$$T_{Y+\mathbb{R}^Q_+}(f(\hat{x})) \cap (-\mathbb{R}^Q_+) = \{0\} \tag{2.24}$$

b) (Benson, 1979, [Ben79])

$\hat{x} \in X$ is called **properly Pareto optimal** if

$$\mathrm{cl}(\mathrm{cone}(Y + \mathbb{R}^Q_+ - f(\hat{x}))) \cap (-\mathbb{R}^Q_+) = \{0\} \tag{2.25}$$

As we observed in Proposition 2.25 it is immediate from the definitions of conical hulls and tangent cones that

$$T_{Y+\mathbb{R}^Q_+}(f(\hat{x})) \subset \mathrm{cl}(\mathrm{cone}(Y + \mathbb{R}^Q_+ - f(\hat{x}))) \tag{2.26}$$

so the latter definition is stronger.

Theorem 2.26.

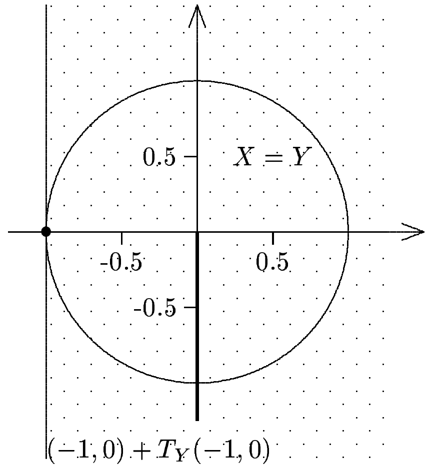
a) If $\hat{x}$ is properly Pareto optimal in Benson's sense, it is also properly Pareto optimal in Borwein's sense.

b) If X is convex and $f_i : \mathbb{R}^n \to \mathbb{R}$ are convex then both definitions coincide.

Proof: Immediate from Proposition 2.25.

□

Example 2.8. Consider $X = \{(x_1, x_2) \;:\; x_1^2 + x_2^2 \leq 1\},\ f_1(x) = x_1,\ f_2(x) = x_2$. Then $(-1, 0),\ (0, -1)$ are Pareto optimal, but not properly Pareto optimal in the sense of Borwein (and thus not in the sense of Benson).



$$T_Y(-1,0) \cap (-\mathbb{R}^Q_+) = \{(y_1, y_2) \in \mathbb{R}^2 \;:\; y_1 = 0,\; y_2 \leq 0\}$$

$$T_Y(-1,0) = \{(y_1, y_2) \in \mathbb{R}^2 \;:\; y_1 \geq 0\} \tag{2.27}$$

Figure 2.15: Benson's Proper Efficiency

Are proper Pareto optimal points in Benson's or Borwein's sense always Pareto optimal ?

Proposition 2.27. *If $\hat{x}$ is properly Pareto optimal in the sense of Borwein, then $\hat{x}$ is also Pareto optimal.*

Proof: Exercise 17.

□

Note. In these definitions, $\mathbb{R}^Q_+$ can be replaced by an arbitrary closed convex cone K.

Theorem 2.28. *$\hat{x}$ is properly Pareto optimal in the sense of Geoffrion (Definition 2.8) if and only if it is properly Pareto optimal in the sense of Benson.*

Proof:

„ $\Longrightarrow$ "

Suppose $\hat{x}$ is Pareto optimal, but not Benson-properly Pareto optimal

$\Longrightarrow \quad \exists\, 0 \neq d \in \mathrm{cl}(\mathrm{cone}(Y + \mathbb{R}^Q_+ - f(\hat{x}))) \cap (-\mathbb{R}^Q_+)$

Wlog assume $d_1 < -1,\ d_i \leq 0,\ i = 2, \ldots, Q$ (otherwise reorder components of f, rescale d). So there are sequences $t_k \in \mathbb{R}_+ \setminus \{0\},\ x^k \in X,\ r^k \in \mathbb{R}^Q_+$ s.t. $t_k \cdot (f(x^k) + r^k - f(\hat{x})) \to d$.

After choosing a subsequence we can assume that $\tilde{Q} = \{i \in \{1, \ldots, Q\} \;:\; f_i(x^k) > f_i(\hat{x})\}$ is the same for all k and nonempty ($\hat{x}$ is Pareto optimal).

Let $M > 0 \quad \Longrightarrow \quad \exists\, k_0$ s.t. $\forall\, k \geq k_0$

$$f_1(x^k) - f_1(\hat{x}) < -\tfrac{1}{2 \cdot t_k} \tag{2.28}$$

$$\text{and} \qquad f_i(x^k) - f_i(\hat{x}) \leq -\tfrac{1}{2 \cdot M t_k} \tag{2.29}$$

$$\implies \quad \forall\, i \in \tilde{Q} \quad \forall\, k \geq k_0 \quad 0 < f_i(x^k) - f_i(\hat{x}) \leq \frac{1}{2 \cdot M t_k}$$

$$\implies \quad \frac{f_1(\hat{x}) - f_1(x^k)}{f_i(x^k) - f_i(\hat{x})} > \frac{\frac{1}{2 \cdot t_k}}{\frac{1}{2 \cdot M t_k}} = M \tag{2.30}$$

$\implies$ $\hat{x}$ is not properly Pareto optimal in Geoffrin's sense.

„ $\Longleftarrow$ “

Suppose $\hat{x}$ is Pareto optimal, but not properly Pareto optimal in the sense of Geoffrion. Let $M_k > 0$ be an unbounded sequence of positive real numbers. Wlog we assume that $\forall\, M_k\ \exists\, x^k \in X$ s.t. $f_1(x^k) < f_1(\hat{x})$ and

$$\frac{f_1(\hat{x}) - f_1(x^k)}{f_i(x^k) - f_i(\hat{x})} > M_k \quad \forall\, i \in \{2, \ldots, Q\} \text{ s.t. } f_i(x^k) > f_i(\hat{x}) \tag{2.31}$$

Again, choosing a subsequence we can assume

$$\tilde{Q} = \{i \in \{1, \ldots, Q\} \ : \ f_i(x^k) > f_i(\hat{x})\} \tag{2.32}$$

is constant for all k and nonempty.

Define $t_k := (f_1(\hat{x}) - f_1(x^k))^{-1} \quad \implies \quad t_k > 0\ \forall\, k$.

Define $r_i^k := \begin{cases} 0 & , i = 1,\ i \in \tilde{Q} \\ f_i(\hat{x}) - f_i(x^k) & , \text{else} \end{cases}$

$\implies \quad r^k \in \mathbb{R}_+^Q$

$$t_k \cdot (f_i(x^k) + r_i^k - f_i(\hat{x})) \begin{cases} = -1 & , i = 1 \\ = 0 & , i \neq 1,\ i \notin \tilde{Q} \\ \in (0, M_k^{-1}) & , i \in \tilde{Q} \end{cases} \tag{2.33}$$

If we use $d_i = \lim_{k \to \infty} t_k \cdot (f_i(x^k) + r_i^k - f_i(\hat{x}))$

$\implies \quad d_1 = -1,\ d_i = 0,\ i \neq 1,\ i \notin \tilde{Q},\ d_i = 0,\ i \in \tilde{Q} \quad (M_k \to \infty)$

$d = (-1, 0, \ldots, 0) \in \text{cl}(\text{cone}(f(X) + \mathbb{R}_+^Q - f(\hat{x}))) \cap (-\mathbb{R}_+^Q)$

□

Despite Theorem 2.28, Benson's proper efficiency is more general than Geoffrion's, because it still can be used when a closed convex cone K is used as ordering cone of $\mathbb{R}^Q$ instead of $\mathbb{R}_+^Q$.

In multicriteria optimization we will often encounter problems, where X is given explicitely as

$$X = \{x \in \mathbb{R}^n \ : \ (g_1(x), \ldots, g_m(x)) \leq 0\} \tag{2.34}$$

Let us assume that $f_i,\ i = 1, \ldots, Q$ and $g_j,\ j = 1, \ldots, m$ are continuously differentiable and consider

$$\min_{x \in X} f(x) \tag{2.35}$$

Definition 2.11 (Kuhn + Tucker, 1951, [KT51]).

A solution $\hat{x} \in X$ is called **properly Pareto optimal** if it is Pareto optimal, and if there is no $h \in \mathbb{R}^n$ s.t.

$$\langle \nabla f_i(\hat{x}), h \rangle \leq 0 \qquad \forall\, i = 1, \dots, Q \tag{2.36}$$

$$\langle \nabla f_i(\hat{x}), h \rangle < 0 \qquad \text{for some } i \tag{2.37}$$

$$\text{and} \qquad \langle \nabla g_j(\hat{x}), h \rangle \leq 0 \qquad \forall\, j \in J(\hat{x}) = \{j \;:\; g_j(\hat{x}) = 0\} \tag{2.38}$$

Theorem 2.29. *If $\hat{x}$ is properly Pareto optimal in the sense of Kuhn-Tucker there exist $\hat{\mu} \in \mathbb{R}^Q$, $\hat{\lambda} \in \mathbb{R}^m$ such that*

$$i) \qquad \sum_{i=1}^{Q} \hat{\mu}_i \nabla f_i(\hat{x}) + \sum_{j=1}^{m} \hat{\lambda}_j \nabla g_j(\hat{x}) = 0 \tag{2.39}$$

$$ii) \qquad \sum_{j=1}^{m} \hat{\lambda}_j g_j(\hat{x}) = 0 \tag{2.40}$$

$$iii) \qquad \hat{\mu} \gg 0, \; \hat{\lambda} \geq 0 \tag{2.41}$$

Proof: $\hat{x}$ is properly Pareto optimal

$\Longrightarrow\ \nexists\, h \in \mathbb{R}^n$ s.t.

$$\langle \nabla f_i(\hat{x}), h \rangle \leq 0 \qquad \forall\, i = 1, \dots, Q \tag{2.42}$$

$$\langle \nabla f_{i^*}(\hat{x}), h \rangle < 0 \qquad \text{for some } i^* \tag{2.43}$$

$$\text{and} \qquad \langle \nabla g_j(\hat{x}), h \rangle \leq 0 \qquad \forall\, j \in J(\hat{x}) \tag{2.44}$$

We use Tucker's Theorem of the alternative to get $\mu_i > 0 \quad i = 1, \dots, Q,\ \hat{\lambda}_j \geq 0 \quad j \in J(\hat{x})$ s.t.

$$\sum_{i=1}^{Q} \hat{\mu}_i \nabla f_i(\hat{x}) + \sum_{j \in J(\hat{x})} \hat{\lambda}_j \nabla g_j(\hat{x}) = 0 \tag{2.45}$$

Letting $\hat{\lambda}_j = 0 \quad \forall\, j \in \{1, \dots, Q\} \setminus J(\hat{x})$, the proof is completed.

□

Theorem 2.30 (Tucker's Theorem of the alternative). [Man69, p.25]

Let B, C and D be $Q \times n$, $k \times n$ and $o \times n$ matrices. Then either

a) $Bx < 0,\ Cx \leq 0,\ Dx = 0$ *has a solution* $x \in \mathbb{R}^n$

or b) $B^t y_1 + C^t y_2 + D^t y_3 = 0,\ y_1 \gg 0,\ y_2 \geq 0$ *has a solution* $y_1,\ y_2,\ y_3$

but never both.

Proof: See [Man69, p.29].

□

Remark. Consider

$$D = 0, \quad B = \begin{pmatrix} \nabla f_1(\hat{x}) \\ \vdots \\ \nabla f_Q(\hat{x}) \end{pmatrix}, \quad C = (\nabla J_j(\hat{x}),\ j \in J(\hat{x})), \quad k = |J(\hat{x})|, \quad h = x$$

$$y^1 = \hat{\mu}, \quad y^2 = \hat{\lambda}, \quad y^3 = 0$$

in Theorem 2.29.

Under additional assumptions, we can show that a properly Pareto optimal point in Geoffrion's sense is properly Pareto optimal in Kuhn-Tucker's sense.

Definition 2.12. A differentiable MOP satisfies the KT constraint qualification at $\hat{x} \in X$ if $\forall\ h \in \mathbb{R}^n$ s.t. $\langle \nabla g_j(\hat{x}), h\rangle \leq 0 \ \ \forall\ j \in J(\hat{x}) \ \ \exists\ \bar{t} > 0$, a function $\theta : [0, \bar{t}] \to \mathbb{R}^n$, and $\alpha > 0$ s.t. $\theta(0) = \hat{x}, \quad g(\theta(\bar{t})) \leq 0 \quad \theta'(0) = \alpha h$.

Theorem 2.31. *If a differentiable MOP satisfies the KT constraint qualification at $\hat{x}$ and $\hat{x}$ is Geoffrion properly Pareto optimal then it is KT properly Pareto optimal.*

Proof: Suppose $\hat{x}$ is Pareto optimal, but not KT properly Pareto optimal.
$\Longrightarrow \exists\ h \in \mathbb{R}^n$ s.t. (wlog)

$$\langle \nabla f_1(\hat{x}), h\rangle < 0 \tag{2.46}$$

$$\langle \nabla f_i(\hat{x}), h\rangle \leq 0 \qquad \forall\ i = 2, \ldots, Q \tag{2.47}$$

$$\langle \nabla g_j(\hat{x}), h\rangle \leq 0 \qquad \forall\ j \in J(\hat{x}) \tag{2.48}$$

Using the function θ from the constraint qualification we take a sequence $\{t_k\} \to 0$, and if necessary a subsequence s.t.

$$\tilde{Q} = \{i \ :\ f_i(\theta(t_k)) > f_i(\hat{x})\} \tag{2.49}$$

is constant. Since for $i \in \tilde{Q}$

$$\begin{aligned} & f_i(\theta(t_k)) - f_i(\hat{x}) = t_k \langle \nabla f_i(\hat{x}), \alpha h\rangle + o(t_k) > 0 \\ \text{and} \quad & \langle \nabla f_i(\hat{x}), h\rangle \leq 0 \\ \Longrightarrow \quad & \langle \nabla f_i(\hat{x}), \alpha h\rangle = 0 \qquad \forall\ i \in \tilde{Q} \end{aligned} \tag{2.50}$$

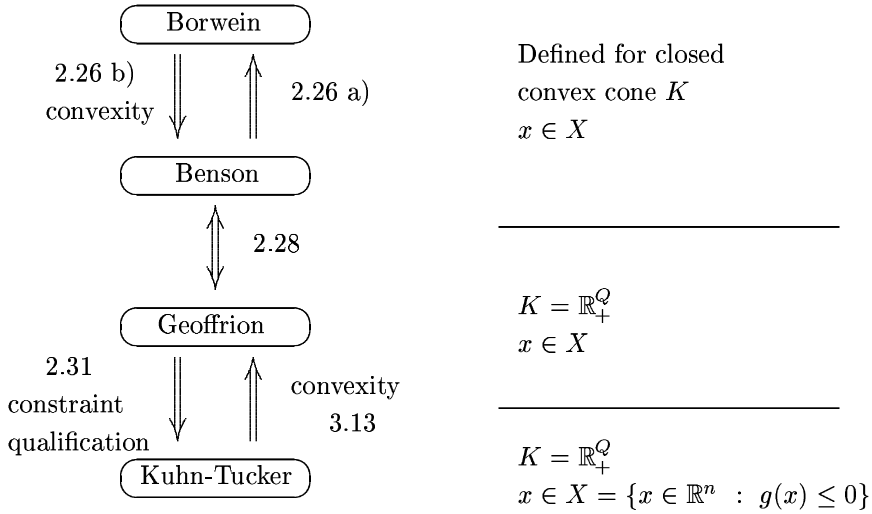
But since $\langle \nabla f_1(\hat{x}), h\rangle < 0$

$$\Longrightarrow \quad \frac{f_1(\hat{x}) - f_1(\theta(t_k))}{f_i(\theta(t_k)) - f_i(\hat{x})} = \frac{-\langle \nabla f_1(\hat{x}), \alpha h\rangle + \frac{o(t_k)}{t_k}}{\langle \nabla f_i(\hat{x}), \alpha h\rangle + \frac{o(t_k)}{t_k}} \to \infty \tag{2.51}$$

$\Longrightarrow \hat{x}$ is not Geoffrion properly Pareto optimal.

□

Let us summarize the definitions of proper Pareto optimality:



In order to derive further results on properly Pareto optimal points and on topological properties, we have to investigate weighted sum scalarizations in greater detail.

2.4 Exercises to Chapter 2

5. Prove, or give a counterexample to the converse inclusion in Proposition 2.4.

6. Given a cone $K \subset \mathbb{R}^n$ and the related order $\leq_K$ we say $y^* \in Y$ is K-efficient if $\nexists\, y \in Y,\ y \neq y^*$ s.t. $y^* \in y - K$. Let K_1, K_2 be two cones in $\mathbb{R}^n$. $K_1 \subset K_2$. Then if y^* is K_2-efficient it is also K_1-efficient. Illustrate this „large cone - fewer efficient points“ result graphically.

7. Let $f : \mathbb{R}^n \to \mathbb{R}^Q$ be the criteria in a multi-criteria optimization problem. Show that any optimal solution x^* of the problem

$$\operatorname*{lexmin}_{x \in X}(f_1(x), \dots, f_Q(x))$$

is Pareto optimal.

8. Prove that $(\alpha Y)_{\text{eff}} = \alpha(Y_{\text{eff}})$ where $Y \subset \mathbb{R}^n$ is a nonempty set and $\alpha > 0$.

9. Show that a function $f : \mathbb{R}^n \to \mathbb{R}^Q$ is $\mathbb{R}^Q_+$-semicontinuous if and only if $f_i : \mathbb{R}^n \to \mathbb{R}$ are lower semicontinuous $\forall\, i = 1, \dots, Q$.

   $f_i : \mathbb{R}^n \to \mathbb{R}$ is called **lower semicontinuous** if $f_i(x^*) \leq \liminf_{x \to x^*} f_i(x) = l$, where l satisfies $\forall\, \varepsilon > 0 \quad \exists\, U(x^*, \varepsilon)$

$$\begin{aligned} \text{s.t.} \quad & f(x) \geq l - \varepsilon \quad \forall\, x \in U(x^*, \varepsilon) \quad \text{and} \\ & \exists\, x \in U(x^*, \varepsilon) \quad f(x) \leq l + \varepsilon \end{aligned}$$

10. Let $Y \subset \mathbb{R}^n$ be a convex set. The recession cone (or asymptotic cone) of Y, Y_∞ is defined as

$$Y_\infty := \{d \in \mathbb{R}^n : \exists\, y \text{ s.t. } y + \alpha d \in Y \quad \forall\, \alpha > 0\}$$

i.e. the set of directions in which Y extends infinitely.

a) Show that Y is bounded if and only if $Y_\infty = \{0\}$.

b) Let $Y = \{(y_1, y_2) \in \mathbb{R}^2 : y_2 \geq y_1^2\}$. Determine Y_∞.

11. A set $Y \subset \mathbb{R}^n$ is called

- $\mathbb{R}^n_+$-closed, if $Y + \mathbb{R}^n_+$ is closed and
- $\mathbb{R}^n_+$-bounded, if $Y_\infty \cap (-\mathbb{R}^n_+) = \{0\}$.

Give examples of sets $Y \subset \mathbb{R}^2$ that are

- $\mathbb{R}^2_+$-compact, $\mathbb{R}^2_+$-bounded, not $\mathbb{R}^2_+$-closed.
- $\mathbb{R}^2_+$-bounded, $\mathbb{R}^2_+$-closed, not $\mathbb{R}^2_+$-compact.

12. Prove the following existence result.

Let $\emptyset \neq Y \subset \mathbb{R}^Q$ such that Y is $\mathbb{R}^Q_+$-compact. Show that $Y_{\text{w-eff}} \neq \emptyset$.

(Do not use Corollary 2.10 and the fact that $Y_{\text{eff}} \subset Y_{\text{w-eff}}$)

13. Recall the definition of K-efficiency from Exercise 6:

$y^* \in Y$ is K-efficient if $\nexists\, y^* \in Y$ s.t. $y \in y + K$.

Verify that Proposition 2.1 is still true if K is a pointed, convex cone. Give examples that the inclusion $Y_{\text{K-eff}} \subset (Y + K)_{\text{K-eff}}$ is not true when K is not pointed and when K is not convex.

14. Let $[a, b] \subset \mathbb{R}$ be a compact interval.

Suppose that $f_i : \mathbb{R} \to \mathbb{R}$ are convex $\forall\, i = 1, \ldots, Q$.

Let

$$\begin{aligned} x_i^m &= \min\left\{x \in [a, b] : f_i(x) = \min_{x \in [a,b]} f_i(x)\right\} \quad \text{and} \\ x_i^M &= \max\left\{x \in [a, b] : f_i(x) = \min_{x \in [a,b]} f_i(x)\right\}. \end{aligned}$$

Using Theorem 2.16 show that

$$\begin{aligned} X_{\text{Par}} &= \left[\min_{i=1,\ldots,Q} x_i^M, \max_{i=1,\ldots,Q} x_i^m\right] \cup \left[\max_{i=1,\ldots,Q} x_i^m, \min_{i=1,\ldots,Q} x_i^M\right] \\ X_{\text{w-Par}} &= \left[\min_{i=1,\ldots,Q} x_i^m, \max_{i=1,\ldots,Q} x_i^M\right] \end{aligned}$$

15. Use the result of Exercise 14 to give an example of a problem with $X \subset \mathbb{R}$ where $X_{\text{s-Par}} \subset X_{\text{Par}} \subset X_{\text{w-Par}}$, with strict inclusions. Use 2 or 3 objective functions.

16. Let $X = \{x \in \mathbb{R} : x \geq 0\}$ and $f_1(x) = e^x$,

$$f_2(x) = \begin{cases} \frac{1}{x+1} & 0 \leq x \leq 5 \\ (x-5)^2 + \frac{1}{6} & x \geq 5 \end{cases}$$

Using the result of Exercise 14, determine X_{Par}. Which of these points are strictly Pareto optimal ? Can you prove a sufficient condition on f for $x \in \mathbb{R}$ to be a strictly Pareto optimal point of $\min\limits_{x \in X \subset \mathbb{R}} f(x)$? Derive a conjecture from the example and try to prove it.

17. Show that if $\hat{x}$ is properly Pareto optimal in the sense of Borwein, then $\hat{x}$ is Pareto optimal.

18. Consider the following example:

$$x = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 \geq 0\} \cup \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 1\} \cup \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq 1\}$$

With $f_1(x) = x_1$, $f_2(x) = x_2$.

Show that $x = 0$ is properly Pareto optimal in the sense of Borwein, but not in the sense of Benson.

19. Consider the problem

$$\begin{aligned} \min \quad & \left[(x_1-2)^2 + (x_2-1)^2,\ x_1^2 + (x_2-3)^2\right] \\ \text{s.t.} \quad & g_1(x) = x_1^2 - x_2 \leq 0 \\ & g_2(x) = x_1 + x_2 - 2 \leq 0 \\ & g_3(x) = -x_1 \leq 0 \end{aligned}$$

Use the conditions of Theorem 2.29 to find at least one candidate for a properly Pareto optimal point $\hat{x}$ (in the sense of Kuhn-Tucker). Try to determine all.

20. Consider an MOP $\min\limits_{x \in X} f(x)$ with Q objectives. Add a new objective f^{Q+1}. Is the Pareto set of the new problem bigger, smaller or the same than that of the original problem ?

Chapter 3

Weighted Sum Scalarization

In this chapter we will investigate to what extent an MOP

$$\min_{x \in X} (f_1(x), \dots, f_Q(x)) \tag{3.1}$$

can be solved by solving scalarized problems of the type

$$\min_{x \in X} \sum_{i=1}^{Q} \lambda_i f_i(x) \tag{3.2}$$

In terms of our classification, this problem is written as

$$(X, f, \mathbb{R}^Q)/\langle \lambda, \cdot \rangle/(\mathbb{R}, \leq) \tag{3.3}$$

where $\langle \lambda, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^Q$.

Again, we will focus on the objective space Y. We shall see the relations between solutions of scalarized problems and Y_{eff}, $Y_{\text{w-eff}}$ and properly efficient points.

Let $Y \subset \mathbb{R}^Q$. For a fixed $\lambda \in \mathbb{R}^Q$ we denote by

$$\text{Opt}(\lambda, Y) := \{y^* \in Y \ : \ \langle \lambda, y^* \rangle = \inf_{y \in Y} \langle \lambda, y \rangle\} \tag{3.4}$$

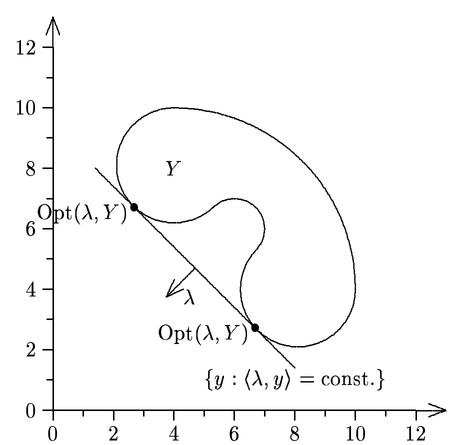


Figure 3.1: $\text{Opt}(\lambda, Y)$

It is obvious that we can assume $\|\lambda\| = 1$, usually $\sum \lambda_i = 1,\ \lambda_i \geq 0$. As we already know, sometimes we have to assume that all $\lambda_i > 0$. So we distinguish:

$$S(Y) = \bigcup_{\lambda \in \operatorname{int} \mathbb{R}^Q_+} \operatorname{Opt}(\lambda, Y) = \bigcup_{\substack{\lambda:\lambda_i>0 \\ \sum \lambda_i = 1}} \operatorname{Opt}(\lambda, Y) \tag{3.5}$$

and

$$S_0(Y) = \bigcup_{\lambda \in \mathbb{R}^Q_+ \setminus \{0\}} \operatorname{Opt}(\lambda, Y) = \bigcup_{\substack{\lambda:\lambda_i \geq 0 \\ \sum \lambda_i = 1}} \operatorname{Opt}(\lambda, Y) \tag{3.6}$$

Obviously $S(Y) \subset S_0(Y)$.

3.1 Scalarization and Efficiency

Theorem 3.1. $S(Y) \subset Y_{\text{eff}}$.

Proof: Assume $y^* \in S(Y) \implies \exists\, \lambda \in \mathbb{R}^Q_+,\ \lambda_i > 0$ s.t. $\sum\limits_{i=1}^{Q} \lambda_i y_i^* \leq \sum\limits_{i=1}^{Q} \lambda_i y_i$.

Suppose $y_i^* \notin Y_{\text{eff}} \implies \exists\, y' \in Y \quad y_i^* \geq y_i' \quad \forall\, i = 1, \ldots, Q$ and strict inequality for one i.

$\implies \lambda_i y_i^* \geq \lambda_i y_i'$ and strict inequality for one i, since $\lambda_i > 0$.

$\implies \sum\limits_{i=1}^{Q} \lambda_i y_i^* > \sum\limits_{i=1}^{Q} \lambda_i y_i'$ ↯ Contradiction.

□

Theorem 3.2. *$Y_{\text{eff}} \subset S_0(Y)$ when Y is an $\mathbb{R}^Q_+$-convex set.*

In order to prove this result, we need a separation theorem.

Theorem 3.3. *Let $S_1, S_2 \subset \mathbb{R}^Q$ be nonempty convex sets. Then $\exists\, x^* \in \mathbb{R}^Q$ s.t.*

$$\inf_{x \in S_1} \langle x, x^* \rangle \geq \sup_{x \in S_2} \langle x, x^* \rangle \tag{3.7}$$

and

$$\sup_{x \in S_1} \langle x, x^* \rangle > \inf_{x \in S_2} \langle x, x^* \rangle \tag{3.8}$$

if and only if $\operatorname{ri}(S_1) \cap \operatorname{ri}(S_2) = \emptyset$. S_1 and S_2 are said to be properly separated by a hyperplane with normal x^*.

Remark. $\operatorname{ri}(S_i)$ is the relative interior of S_i, i.e. the interior in the space of appropriate dimension $\dim(S_i) \leq Q$.

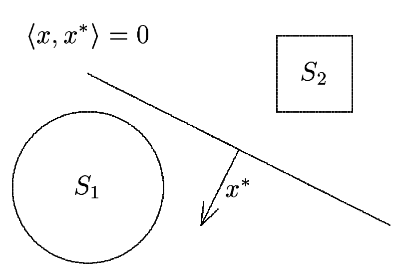


Figure 3.2: Properly Separated S_1 and S_2

Proof of Theorem 3.2. Let $y^* \in Y_{\text{eff}} \implies y^* \in (Y+\mathbb{R}^Q_+)_{\text{eff}} \implies (Y+\mathbb{R}^Q_+-y^*)\cap(-\mathbb{R}^Q_+) = \{0\}$.

Both sets are convex and the intersection of their relative interior is empty.

By Theorem 3.3 $\exists\ \lambda \in \mathbb{R}^Q$ s.t.

$$\langle \lambda, y+d-y^* \rangle \geq 0 \geq \langle y, -d' \rangle \quad \forall\ y \in Y,\ d \in \mathbb{R}^Q_+,\ d' \in \mathbb{R}^Q_+$$

Now $\langle \lambda, -d' \rangle \leq 0 \quad \forall\ d' \in \mathbb{R}^Q_+ \implies \lambda_i \geq 0$ (choose $d' = (0,\dots,0,\underset{i}{1},0,\dots,0)$).

On the other hand

$$\langle \lambda, y+d-y^* \rangle \geq 0 \quad \text{implies (take } d=0)$$
$$\langle \lambda, y \rangle \geq \langle \lambda, y^* \rangle \quad \forall\ y \in Y$$

$\implies y^* \in \text{Opt}(\lambda, Y) \subset S_0(Y)$.

□

So for ($\mathbb{R}^Q_+$-) convex sets, we have the inclusions

$$S(Y) \subset Y_{\text{eff}} \subset S_0(Y). \tag{3.9}$$

It is possible to find examples where both inclusions are strict, see the Exercise 22.

Theorem 3.1 can be extended by the following Proposition.

Proposition 3.4. *If y^* is the unique element in* $\text{Opt}(\lambda, Y)$ *for some $\lambda \in \mathbb{R}^Q_+$ then $y^* \in Y_{\text{eff}}$.*

Proof: Exercise 22.

□

In the next two sections we will discuss which subset / superset of Y_{eff} is described by $S(Y)$ respectively $S_0(Y)$.

3.2 Scalarization and Weak / Proper Efficiency

Theorem 3.5. $S_0(Y) \subset Y_{\text{w-eff}}$.

Proof: Let $\lambda \in \mathbb{R}^Q_+,\ \sum_{i=1}^{Q} \lambda_i = 1,\ \lambda_i \geq 0$ and $y^* \in \text{Opt}(\lambda, Y)$.

$\implies \sum_{i=1}^{Q} \lambda_i y_i^* \leq \sum_{i=1}^{Q} \lambda_i y_i \quad \forall\ y \in Y$.

Suppose $y^* \notin Y_{\text{w-eff}} \implies \exists\ y^* \in Y \quad y_i' < y_i^* \quad \forall\ i$

$\implies \sum_{i=1}^{Q} \lambda_i y_i' < \sum_{i=1}^{Q} \lambda_i y_i^*$ since $\exists\ i\ :\ \lambda_i > 0$ $\quad\lightning$ Contradiction

□

For convex sets we can prove the converse inclusion.

Theorem 3.6. *If Y is $\mathbb{R}^Q_+$-convex* $Y_{\text{w-eff}} = S_0(Y)$.

Proof: We only have to show $Y_{\text{w-eff}} \subset S_0(Y)$.

We observe that $Y_{\text{w-eff}} \subset (Y + \text{int}\,\mathbb{R}^Q_+)_{\text{w-eff}}$. (Proof is the same as that of Proposition 2.1)

So if $y^* \in Y_{\text{w-eff}} \implies \underbrace{(Y + \text{int}\,\mathbb{R}^Q_+ - y^*)}_{\text{convex}} \cap (-\,\text{int}\,\mathbb{R}^Q_+) = \emptyset$.

Therefore we can proceed exactly as in the proof of Theorem 3.2 to get $\lambda \in \mathbb{R}^Q_+$ s.t.

$$\sum_{i=1}^{Q} \lambda_i y_i^* \leq \sum_{i=1}^{Q} \lambda_i y_i \quad \forall\, y \in Y.$$

(Note that the 0's in the choices of d', d in the proof of Theorem 3.2 can be replaced by arbitrary small ε.)

□

We will now deal with properly Pareto optimal points in the sense of Benson / Geoffrion and denote the set of properly efficient points by $Y_{\text{p-eff}}$.

From Theorems 2.22 and 2.28 we immediately derive

Corollary 3.7. $S(Y) \subset Y_{\text{p-eff}}$.

As a generalization of Theorem 2.23 we can show the converse inclusion for $\mathbb{R}^Q_+-$convex sets.

Note that if X is convex and all f_i are convex then $Y = f(X)$ is a convex set.

Theorem 3.8. *If Y is $\mathbb{R}^Q_+$-convex $Y_{\text{p-eff}} \subset S(Y)$.*

Proof: Let $y^* \in Y_{\text{p-eff}}$.

$\implies \text{cl}(\text{cone}(Y + \mathbb{R}^Q_+ - y^*)) \cap (-\mathbb{R}^Q_+) = \{0\}$.

From Proposition 2.25 $\text{cl}(\text{cone}(Y + \mathbb{R}^Q_+ - y^*))$ is a closed convex cone.

We show $\exists\, \lambda \in \text{int}\,\mathbb{R}^Q_+$ s.t.

$$\langle \lambda, d \rangle \geq 0 \quad \forall\, d \in \text{cl}(\text{cone}(Y + \mathbb{R}^Q_+ - y^*)) =: K \tag{3.10}$$

Then, since $Y - y^* \subset \text{cl}(\text{cone}(Y + \mathbb{R}^Q_+ - y^*))$ we especially get

$$\begin{aligned} & \langle \lambda, y - y^* \rangle \geq 0 \quad \forall\, y \in Y \\ \implies & \langle \lambda, y \rangle \geq \langle \lambda, y^* \rangle \quad \forall\, y \in Y \\ \implies & y^* \in S(Y) \end{aligned}$$

<u>Proof of the claim (3.10):</u>

Assume no such λ exists. Since both $\text{int}\,\mathbb{R}^Q_+$ and $K^0 := \{\mu \in \mathbb{R}^n : \langle \mu, d \rangle \geq 0 \;\; \forall\, d \in K\}$ are convex sets we apply Theorem 3.3 again to get $x^* \in \mathbb{R}^Q_+$, $x^* \neq 0$ and $\beta \in \mathbb{R}$ s.t.

$$\begin{aligned} \langle x^*, \mu \rangle \leq \beta & \quad \forall\, \mu \in \text{int}\,\mathbb{R}^Q_+ \\ \langle x^*, \mu \rangle \geq \beta & \quad \forall\, \mu \in K^0 \end{aligned}$$

Using $d = \lambda d'$, $\lambda \to \infty$ we get $\beta = 0$.

$$\begin{aligned} \implies & \langle x^*, \mu \rangle \leq 0 && \forall\, \mu \in \text{int}\,\mathbb{R}^Q_+ \\ \implies & x_i^* \leq 0 && \forall\, i \quad (\text{use } \mu = (\varepsilon, \dots, \varepsilon, \underset{i}{1}, \varepsilon, \dots, \varepsilon),\; \varepsilon \to 0) \\ \implies & x^* \in -\mathbb{R}^Q_+ \setminus \{0\} && \end{aligned} \tag{3.11}$$

Let $K^{00} := \{x \in \mathbb{R}^n \ : \ \langle x, \mu \rangle \geq 0 \quad \forall \ \mu \in K^0\}$, we show $K^{00} \subset \operatorname{cl} K = K$.

Then, since

$$x^* \in K^{00} \Longrightarrow x^* \in K \tag{3.12}$$

Finally (3.11), (3.12) $\Longrightarrow$ $x^* \in K \cap (-\mathbb{R}^Q_+),\ x^* \neq 0$ ⚡ Contradiction

Therefore the desired λ exists in (3.10).

<u>Proof of $K^{00} \subset \operatorname{cl} K = K$:</u>

Let $x \in \mathbb{R}^n,\ x \notin K$. Using Theorem 3.3 once more we get $x_0 \in \mathbb{R}^n,\ x_0 \neq 0,\ \alpha \in \mathbb{R}$ with $\langle d, x_0 \rangle > \alpha \quad \forall\ d \in K$ and $\langle x, x_0 \rangle < \alpha$. Then $0 \in K \Longrightarrow \alpha < 0 \Longrightarrow \langle x, x_0 \rangle < 0$. Taking $d = \lambda d',\ \lambda \to \infty$ we get $\langle d, x_0 \rangle \geq 0 \quad \forall\ d \in K \Longrightarrow x_0 \in K^0$.

So $\langle x, x_0 \rangle < 0$ implies $x \notin K^{00}$, hence $K^{00} \subset K$.

□

Example 3.1. $Y = \{(y_1, y_2) : y_1^2 + y_2^2 \leq 1\}$

$Y_{\text{eff}} = \{(y_1, y_2) : y_1^2 + y_2^2 = 1,\ y_1 \leq 0,\ y_2 \leq 0\}$

$Y_{\text{p-eff}} = Y_{\text{eff}} \setminus \{\binom{-1}{0}, \binom{0}{-1}\}$

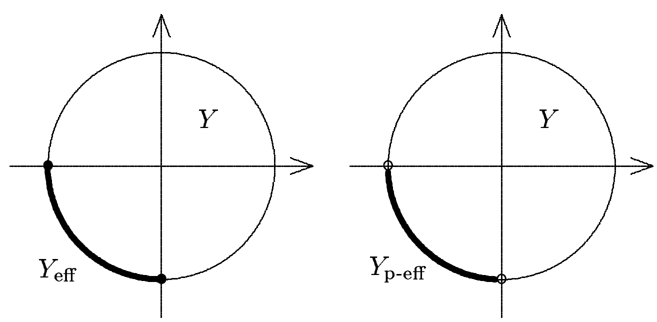


Figure 3.3: Y_{eff} and $Y_{\text{p-eff}}$

But $(-1, 0)$ and $(0, -1)$ are unique solutions of

$$\min_{y \in Y} \lambda_1 y_1 + \lambda_2 y_2 \tag{3.13}$$

for $\lambda = (1, 0)$ and $\lambda = (0, 1)$, respectively, and therefore contained in Y_{eff}.

Our results up to now are

$$S(Y) \subset Y_{\text{p-eff}} \quad \text{and} \quad S_0(Y) \subset Y_{\text{w-eff}} \tag{3.14}$$

in general and

$$S(Y) = Y_{\text{p-eff}} \subset Y_{\text{eff}} \subset Y_{\text{w-eff}} = S_0(Y) \tag{3.15}$$

when Y is $\mathbb{R}^Q_+$-convex.

Therefore, a characterization of Y_{eff} is not possible. But we can show that $Y_{\text{p-eff}}$ is dense in Y_{eff}.

Theorem 3.9. *If $Y \neq \emptyset$ is $\mathbb{R}^Q_+$-closed and $\mathbb{R}^Q_+$-convex we have*

$$S(Y) \subset Y_{\text{eff}} \subset \operatorname{cl} S(Y) = \operatorname{cl} Y_{\text{p-eff}} \tag{3.16}$$

Proof: The only inclusion we have to show is $Y_{\text{eff}} \subset \operatorname{cl} S(Y)$.

Since $Y_{\text{eff}} = (Y + \mathbb{R}^Q_+)_{\text{eff}}$ and $S(Y) = S(Y + \mathbb{R}^Q_+)$, we only prove it for a closed convex Y.

Wlog we use $\hat{y} = 0 \in Y_{\text{eff}}$.

Case 1: Y is compact convex.

Choose $\overline{\lambda} \in \operatorname{int} \mathbb{R}^Q_+$ and $C(\varepsilon) := \varepsilon\overline{\lambda} + \mathbb{R}^Q_+$ for $\varepsilon > 0$.

If ε is sufficiently small $C(\varepsilon) \cap B(0,1)$ is nonempty. So both Y and $C(\varepsilon) \cap B$ are nonempty, convex compact.

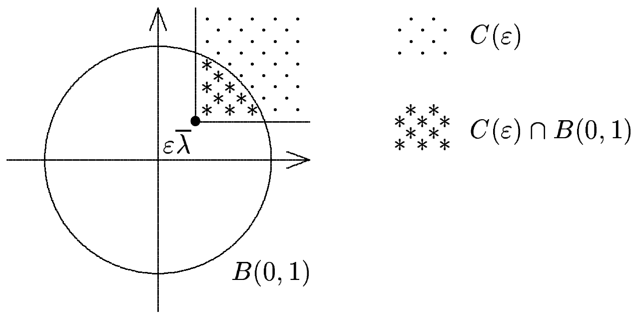


Figure 3.4: Illustration of the Proof of Theorem 3.9

Applying the Sion-Kakutani minimax Theorem 3.10 we get the existence of $y(\varepsilon) \in Y$ and $\lambda(\varepsilon) \in C(\varepsilon) \cap B(0,1)$ such that

$$\langle \lambda, y(\varepsilon)\rangle \leq \langle \lambda(\varepsilon), y(\varepsilon)\rangle \leq \langle \mu(\varepsilon), y\rangle \quad \forall\, y \in Y \quad \forall\, \lambda \in C(\varepsilon) \cap B(0,1) \tag{3.17}$$

Theorem 3.10 (Minimax Theorem of Sion-Kakutani).

Let $C, D \subset \mathbb{R}^n$ be nonempty compact convex sets and $\Phi : C \times D \to \mathbb{R}$ be a continuous mapping s.t. $\Phi(\cdot, y)$ is convex $\forall\, y \in D$ and $\Phi(x, \cdot)$ is concave $\forall\, x \in C$. Then

$$\max_{y\in D} \min_{x\in C} \Phi(x,y) = \min_{x\in C} \max_{y\in D} \Phi(x,y) \tag{3.18}$$

Proof: See [SW70, p.232].

□

We use $C = C(\varepsilon) \cap B(0,1)$, $D = Y$, and $\Phi = \langle \cdot, \cdot \rangle$.

From (3.17) using $0 \in Y \implies \langle \lambda, y(\varepsilon)\rangle \leq 0 \quad \forall\, \lambda \in C(\varepsilon) \cap B$

Y is compact $\implies \exists$ sequence $\varepsilon^k \to 0$ s.t. $\{y^k\} := \{y(\varepsilon^k)\} \to \overline{y} \in Y$ if $k \to \infty$.

For any $\lambda \in \operatorname{int} \mathbb{R}^Q_+ \cap B(0,1)$ $\exists\, \overline{\varepsilon} > 0$ s.t.

$$\lambda \in C(\varepsilon) \cap B(0,1) \quad \forall\, \varepsilon \leq \overline{\varepsilon} \implies \langle \lambda, y^k\rangle \leq 0$$

for k large enough.

Therefore $\langle \lambda, \overline{y}\rangle \leq 0 \quad \forall\, \lambda \in \operatorname{int} \mathbb{R}^Q_+ \implies \overline{y} \in -\mathbb{R}^Q_+$.

So $\overline{y} \leq 0$ but since $\hat{y} = 0 \in Y_{\text{eff}}$ we get $\overline{y} = 0$.

Now we show $\overline{y} = \hat{y} = 0 \in \operatorname{cl} S(Y)$.

So let $\lambda^k := \frac{\lambda(\varepsilon^k)}{\|\lambda(\varepsilon^k)\|} \in \operatorname{int} \mathbb{R}^Q_+ \cap \delta B(0,1)$ where $\lambda(\varepsilon^k)$ is the λ associated with ε^k and $y(\varepsilon^k)$ to satisfy (3.17).

Therefore we have $\langle \lambda^k, y(\varepsilon^k)\rangle \leq \langle \lambda^k, y\rangle \quad \forall\ y \in Y$ i.e. $y^k = y(\varepsilon^k) \subset \mathrm{Opt}(\lambda^k, Y) \subset S(Y)$ since $\overline{y} = \lim y^k$ consequently $\hat{y} = \overline{y} \in \mathrm{cl}\, S(Y)$.

Case 2: Y is closed convex (not necessarily compact)

Again let $\hat{y} = 0 \in Y_{\text{eff}}$.

$Y \cap B(0,1)$ is nonempty, convex compact, $0 \in (Y \cap B(0,1))_{\text{eff}}$.

Case 1 $\Longrightarrow\ \exists\ \{\lambda^k\} \subset \mathrm{int}\, \mathbb{R}^Q_+,\ \|\lambda^k\| = 1$ and $y^k \in \mathrm{Opt}(\lambda^k, Y \cap B),\ y^k \to 0$.

We show that $y^k \in \mathrm{Opt}(\lambda^k, Y)$, which completes the proof.

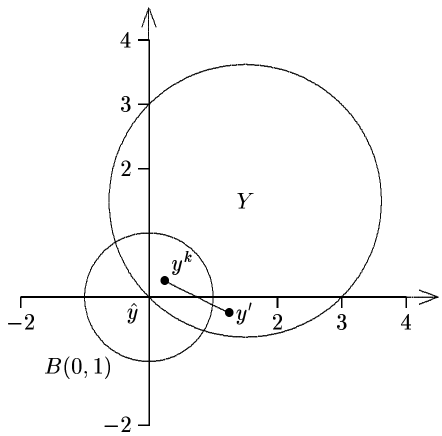


Figure 3.5: Illustration of the Proof of Theorem 3.9

Note that for k large enough $y^k \in \mathrm{int}\, B$ (since $y^k \to 0$) and suppose $\exists\, y' \in Y \quad \langle \lambda^k, y'\rangle < \langle \lambda^k, y^k\rangle$

$\Longrightarrow\ \alpha y' + (1-\alpha) y^k \in Y \cap B$ for sufficiently small α

$\Longrightarrow\ \langle \lambda^k, \alpha y' + (1-\alpha)y^k\rangle = \alpha\langle \lambda^k, y'\rangle + (1-\alpha)\langle \lambda^k, y^k\rangle < \langle \lambda^k, y^k\rangle$ contradicting $y^k \in \mathrm{Opt}(\lambda^k, Y)$.

□

Example 3.2. [ABB53]

The inclusion $\mathrm{cl}\, Y_{\text{p-eff}} \subset Y_{\text{eff}}$ is not always satisfied.

$Y' = \{(y_1, y_2, y_3) : (y_1 - 1)^2 + (y_2 - 1)^2 = 1,\ y_1 \leq 1,\ y_2 \leq 1,\ y_3 = 1\}$

$Y = \mathrm{conv}(Y' \cup \{(1,0,0)\})$

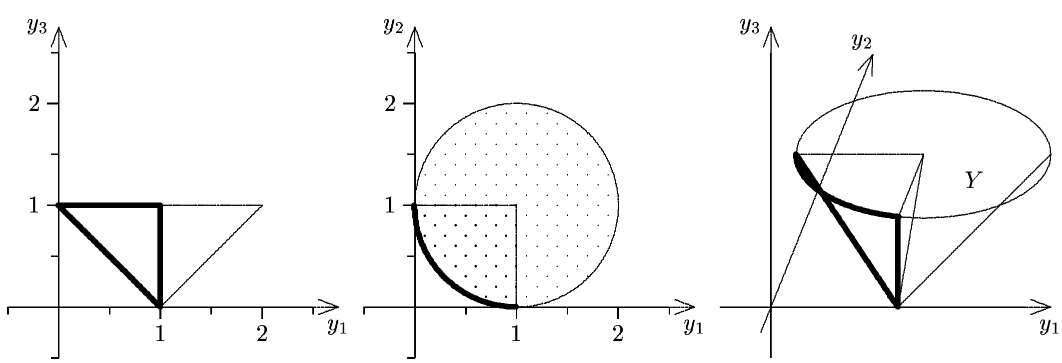


Figure 3.6: The Set Y in Example 3.2

Y is closed, convex $\quad \hat{y} = (1,0,1) \notin Y_{\text{eff}}$ because $(1,0,0) < \hat{y}$. From Theorem 3.8 $Y_{\text{p-eff}} = S(Y)$.

We show that for all $\overline{y} \in Y'$ with $\overline{y}_1 < 1,\ \overline{y}_2 < 1,\ \overline{y} \in Y_{\text{p-eff}}$.

Let $\quad \overline{y} = (1-\cos\theta, 1-\sin\theta, 1) \quad$ for $\quad 0 < \theta < \frac{\pi}{2}$

$\mu = (1-\alpha)(\cos\theta, \sin\theta, 0) + \alpha(0,0,1) \quad 0 < \alpha < 1 \quad$ so $\mu \in \operatorname{int} \mathbb{R}^Q_+$

Let's look at $\langle \mu, y - \overline{y} \rangle \qquad y = (1-\cos\theta', 1-\sin\theta', 1),\ 0 \le \theta' \le \frac{\pi}{2}$

$$\begin{aligned} \langle \mu, y - \overline{y} \rangle &= (1-\alpha)\left[\cos\theta(\cos\theta - \cos\theta') + \sin\theta(\sin\theta - \sin\theta')\right] \\ &= (1-\alpha)(1 - (\cos\theta\cos\theta' + \sin\theta\sin\theta')) \qquad (3.19) \\ &= (1-\alpha)(1-\cos(\theta-\theta')) \ge 0 \end{aligned}$$

$$\begin{aligned} \langle \mu, (1,0,0) - \overline{y} \rangle &= (1-\alpha)\left[\cos^2\theta - \sin\theta(1-\sin\theta)\right] - \alpha \\ &= (1-\alpha)(1-\sin\theta) - \alpha \quad > 0 \quad \text{for small } \alpha \qquad (3.20) \end{aligned}$$

So by convex combinations of (3.19) and (3.20) we get $\langle \mu, y - \overline{y} \rangle \ge 0 \quad \forall\, y \in Y$ and $\overline{y} \in S(Y)$. Furthermore, for $\theta \to 0$ we get $\overline{y} \to \hat{y}$ which is therefore in $\operatorname{cl} S(Y)$.

To conclude this section, we present some results of the Kuhn-Tucker type.

We already have Theorem 2.29, necessary conditions for KT proper efficiency and Theorem 2.31 (Geoffrion $\implies$ KT under the constraint qualification).

An immediate consequence:

Corollary 3.11. *If $\hat{x}$ is properly Pareto optimal in Geoffrion's sense and the KT constraint qualification is satisfied at $\hat{x}$ then the condition of Theorem 2.29 holds.*

For the missing link in the relations of proper Pareto optimality definitions, we use the single objective Kuhn-Tucker optimality conditions.

Theorem 3.12. *Let $f, g_j : \mathbb{R}^n \to \mathbb{R}$ be convex, continuously differentiable functions and suppose that $\exists\ \hat{x} \in X,\ \hat{\lambda} \ge 0$ such that*

$$\nabla f(\hat{x}) + \sum_{j=1}^{m} \hat{\lambda}_j \nabla g_j(\hat{x}) = 0 \qquad (3.21)$$

$$\sum_{j=1}^{m} \hat{\lambda}_j g_j(\hat{x}) = 0 \qquad (3.22)$$

then $\hat{x}$ is a locally, thus globally, optimal solution of $\min_{x \in X} f(x)$, where $X = \{x \in \mathbb{R}^n : g_j(x) \le 0,\ j = 1, \ldots, m\}$.

Proof: See e.g. [BSS93].

□

Using $f = \sum_{i=1}^{Q} \hat{\mu}_i f_i(\hat{x})$ we have the following result:

Theorem 3.13. *Let $X = \{x \in \mathbb{R}^n \ : \ g_j(x) \le 0 \quad \forall\, j = 1, \dots, m\}$. Assume $f_i, g_j : \mathbb{R}^n \to \mathbb{R}$ are convex, continuously differentiable functions and for $\hat{x} \in X \quad \exists\, \hat{\mu} \gg 0,\ \hat{\lambda} \ge 0$ s.t.*

$$\sum_{i=1}^{Q} \hat{\mu}_i \nabla f_i(\hat{x}) + \sum_{j=1}^{m} \hat{\lambda}_j \nabla g_j(\hat{x}) = 0 \tag{3.23}$$

$$\sum_{j=1}^{m} \hat{\lambda}_j g_j(\hat{x}) = 0 \tag{3.24}$$

then $\hat{x}$ is properly Pareto optimal in the sense of Geoffrion.

Proof: Let $f = \sum_{i=1}^{Q} \hat{\mu}_i \nabla f_i(x)$, which is convex. By Theorem 3.12 $\hat{x}$ is an optimal solution of $\min_{x \in X} \sum_{i=1}^{Q} \hat{\mu}_i f_i(x)$. Since $\hat{\mu}_i > 0 \quad \forall\, i = 1, \dots, Q$ by Theorem 2.23 $\hat{x}$ is properly Pareto optimal in the sense of Geoffrion.

□

Corollary 3.14. *Let $f_i, g_j : \mathbb{R}^n \to \mathbb{R}$ be convex, continuously differentiable functions and suppose $\hat{x}$ is properly Pareto optimal in the sense of Kuhn-Tucker. Then $\hat{x}$ is properly Pareto optimal in the sense of Geoffrion.*

Proof: Theorem 2.29 and Theorem 3.13.

□

Finally, we can discuss KT conditions for weak Pareto optimality.

Theorem 3.15. *Suppose that the KT constraint qualification is satisfied at $\hat{x} \in X$. Then if $\hat{x}$ is weakly Pareto optimal there exists $\hat{\mu} \in \mathbb{R}^Q$ and $\hat{\lambda} \in \mathbb{R}^m$ s.t.*

i) $$\sum_{i=1}^{Q} \hat{\mu}_i \nabla f_i(\hat{x}) + \sum_{j=1}^{m} \hat{\lambda}_j g_j(\hat{x}) = 0 \tag{3.25}$$

ii) $$\sum_{j=1}^{m} \hat{\lambda}_j g_j(\hat{x}) = 0 \tag{3.26}$$

iii) $$\hat{\mu} > 0,\ \hat{\lambda} \ge 0 \tag{3.27}$$

Proof: Let $\hat{x} \in X_{\text{w-Par}}$. We show $\nexists\, h$ such that

$$\begin{aligned} \langle \nabla f_i(\hat{x}), h \rangle &< 0 \qquad \forall\, i = 1, \dots, Q \\ \langle \nabla g_j(\hat{x}), h \rangle &< 0 \qquad \forall\, j \in J(\hat{x}) := \{j : g_j(\hat{x}) = 0\} \end{aligned} \tag{3.28}$$

Suppose that such an $h \in \mathbb{R}^n$ exists.
From the KT constraint qualification $\exists$ continuously differentiable $\theta : [0, \bar{t}] \to \mathbb{R}^n$ such that $\theta(0) = \hat{x},\ g(\theta(t)) \le 0,\ \theta'(0) = \alpha h,\ \alpha > 0$. Since $f_i(\theta(t)) = f_i(\hat{x}) + t\langle \nabla f_i(\hat{x}), \alpha h \rangle + \theta(t)$ and using $\langle \nabla f_i(\hat{x}), h \rangle < 0$
$\Longrightarrow \quad f_i(\theta(t)) < f_i(\hat{x}) \quad \forall\, i = 1, \dots, Q$ and for t sufficiently small, which contradicts $\hat{x} \in X_{\text{w-Par}}$.
It remains to show that (3.28) implies the conditions. To that end we use Motzkin's Theorem of the alternative.

Theorem 3.16. *Let B, C, D be $Q \times n$, $k \times n$ and $0 \times n$ matrices, respectively. Then either*

$$a) \quad Bx \ll 0, \quad Cx \leq 0, \quad Dx = 0 \text{ has a solution } x \in \mathbb{R}^n \tag{3.29}$$

$$\text{or} \quad b) \quad B^t y^1 + C^t y^2 + D^t y^3 = 0 \text{ has a solution } y^1 > 0,\ y^2 \geq 0 \tag{3.30}$$

Proof: See [Man69, p.28].

□

Therefore, using $B = \begin{pmatrix} \nabla f_1(\hat{x}) \\ \vdots \\ \nabla f_Q(\hat{x}) \end{pmatrix}$, $C = (\nabla g_j(\hat{x}))_{j \in J(\hat{x})}$, $D = 0$, $\hat{x} = h$, $y^1 = \hat{\mu}$, $y^2 = \hat{\lambda}$, $y^3 = 0$ in the proof of Theorem 3.15 completes the proof.

□

Corollary 3.17. *Under the conditions of Theorem 3.15 and the additional assumption that all functions are convex i), ii), iii) in 3.15 are sufficient for $\hat{x}$ to be weakly Pareto optimal.*

Proof: By Theorem 3.15 (3.25) - (3.27) imply that $\hat{x}$ is optimal for $\min_{x \in X} \sum \hat{\mu}_i f_i(x)$. Since $\hat{\mu} \in \mathbb{R}^Q_+ \setminus \{0\}$ this implies $f(\hat{x}) \in \text{Opt}(\hat{\mu}, f(X))$. By Theorem 3.5 we get $f(\hat{x}) \in S_0(Y) \subset Y_{\text{w-eff}} \implies \hat{x} \in X_{\text{w-Par}}$.

□

We close the section by examples showing that Geoffrion's and Kuhn-Tucker's definitions are something different.

Example 3.3 (Kuhn-Tucker, but not Geoffrion).

$X = \{x \in \mathbb{R} \,:\, x \geq 0\}$ $\quad f_1(x) = -x^2$ $\quad f_2(x) = x^3$ $\quad f_2 = (-f_1)^{\frac{3}{2}}$ $\quad \hat{x} = 0$

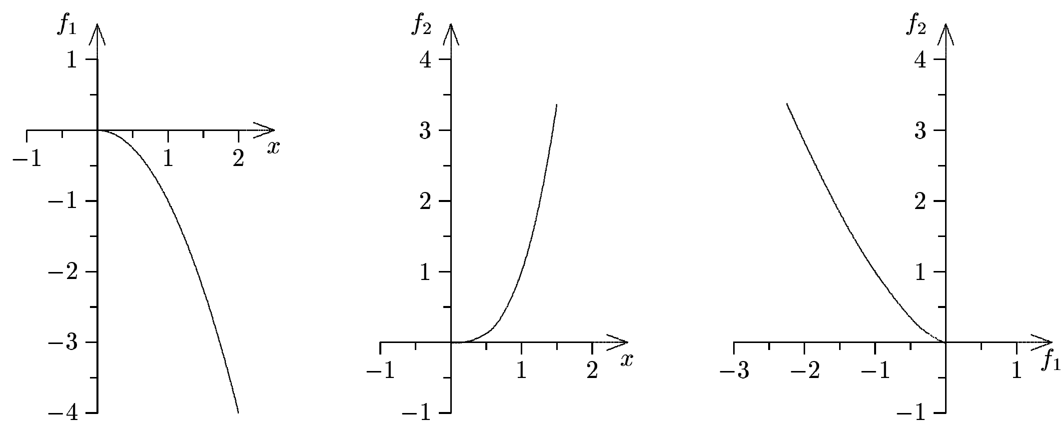


Figure 3.7: Objective Functions of Example 3.3

Kuhn-Tucker:

$$\begin{array}{rrcl rcl} & \nabla f_1 & = & -2x & \nabla f_2(\hat{x}) & = & 0 \\ & \nabla f_2 & = & 3x^2 & \nabla f_2(\hat{x}) & = & 0 \\ g(x) = -x \leq 0 & \nabla g & = & -1 & \nabla g(\hat{x}) & = & -1 \end{array}$$

Choose $\hat{\mu}_1 = \hat{\mu}_2 = 1, \ \hat{\lambda} = 0$ which satisfies the conditions.

Not Geoffrion:

Let $\varepsilon > 0$

$$\frac{f_1(\hat{x}) - f_1(\varepsilon)}{f_2(\varepsilon) - f_2(\hat{x})} = \frac{0 + \varepsilon^2}{\varepsilon^3 - 0} = \frac{1}{\varepsilon} \xrightarrow{\varepsilon \to 0} \infty$$

Example 3.4 (Geoffrion, but not Kuhn-Tucker).

Exercise 25.

3.3 Connectedness of Y_{eff} and X_{Par}

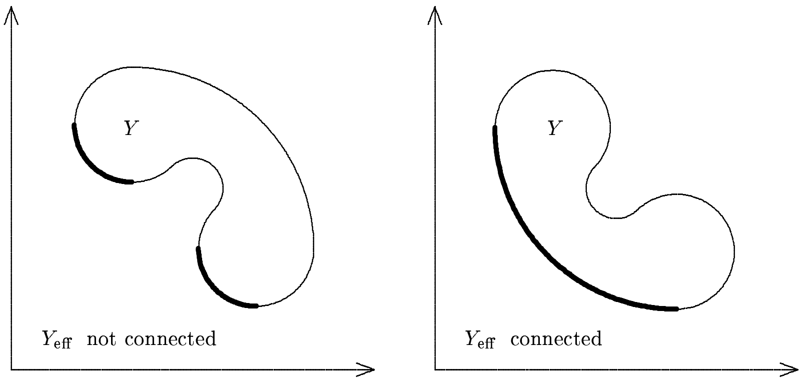


Figure 3.8: Connectedness of Y_{eff}

Connectedness is especially important for finding a best compromise solution: The whole sets $Y_{\text{eff}}/X_{\text{Par}}$ can be searched by moving slightly from a given point.

Definition 3.1. $A \subset \mathbb{R}^Q$ is called **not connected** if it can be written as $A = A_1 \cup A_2$, with $\text{cl}\, A_1 \cap A_2 = A_1 \cap \text{cl}\, A_2 = \emptyset$, equivalently A is **not connected** if $\exists$ open sets O_1, O_2 s.t. $A \subset O_1 \cup O_2, \ A \cap O_1 \neq \emptyset, \ A \cap O_2 \neq \emptyset, \ A \cap O_1 \cap O_2 = \emptyset$.

Otherwise, A is called **connected**.

We use some facts about connected sets stated below.

Lemma 3.18.

a) If A is connected and $A \subset B \subset \text{cl}\, A$ then B is connected.

b) If $\{A_i \ : \ i \in I\}$ is a family of connected sets with $\bigcap_{i \in I} A_i \neq \emptyset$ then $\bigcup_{i \in I} A_i$ is connected.

Now we consider $\text{Opt}(\lambda, Y)$ and $S(Y)$.

From Theorem 3.9 we know $S(Y) \subset Y_{\text{eff}} \subset \text{cl}\, S(Y)$. We prove connectedness of $S(Y)$ in the case that Y is compact.

Proposition 3.19. *If Y is compact convex then $S(Y)$ is connected.*

Proof: Suppose $S(Y)$ is not connected.

$\Longrightarrow \exists$ open sets Y_1, Y_2 s.t. $Y_i \cap S(Y) \neq \emptyset, \ i = 1, 2, \quad Y_1 \cap Y_2 \cap S(Y) = \emptyset, \quad S(Y) \subset Y_1 \cup Y_2$.

Let $M_i := \{\lambda \in \text{int}\, \mathbb{R}^Q_+ \ : \ Opt(\lambda, Y) \cap Y_i \neq \emptyset\}, \ i = 1, 2$.

We know that $\text{Opt}(\lambda, Y)$ is connected (it is convex and every convex set is connected).

$\Longrightarrow\ M_i = \{\lambda \in \text{int}\,\mathbb{R}^Q_+ \ :\ \text{Opt}(\lambda, Y) \subset Y_i\},\ i = 1, 2$

$\Longrightarrow\ M_1 \cap M_2 = \emptyset.$

But since $Y_i \cap S(Y) \neq \emptyset$ we also have $M_i \cap \text{int}\,\mathbb{R}^Q_+ \neq \emptyset,\ i = 1, 2$

and from $S(Y) \subset Y_1 \cup Y_2$ follows $\text{int}\,\mathbb{R}^Q_+ \subset M_1 \cup M_2$ (indeed, it's equality)

By Lemma 3.20 M_i are open $\Longrightarrow$ $\text{int}\,\mathbb{R}^Q_+$ is not connected. ↯ Contradiction

□

Lemma 3.20. $M_i = \{\lambda \in \text{int}\,\mathbb{R}^Q_+ \ :\ S(Y) \subset Y_i\}$ *in the proof of Proposition 3.19 are open.*

Proof: We will show it for M_1.

If M_1 is not open $\Longrightarrow\ \exists\ \hat{\lambda} \in M_1$ and $\lambda^k \in \text{int}\,\mathbb{R}^Q_+ \setminus M_1 = M_2,\ k \geq 1$ s.t. $\lambda^k \to \hat{\lambda}$.

Let $y^k \in \text{Opt}(\lambda^k, Y),\ k \geq 1$.

So we can assume (taking a subsequence if necessary) that $y^k \to \hat{y} \in Y,\ \hat{y} \in \text{Opt}(\hat{\lambda}, Y)$ [otherwise $\exists\ y' \in Y$ s.t $\langle\hat{\lambda}, y'\rangle < \langle\hat{\lambda}, \hat{y}\rangle$ so by continuity we would have $\langle\hat{\lambda}, y'\rangle < \langle\lambda^k, y^k\rangle$ for sufficiently large k, contradicting $y^k \in \text{Opt}(\lambda^k, Y)$]

Now we have $y^k \in \text{Opt}(\lambda^k, Y) \subset Y_2 \cap S(Y)$ and $Y_1 \cap Y_2 \cap S(Y) = \emptyset$, so $y^k \in Y_1^c\ \ \forall\ k \geq 1$.

Since Y_1^c is closed $\Longrightarrow\ \hat{y} = \lim y^k \in Y_1^c$ or $\hat{y} \notin Y_1$ contradicting $\hat{\lambda} \in M_1$.

□

Theorem 3.21 (Naccache, 1978, [Nac78]).

If Y is closed, convex and $\mathbb{R}^Q_+$-compact then Y_{eff} is connected.

Proof: Choose $d \in \text{int}\,\mathbb{R}^Q_+$ and define $y(\alpha) = \alpha \cdot d,\ \alpha \in \mathbb{R}$.

Claim: $\forall\ y \in \mathbb{R}^Q\ \ \exists\ \alpha > 0$ s.t. $y \in y(\alpha) - \mathbb{R}^Q_+$. If this is not true we have two convex sets $\{y - \alpha d\ :\ \alpha > 0\}$ and $-\mathbb{R}^Q_+$ which can be separated (Theorem 3.3).

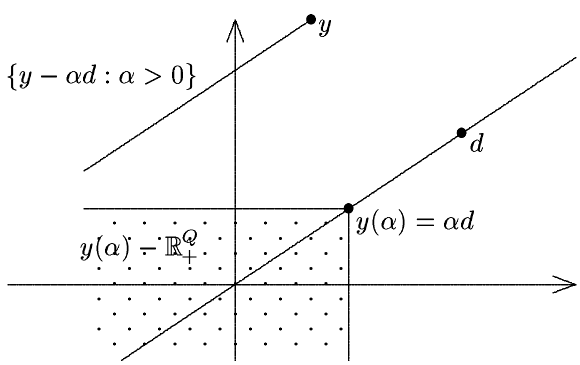


Figure 3.9: Illustration of the Proof of Theorem 3.21

$\Longrightarrow\ \exists\ \lambda \in \mathbb{R}^Q \setminus \{0\}$ s.t.

$$\langle\lambda, y - \alpha d\rangle \geq 0 \quad \forall\ \alpha > 0$$
$$\langle\lambda, -d'\rangle \leq 0 \quad \forall\ d' \in \mathbb{R}^Q_+$$

$\Longrightarrow\ \langle\lambda, d'\rangle \geq 0\ \ \forall\ d' \in \mathbb{R}^Q_+$, especially $\langle\lambda, d\rangle > 0$ because $d \in \text{int}\,\mathbb{R}^Q_+$.

Then $\langle\lambda, y - \alpha d\rangle < 0$ for α sufficiently large, a contradiction to the first inequality.

$\Longrightarrow$ Especially for $y \in Y_{\text{eff}}$ we can choose $\hat{\alpha} > 0$ s.t. $y \in y(\hat{\alpha}) - \mathbb{R}^Q_+$.

$\Longrightarrow$ $(y(\hat{\alpha}) - \mathbb{R}^Q_+) \cap Y_{\text{eff}} \neq \emptyset$.

Denote $E(\alpha) := [(y(\alpha) - D) \cap Y]_{\text{eff}}$.

The claim above implies that $Y_{\text{eff}} = \bigcup_{\alpha \geq \hat{\alpha}} E(\alpha)$.

Because $(y(\alpha) - D) \cap Y$ is compact, using Theorem 3.9 for $E(\alpha)$, Proposition 3.19 and Lemma 3.18 a) we get that $E(\alpha)$ is connected.

But noting that $E(\alpha) \supset E(\hat{\alpha})$ for $\alpha > \hat{\alpha}$, i.e. $\bigcap_{\alpha \geq \hat{\alpha}} E(\alpha) = E(\hat{\alpha}) \neq \emptyset$ we have expressed Y_{eff} as union of a family of connected sets with nonempty intersection. Lemma 3.18 b) $\Longrightarrow$ Y_{eff} is connected.

□

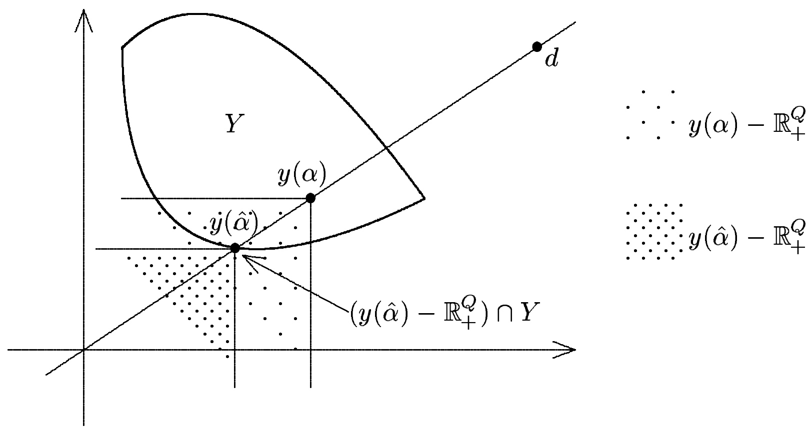


Figure 3.10: Illustration of the Proof of Theorem 3.21

Let us now turn to $X_{\text{w-Par}}$. We shall show that $X_{\text{w-Par}}$ is connected under convexity assumptions.

Let $X \subset \mathbb{R}^n$ be convex and compact and $f_i : \mathbb{R}^n \to \mathbb{R}$ be convex. We will use Theorem 3.6 ($Y_{\text{w-eff}} = S_0(Y)$) and the following fact:

Lemma 3.22. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex on the closed convex X. Then the set $\{x \in X : f(x) = \inf_{x \in X} f(x)\}$ is closed and convex.*

We also need a theorem providing a result on connectedness of preimages of sets.

Theorem 3.23. *Let $V \subset \mathbb{R}^n$, $W \subset \mathbb{R}^Q$ s.t. V is compact and W is connected, and let $g : V \times W \to \mathbb{R}$ be continuous. Denote by $X(w) = \operatorname{argmin}\{g(v,w) : v \in V\}$. If $X(w)$ is connected $\forall\, w \in W$ then $\bigcup_{w \in W} X(w)$ is connected.*

Proof: See [War83].

□

Theorem 3.24. *Let X be a compact convex set and assume $f_i : \mathbb{R}^n \to \mathbb{R}$ are convex. Then $X_{\text{w-Par}}$ is connected.*

Proof: Since f_i are continuous and X is compact, $Y = f(X)$ is compact.

Using Theorem 3.6 we have $Y_{\text{w-eff}} = S_0(Y)$, in terms of f this means

$$X_{\text{w-Par}} = \bigcup_{\lambda \in \mathbb{R}^Q_+} \{x^* : \sum_{i=1}^{Q} \lambda_i f_i(x^*) \leq \sum_{i=1}^{Q} \lambda_i f_i(x) \quad \forall\, x \in X\}$$
$$=: \bigcup_{\lambda \in \mathbb{R}^Q_+} X(\lambda)$$

Noting that $\langle \lambda, f(x) \rangle$ is continuous on $\mathbb{R}^Q_+ \times X$, $\mathbb{R}^Q_+$ is connected, X is compact, and that by Lemma 3.22 $X(\lambda)$ is nonempty and convex (hence connected) we can apply Theorem 3.23 to get: $X_{\text{w-Par}}$ is connected.

□

Remark. The proof works in the same way to see that X_{Par} is connected under the same assumptions, if we take into account Theorem 3.9.

Obviously, Theorem 3.24 has consequences for $Y_{\text{w-eff}}/Y_{\text{eff}}$.

Corollary 3.25. *If X is convex compact and $f_i : \mathbb{R}^n \to \mathbb{R}$ are convex then Y_{eff} and $Y_{\text{w-eff}}$ are connected.*

Proof: The image of a connected set under a continuous mapping is connected.

□

3.4 Exercises to Chapter 3

21. Prove that if Y is closed then $\operatorname{cl} S(Y) \subset S_0(Y)$.

    Hint: Choose sequences λ_k, y^k s.t. $y^k \in \operatorname{Opt}(\lambda_k, Y)$ and show that $\lambda_k \to \hat{\lambda}$ and $y^k \to \hat{y}$ with $\hat{y} \in \operatorname{Opt}(\hat{\lambda}, Y)$, $\hat{\lambda} > 0$.

22. Prove Proposition 3.4, i.e. show that if y^* is the unique member of $\operatorname{Opt}(\lambda, Y)$ for some $\lambda \in \mathbb{R}^Q_+ \setminus \{0\}$ then $y^* \in Y_{\text{eff}}$.

23. Give one example of a set for each of the following situations:

    i) $S(Y) \subset Y_{\text{eff}} \subset S_0(Y)$ with both inclusions strict.

    ii) Denote by $S_0'(Y) = \left\{y' \in Y : y' \text{ is the unique member of } \operatorname{Opt}(\lambda, Y),\ \lambda \in \mathbb{R}^Q_+ \setminus \{0\}\right\}$.
    $S(Y) \cup S_0'(Y) = Y_{\text{eff}} = S_0(Y)$

24. Let $Y = \left\{(y_1, y_2) : y_1^2 + y_2^2 \leq 1\right\}$ and $K = \left\{(y_1, y_2) = y_2 \leq \frac{1}{2} y_1\right\}$.

    a) Show that $\hat{y} = (-1, 0)$ is properly efficient in Benson's sense,
    i.e. $(\operatorname{cl}(\operatorname{cone}(Y + K - \hat{y}))) \cap (-K) = \{0\}$.

    b) Show that $\hat{y} \in \operatorname{Opt}(\lambda, Y)$ for a $\lambda \notin \operatorname{int} \mathbb{R}^Q_+$ and check that this $\lambda \in K^{so} = \{\mu : \langle \mu, d \rangle > 0 \quad \forall\, d \in K\}$.

Conclusion: Proper Pareto optimality is related to scalarisation with vectors in K^{so}.

25. Let

$$\begin{aligned} X &= \{(x_1, x_2) \in \mathbb{R}^2 : -x_1 \le 0,\ -x_2 \le 0,\ (x_1-1)^3 + x_2 \le 0\} \\ f_1(x) &= -3x_1 - 2x_2 + 3 \\ f_2(x) &= -x_1 - 3x_2 + 1 \end{aligned}$$

Graph X and $Y = f(X)$. Show that $\hat{x} = (1, 0)$ is properly Pareto optimal in Geoffrion's sense, but not in Kuhn-Tucker's. (You may equivalently use Benson instead of Geoffrion.)

26. Let $K \subset \mathbb{R}^Q$ be a cone.

The polar cone K° of K is defined as follows:

$$K^\circ := \{x \in \mathbb{R}^Q : \langle x, d\rangle \ge 0 \quad \forall\, d \in K\}.$$

Prove the following:

a) K° is a closed convex cone containing 0.

b) $K \subset (K^\circ)^\circ =: K^{\circ\circ}$

c) $K_1 \subset K_2 \Rightarrow K_2^\circ \subset K_1^\circ$

d) $K^\circ = (K^{\circ\circ})^\circ$

27. Comparing scalarizations with respect to polar cones and K-efficiency. Let K be a convex pointed cone and $\lambda \in K^\circ$.

$$\mathrm{Opt}_K(\lambda, Y) = \left\{ y^* \in Y : \langle \lambda, y^* \rangle = \inf_{y \in Y} \langle \lambda, y \rangle \right\}.$$

a) Show that $S_{K^\circ}(Y) := \bigcup_{\lambda \in K^\circ \setminus \{0\}} \mathrm{Opt}(\lambda, Y) \subset Y_{\text{Kw-eff}}$

where $y^* \in Y_{\text{Kw-eff}}$ if $(Y + \operatorname{int} K - y^*) \cap (-\operatorname{int} K) = \emptyset$

b) Let $K^{so} := \{x \in \mathbb{R}^Q : \langle x, d\rangle > 0 \quad \forall\, d \in K \setminus \{0\}\}$.

Show $S_{K^{so}}(Y) := \bigcup_{\lambda \in K^{so}} \mathrm{Opt}(\lambda, Y) \subset Y_{\text{K-eff}}$.

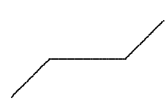
Hint: Look at the proofs of Theorems 3.5 and 3.2, respectively.

28. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called quasi-convex if $f(\alpha x + (1-\alpha)y) \le \max\{f(x), f(y)\}$ $\forall\, \alpha \in (0, 1)$.

(It is known that f is quasi-convex iff $L_{\le}(f(x))$ is convex for all x.)

Give an Example of a multi-criteria optimization problem with $X \subset \mathbb{R}$ convex, $f_i : \mathbb{R} \to \mathbb{R}$ quasi-convex s.t. X_{Par} is not connected.

Hint: Monotone increasing/decreasing functions are quasi-convex, especially those that look like



Chapter 4

Methodology

In this chapter we will discuss methods to find Pareto optimal solutions. We have already investigated in detail the most popular, namely weighted sum scalarization. First we discuss lower and upper bounds for efficient points.

4.1 Bounds on the Efficient Set

We assume that X_{Par} and Y_{eff} are nonempty.

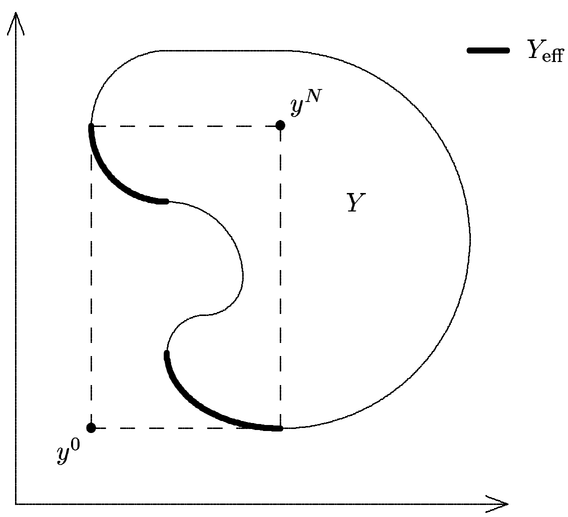


Figure 4.1: Y, Y_{eff}, y^0 and y^N

We are interested in the range of possible values of the objective functions.

A lower bound is given by the optima of each objective individually.

Let

$$y_i^0 := \inf_{x \in X} f_i(x) = \inf_{y \in Y} y_i \tag{4.1}$$

Then $y^0 = (y_1^0, \dots, y_Q^0)$ is called the **ideal point** of the problem $\min_{x \in X}(f_1(x), \dots, f_Q(x))$.

As a consequence we have $\quad y_i^0 \leq y_i \quad \forall\, y \in Y_{\text{eff}}$.

An **upper bound** is defined as follows

$$y_i^N := \sup_{x \in X_{\text{Par}}} f_i(x) = \sup_{y \in Y_{\text{eff}}} y_i \tag{4.2}$$

$y^N = (y_1^N, \dots, y_Q^N)$ is called the **Nadir point** of the multicriteria optimization problem.

Obviously $y_i \le y_i^N \quad \forall\, y \in Y_{\text{eff}}$.

The ideal point is found by solving Q single objective optimization problems.

But determination of y^N would require knowledge of X_{Par}. There is no known method to determine y^N for a general MOP.

An estimation is as follows. Assume that minimizers of f_i over X exist. Then

① Determine $x^i,\ i = 1, \dots, Q$ s.t. $f_i(x^i) = \min_{x \in X} f_i(x)$.

② Make the following payoff table

	x^1	$\cdots$	$\cdots$	x^Q
f_1	y_1^0	$f_1(x^2)$	$\cdots$	$f_1(x^Q)$
$\vdots$	$f_2(x^1)$	$\ddots$	$\ddots$	$\vdots$
$\vdots$	$\vdots$	$\ddots$	$\ddots$	$f_{Q-1}(x^Q)$
f_Q	$f_Q(x^1)$	$\cdots$	$f_Q(x^{Q-1})$	y_Q^0

the ideal point y^0

③ Let $\tilde{y}_q^N := \max_{i=1,\dots,Q} f_Q(x^i)$, the largest element in each row is used as estimate for y_q^N.

The problem is that $\tilde{y}^N$ may over- or under-estimate y^N.

Example 4.1 ([KSS97]).

$$\begin{array}{llllllll}
\min & & -11x_2 & -11x_3 & -12x_4 & -9x_5 & -9x_6 & +9x_7 \\
\min & -11x_1 & & -11x_3 & -9x_4 & -12x_5 & -9x_6 & +9x_7 \\
\min & -11x_1 & -11x_2 & & -9x_4 & -9x_5 & -12x_6 & -12x_7 \\
\text{s.t.} & \multicolumn{7}{l}{x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 1} \\
& x_i \ge 0 & & & & & &
\end{array}$$

The individual minimizers are

for f_1 $\quad x_4^1 = 1,\ x_i^1 = 0,\ i \ne 4$

for f_2 $\quad x_5^2 = 1,\ x_i^2 = 0,\ i \ne 5$

for f_3 $\quad x_6^3 = 1,\ x_i^3 = 0,\ i \ne 6$ or $\overline{x}_7^3 = 1,\ \overline{x}_i^3 = 0,\ i \ne 7$.

Payoff table

	x^1	x^2	x^3	$\overline{x}^3$
f_1	-12	-9	-9	9
f_2	-9	-12	-9	9
f_3	-9	-9	-12	-12

y^0

By solving appropriate weighted sum problems it can be seen that all x where $x_i = 1$ for one $i \in \{1, \dots, 6\}$ and 0 else are (properly) Pareto optimal.

$\overline{x}^3 = (0, \dots, 0, 1)$ is obviously weakly Pareto optimal, as a minimizer of one objective.

So choose $x = (1, 0, \dots, 0) \in X_{\text{Par}} \implies f(x) = (0, -11, -11)$

$x = (0, 1, 0, \dots, 0) \in X_{\text{Par}} \implies f(x) = (-11, 0, -11)$

$x = (0, 0, 1, 0, \dots, 0) \in X_{\text{Par}} \implies f(x) = (-11, -11, 0)$

Therefore $y^N = (0, 0, 0)$ (No Pareto point has positive objective values.)

We observe that

1. With $\overline{x}^3$ we overestimate y_1^N (arbitrarily far: replace $+9$ by $C > 0$ arbitrarily large)

   The reason is: $\overline{x}^3$ is weakly Pareto optimal. If we choose Pareto optimal points to determine x^i, overestimation is impossible.

2. With x^3 we underestimate y_1^N severly (arbitrarily far, if we modify the cost coefficients).

In general, it is difficult to be sure that x^i are Pareto optimal. The only case where y^N can be determined is for $Q = 2$. Here the worst value for y_2 is attained when y_1 is minimal and vice versa.

(1) Solve $\min_{x \in X} f_1(x)$, $\min_{x \in X} f_2(x)$. Denote the optimal objective values by y_1^0, y_2^0.

(2) Solve $\min_{x \in X} f_2(x)$ s.t. $f_1(x) = y_1^0$.
Solve $\min_{x \in X} f_1(x)$ s.t. $f_2(x) = y_2^0$.
Denote the values by y_2^N, y_1^N, respectively.

(3) $y^N = (y_1^N, y_2^N)$ is the Nadir point.

If $Q > 2$ we don't know, which objectives to fix in Step (2).

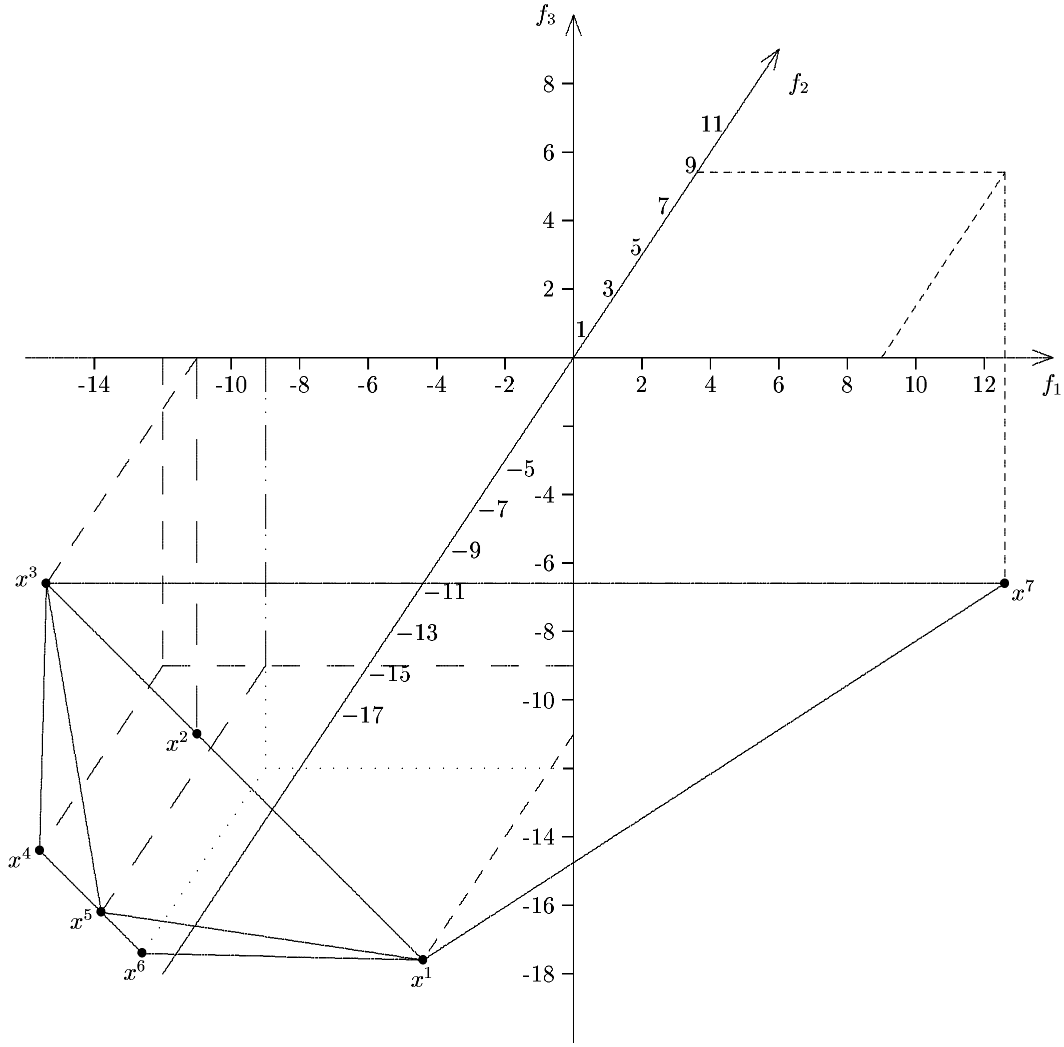


Figure 4.2: Feasible Set of Example 4.1

4.2 The ε-Constraint Method

In this method, only one objective is minimized, whereas constraints are put on the others. It was introduced by Haimes et.al. in 1971, [HLW71].

$$\min_{x \in X} (f_1(x), \dots, f_Q(x)) \tag{4.3}$$

is replaced by

$$\begin{aligned} &\min_{x \in X} f_k(x) \\ &\text{s.t.} \quad f_j(x) \le \varepsilon_j \quad \forall\, j = 1, \dots, Q, \quad j \ne k \end{aligned} \qquad P_k(\varepsilon)$$

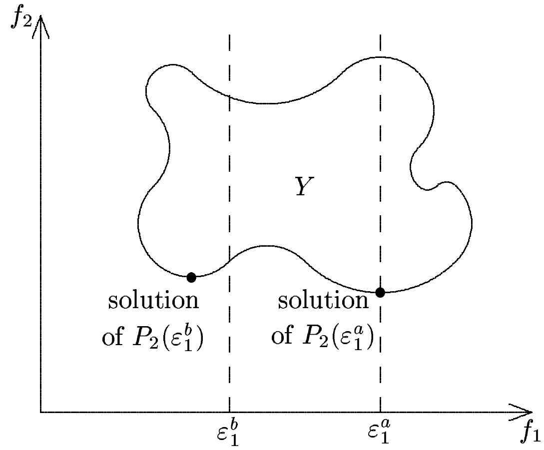


Figure 4.3: Solution of $P_k(\varepsilon)$

Proposition 4.1. *Let $\hat{x}$ be an optimal solution of $P_k(\varepsilon)$ for some k then $\hat{x}$ is weakly Pareto optimal.*

Proof: Assume $\hat{x} \notin X_{\text{w-Par}} \implies \exists\, x \in X$ s.t. $f_i(x) < f_i(\hat{x}) \quad \forall\, i = 1, \dots, Q$, especially $f_k(x) < f_k(\hat{x})$. Since $f_i(x) < f_i(\hat{x}) \leq \varepsilon_i$, $i \neq k$, x is also feasible for $P_k(\varepsilon)$. We have a contradiction to optimality of $\hat{x}$.

□

If the solution of a $P_k(\varepsilon)$ problem is unique, we get Pareto optimal solutions:

Proposition 4.2. *Let $\hat{x}$ be a unique optimal solution of $P_k(\varepsilon)$ for some k then $\hat{x} \in X_{\text{Par}}$.*

Proof: Suppose there exists an $x \in X$, $f_j(x) \leq \varepsilon_j \quad \forall\, j \neq k$. Because $\hat{x}$ solves $P_k(\varepsilon)$, from $f_k(x) \leq f_k(\hat{x})$ it must follow $f_k(x) = f_k(\hat{x})$, thus from uniqueness, $x = \hat{x}$, therefore $\hat{x} \in X_{\text{Par}}$.

□

In general, Pareto optimality of $\hat{x}$ is related to $\hat{x}$ solving $P_k(\varepsilon)$ for all k.

Theorem 4.3. *$\hat{x} \in X_{\text{Par}}$ if and only if $\exists\, \hat{\varepsilon} \in \mathbb{R}^Q$ s.t. $\hat{x}$ solves $P_k(\hat{\varepsilon})$ for all $k = 1, \dots, Q$.*

Proof:

„$\implies$“ Let $\hat{\varepsilon} = f(\hat{x})$. Assume $\hat{x}$ does not solve $P_k(\hat{\varepsilon})$ for some k.
$\implies \exists\, x \in X$ s.t. $f_k(x) < f_k(\hat{x})$ and $f_j(x) \leq \hat{\varepsilon}_j = f_j(\hat{x}) \quad \forall\, j \neq k$
$\implies \hat{x} \notin X_{\text{Par}}$ ↯ Contradiction

„$\impliedby$“ Suppose $\hat{x} \notin X_{\text{Par}} \implies \exists\, q \in \{1, \dots, Q\}$, $\exists\, x \in X$ s.t. $f_q(x) < f_q(\hat{x})$, $f_j(x) \leq f_j(\hat{x}) \quad j \neq q$.
Since $\hat{x}$ solves $P_q(\hat{\varepsilon})$, in particular $\hat{x}$ is feasible for $P_q(\hat{\varepsilon})$, we have $f_j(x) \leq f_j(\hat{x}) \leq \hat{\varepsilon}_j \; \forall\, j \neq q$.
So x is feasible for $P_q(\hat{\varepsilon})$. Therefore $f_q(x) < f_q(\hat{x})$ contradicts the assumption $\implies \hat{x} \in X_{\text{Par}}$.

□

If we denote by $\mathcal{E}_k := \{\varepsilon : \{x \in X : f_j(x) \leq \varepsilon_j, \ j \neq k\} \neq \emptyset\}$ and $X_k(\varepsilon) := \{x : x$ solves $P_k(\varepsilon)$ for $\varepsilon \in \mathcal{E}_k\}$ we can write (using Theorem 4.3): For each $\varepsilon \in \mathbb{R}^Q$

$$\bigcap_{k=1}^{Q} X_k(\varepsilon) \subset X_{\text{Par}} \subset X_k(\varepsilon) \quad \forall \ k = 1, \dots, Q \tag{4.4}$$

We can relate solutions of $P_k(\varepsilon)$ problems to solutions of weighted sum problems:

Theorem 4.4.

a) *Suppose $\hat{x}$ solves $\min_{x \in X} \sum_{i=1}^{Q} \lambda_i f_i(x)$. If $\lambda_k > 0$ there exists $\hat{\varepsilon}$ s.t. $\hat{x}$ solves $P_k(\hat{\varepsilon})$.*

b) *Suppose X is convex and $f_i : \mathbb{R}^n \to \mathbb{R}$ are convex. If $\hat{x}$ solves $P_k(\hat{\varepsilon})$ for some k, there exists $\lambda \in \mathbb{R}^Q_+ \setminus \{0\}$ such that $\hat{x}$ solves $\min_{x \in X} \sum_{i=1}^{Q} \lambda_i f_i(x)$.*

Proof:

a) Let $\hat{\varepsilon} = f(\hat{x})$. From the choice of $\hat{x}$ we have $\sum_{i=1}^{Q} \lambda_i (f_i(x) - f_i(\hat{x})) \geq 0 \quad \forall \ x \in X$.
Suppose $\hat{x}$ does not solve $P_k(\hat{\varepsilon})$. Then $\exists \ x^0 \in X$ s.t. $f_k(x^0) < f_k(\hat{x})$ and $f_i(x^0) \leq f_i(\hat{x}) \quad i \neq k$.
$\Longrightarrow$ because $\lambda_k > 0$

$$\lambda_k \underbrace{(f_k(x^0) - f_k(\hat{x}))}_{<0} + \sum_{i \neq k} \lambda_i \underbrace{(f_i(x^0) - f_i(\hat{x}))}_{\leq 0} < 0 \qquad \lightning$$

b) Suppose $\hat{x}$ solves $P_k(\hat{\varepsilon})$. Therefore $\nexists \ x \in X$ s.t. $f_k(x) < f_k(\hat{x}), \ f_i(x) \leq f_i(\hat{x}) \leq \varepsilon_i \quad \forall \ i \neq k$.
Using convexity of $f_i \quad \Longrightarrow \exists \ p \in \mathbb{R}^Q, \ p > 0$ s.t. $\sum_{i=1}^{Q} p_i (f_i(x) - f_i(\hat{x})) \geq 0 \quad \forall \ x \in X$.
Since $p \in \mathbb{R}^Q_+ \setminus \{0\}$ we get

$$\sum_{i=1}^{Q} p_i f_i(x) \geq \sum_{i=1}^{Q} p_i f_i(\hat{x}) \quad \forall \ x \in X$$

so $\lambda = p$ is the desired weight vector. (Here we again used the generalized Gordon Theorem 2.24, see also [Man69, p.65].)

□

4.3 Benson's Method

This section is from a paper by Benson, 1978 [Ben78]. In this method $x^0 \in X$ is chosen, and efficiency of $f(x^0)$ is tested by maximizing the sum of $f_i(x^0) - f_i(x)$.

$$\begin{array}{lll} \max & \sum_{i=1}^{Q} \varepsilon_i & \\ \text{s.t.} & f_i(x^0) - \varepsilon_i - f_i(x) = 0 & \\ & \varepsilon_i \geq 0 & \qquad P_\varepsilon(x^0) \\ & x \in X & \end{array}$$

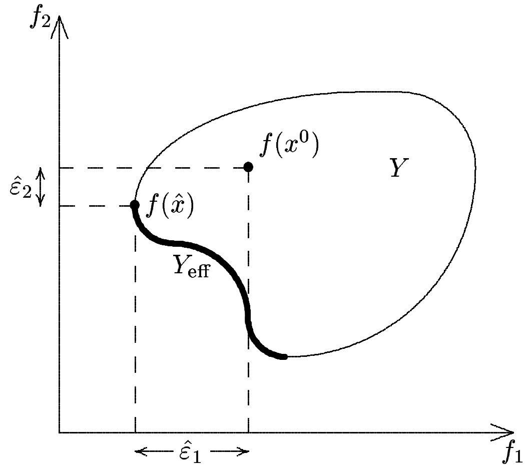


Figure 4.4: Illustration of Benson's Problem

The main result is as follows:

Theorem 4.5. *$x^0 \in X$ is Pareto optimal iff the optimal value of $P_\varepsilon(x^0)$ is 0.*

Proof: $\sum \varepsilon_i = 0 \iff \varepsilon_i = 0 \quad \forall\, i = 1, \ldots, Q$, because $\varepsilon_i \geq 0$

$\iff f_i(x^0) = f_i(x) \quad \forall\, i = 1, \ldots, Q$

$\iff \nexists\, x \in X \quad f_i(x) \leq f_i(x^0) \quad f(x) \neq f(x^0)$

$\iff x^0 \in X_{\text{Par}}$.

□

$P_\varepsilon(x^0)$ is useful for testing x^0 for Pareto optimality, especially in linear problems, as we will see in Chapter 5.

Proposition 4.6. *If problem $P_\varepsilon(x^0)$ has a finite optimal solution and this optimum is attained at (x^*, ε^*) then $x^* \in X_{\text{Par}}$.*

Proof: Suppose $x^* \notin X_{\text{Par}} \implies \exists\, \hat{x}$ s.t. $f_i(\hat{x}) \leq f_i(x^*) \quad \forall\, i, \quad f_q(\hat{x}) < f_q(x^*)$ for some q. So for $\hat{\varepsilon}_i = f_i(x^0) - f_i(\hat{x})$, $(\hat{x}, \hat{\varepsilon})$ is feasible for $P_\varepsilon(x^0)$ because $\hat{\varepsilon}_i = f_i(x^0) - f_i(\hat{x}) \geq f_i(x^0) - f_i(x^*) = \varepsilon_i^* \geq 0$ and $\sum_{i=1}^{Q} \hat{\varepsilon}_i > \sum_{i=1}^{Q} \varepsilon_i^*$ as $\hat{\varepsilon}_k > \varepsilon_k^*$ contradicting the choice of (x^*, ε^*).

□

The question what happens if there is no finite solution of $P_\varepsilon(x^0)$ is answered by Theorem 4.7.

Theorem 4.7. *Assume that f_i are convex, $i = 1, \ldots, Q$ and $X \subset \mathbb{R}^n$ is convex. If $P_\varepsilon(x^0)$ has no finite optimal value then $Y_{\text{p-eff}} = \emptyset$.*

Proof: Since no finite optimum exists $\forall\, \overline{M} \geq 0 \quad \exists\, \overline{x} \in X$ s.t. $f(x^0) - f(\overline{x}) \geq 0$ and

$$\sum_{i=1}^{m} (f_i(x^0) - f_i(\overline{x})) > \overline{M} \tag{4.5}$$

Suppose that x^* is properly efficient (Geoffrion). By Theorem 2.23 $\implies \exists\, \lambda_i > 0 \quad i = 1, \ldots, Q$ s.t. x^* is an optimal solution of $\min_{x \in X} \sum_{i=1}^{Q} \lambda_i f_i(x)$

$\Longrightarrow \sum_{i=1}^{Q} \lambda_i(f_i(x) - f_i(x^*)) \geq 0 \quad \forall\, x \in X.$

Especially, $\sum_{i=1}^{Q} \lambda_i(f_i(x^0) - f_i(x^*)) \geq 0.$

Let $\lambda_m = \min\{\lambda_1, \dots, \lambda_Q\} > 0.$

Given $M \geq 0$ let $\overline{M} := \frac{M}{\lambda_m}$.

(4.5) $\Longrightarrow \exists\, \overline{x}$ s.t. $f_j(x^0) - f_j(\overline{x}) \geq 0 \quad \forall\, j = 1, \dots, Q$ and

$$\lambda_m \sum_{i=1}^{Q} (f_i(x^0) - f_i(\overline{x})) > \frac{M}{\lambda_m} \cdot \lambda_m = M$$

$\Longrightarrow M < \sum_{i=1}^{Q} \lambda_m(f_i(x^0) - f_i(\overline{x})) \leq \sum_{i=1}^{Q} \lambda_i(f_i(x^0) - f_i(\overline{x})) \quad \forall\, M \geq 0.$

Choosing $M = \sum_{i=1}^{Q} \lambda_i(f_i(x^0) - f_i(\overline{x}))$ we get

$\sum_{i=1}^{Q} \lambda_i(f_i(x^0) - f_i(x^*)) < \sum_{i=1}^{Q} \lambda_i(f_i(x^0) - f_i(\overline{x}))$

$\Longrightarrow \sum_{i=1}^{Q} \lambda_i f_i(\overline{x}) < \sum_{i=1}^{Q} \lambda_i f_i(x^*)$ ↯ Contradiction

□

We can combine Theorems 4.7 and 3.9 to get

Corollary 4.8. *Assume $X \subset \mathbb{R}^n$ is convex, $f_i : \mathbb{R}^n \to \mathbb{R}$ are convex $\forall\, i = 1, \dots, Q$ and $f(X)$ is $\mathbb{R}^Q_+$-closed. Then if $P_\varepsilon(x^0)$ has no finite optimal solution $Y_{\text{eff}} = \emptyset$.*

Proof: From Theorem 3.9 $\Longrightarrow Y_{\text{eff}} \subset \text{cl}\, S(Y) = \text{cl}\, Y_{\text{p-eff}}$. From Theorem 4.7 $Y_{\text{p-eff}} = \emptyset \Longrightarrow \text{cl}\, Y_{\text{p-eff}} = \emptyset \Longrightarrow Y_{\text{eff}} = \emptyset$.

□

Example 4.2.

$$\begin{array}{ll} \min & (x^2 - 4, (x-1)^4) \\ \text{s.t.} & -x - 100 \leq 0 \end{array}$$

$$P_\varepsilon(x^0) \qquad \begin{array}{ll} \max & \varepsilon_1 + \varepsilon_2 \\ \text{s.t.} & -x - 100 \leq 0 \\ & (x^0)^2 - 4 - \varepsilon_1 - x^2 + 4 = 0 \\ & (x^0 - 1)^4 - \varepsilon_2 - (x-1)^4 = 0 \\ & \varepsilon_1, \varepsilon_2 \geq 0 \end{array}$$

First, choose $x^0 = 0$.

$$P_\varepsilon(0) \qquad \begin{array}{ll} \max & \varepsilon_1 + \varepsilon_2 \\ \text{s.t.} & -x - 100 \leq 0 \\ & x^2 + \varepsilon_1 = 0 \quad \Longrightarrow \varepsilon_1 = -x^2 \leq 0 \\ & 1 - \varepsilon_2 - (x-1)^4 = 0 \\ & \varepsilon_1, \varepsilon_2 \geq 0 \end{array}$$

We see: $\varepsilon_1 = 0 \implies x = 0 \implies \varepsilon_2 = 0$

Therefore $x^0 = 0,\ \hat{\varepsilon} = (0,0)$ is the only feasible point for $P_\varepsilon(0)$.

From Theorem 4.5 $\quad x^0 = 0 \in X_{\text{Par}}$.

From Exercise 14 we know that $X_{\text{Par}} = X_{\text{s-Par}} = X_{\text{w-Par}} = [0,1]$.

Let us check $x^0 = 2$, to see if $x^0 \notin X_{\text{Par}}$ can be continued.

$$P_\varepsilon(2) \qquad \begin{aligned} \max \quad & \varepsilon_1 + \varepsilon_2 \\ \text{s.t.} \quad & -x - 100 \le 0 \\ & -x^2 + 4 - \varepsilon_1 = 0 \\ & 1 - (x-1)^4 - \varepsilon_2 = 0 \\ & \varepsilon_1, \varepsilon_2 \ge 0 \end{aligned}$$

Here we have $0 \le \varepsilon_1 \le 4, \quad 0 \le \varepsilon_2 \le 1$.

Therefore the optimal value is bounded, and according to Proposition 4.6 an optimal solution of $P_\varepsilon(2)$ defines a Pareto optimal point.

Because $x = 0,\ \varepsilon_1 = 4,\ \varepsilon_2 = 0$ is feasible for $P_\varepsilon(2)$, the optimal value is nonzero. Theorem 4.5 implies $x^0 = 2$ is not Pareto optimal.

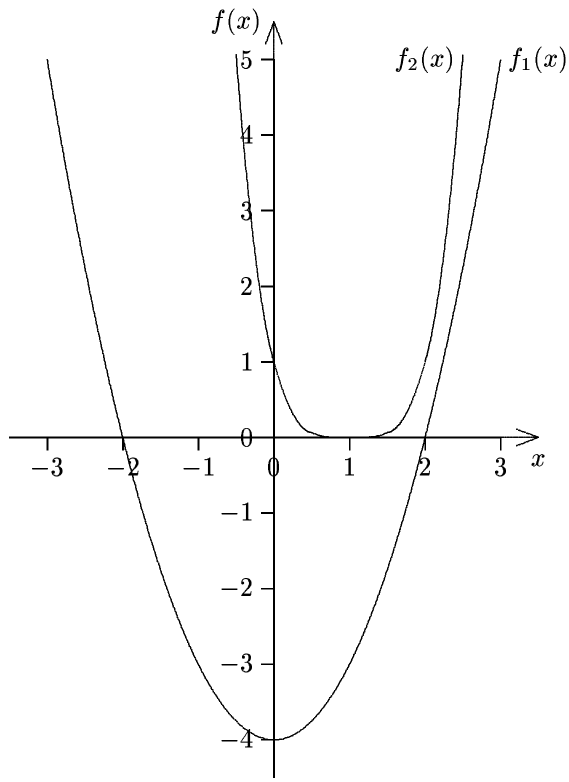


Figure 4.5: Objective Functions of Example 4.2

4.4 Compromise Solutions — Approximation of the Ideal Point

In Exercise 29 we have seen a characterization of $X_{\text{w-Par}}$ by solutions of $\min_{x \in X} \max_{i=1,\dots,Q} \lambda_i f_i(x)$, $\lambda_i \in \operatorname{int} \mathbb{R}^Q_+$. However, we had to assume $\inf_{x \in X} f_i(x) > 0$.
We can avoid this, if we use the ideal point y^0 from Section 4.1. The idea is to find a point $\hat{x}$ s.t. $f(\hat{x})$ is close to y^0.
We use a **norm** as measure of distance.

$$\min_{x \in X} \|f(x) - y^0\| \tag{4.6}$$

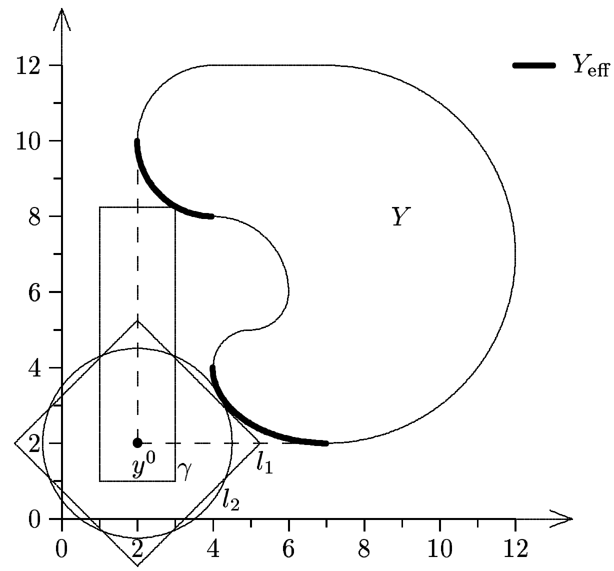


Figure 4.6: Sets $\{y : \|y - y^0\| \le c\}$ for Different Norms: l_1, l_2 and γ

We define two properties of norms:

Definition 4.1.

i) A norm $\|\cdot\| : \mathbb{R}^Q \to \mathbb{R}_+$ is called **monotone**, if for $a, b \in \mathbb{R}^Q$ $|a_i| \le |b_i|$, $i = 1, \dots, Q$ $\|a\| \le \|b\|$ holds, and $|a_i| < |b_i|$ $\forall\, i = 1, \dots, Q \implies \|a\| < \|b\|$.

ii) $\|\cdot\|$ is called **strictly monotone**, if $|a_i| \le |b_i|$, $i = 1, \dots, Q$ and $\exists\, k$ s.t. $|a_k| \ne |b_k| \implies \|a\| < \|b\|$ holds.

We obtain the following results:

Theorem 4.9.

a) *If $\|\cdot\|$ is monotone and $\hat{x}$ solves* (4.6) *then $\hat{x}$ is weakly Pareto optimal. If $\hat{x}$ is unique then $\hat{x} \in X_{\text{Par}}$.*

b) *If $\|\cdot\|$ is strictly monotone and $\hat{x}$ solves* (4.6) *then $\hat{x}$ is Pareto optimal.*

Proof:

a) Suppose $\hat{x}$ solves (4.6) and $\hat{x} \notin X_{\text{w-Par}}$

$\implies \exists\, x \in X \quad \text{s.t.} \quad f_i(x) < f_i(\hat{x}) \quad \forall\, i = 1, \dots, Q$

$\implies 0 \le f_i(x) - y_i^0 < f_i(\hat{x}) - y_i^0, \; i = 1, \dots, Q$

$\implies \|f(x) - y^0\| < \|f(\hat{x}) - y^0\|$ ↯ Contradiction

Now suppose $\hat{x}$ is unique, $\hat{x} \notin X_{\text{Par}}$.

$\implies \exists\, x \in X \quad \text{s.t.} \quad f_i(x) \le f_i(\hat{x}) \quad \forall\, i = 1, \dots, Q$ and $\exists\, k$ s.t. $f_k(x) < f_k(\hat{x})$.

$\implies 0 \le f_i(x) - y_i^0 \le f_i(\hat{x}) - y_i^0$ with strict inequality once.

$\implies \|f(x) - y^0\| \le \|f(\hat{x}) - y^0\|$, from optimality of $\hat{x}$.

$\implies$ equality holds ↯ the uniqueness of $\hat{x}$.

b) Suppose $\hat{x}$ solves (4.6) and $\hat{x} \notin X_{\text{Par}}$

$\implies \exists\, x \in X \quad f_i(x) \le f_i(\hat{x}) \quad i = 1, \dots, Q$ and $\exists\, k$ s.t. $f_k(x) < f_k(\hat{x})$.

$\implies 0 \le f_i(x) - y_i^0 \le f_i(\hat{x}) - y_i \quad i = 1, \dots, Q$ and $0 \le f_k(x) - y_k^0 < f_k(\hat{x}) - y_k^0$

$\implies \|f(x) - y^0\| < \|f(\hat{x}) - y^0\|$ ↯ Contradiction

□

Remark. Let $\|\cdot\| = \|\cdot\|_p$, i.e. $\|y\| = \left(\sum_{i=1}^{Q} |y_i|^p\right)^{\frac{1}{p}}$ for $1 \le p \le \infty$. Then $\|\;\|_p$ is strictly monotone for $1 \le p < \infty$ and monotone for $p = \infty$ ($\|\;\|_p$ is called the $\mathbf{l_p}$**-norm**). Note that $\|y\|_\infty = \max_{i=1,\dots,Q} |y_i|$.

The full strength of the method is obtained when we use weighted norms. We shall only consider l_p-norms now. We consider the following problems:

$$\min_{x \in X} \left(\sum_{i=1}^{Q} w_i (f_i(x) - y_i^0)^p\right)^{\frac{1}{p}} \tag{N_p^w}$$

$$\min_{x \in X} \max_{i=1,\dots,Q} w_i (f_i(x) - y_i^0) \tag{N_∞^w}$$

with $w \in \mathbb{R}_+^Q \setminus \{0\}$.

We obtain the following results:

Theorem 4.10. *A solution $\hat{x}$ of (N_p^w) is Pareto optimal*

i) if it is a unique solution

ii) $w_i > 0 \quad \forall\, i = 1, \dots, Q$.

Proof: Assume $\hat{x}$ solves (N_p^w) but $\hat{x} \notin X_{\text{Par}}$.

i) $\exists\, x \in X \quad f_i(x) \le f_i(\hat{x}) \quad i = 1, \dots, Q, \; f_k(x) < f_k(\hat{x})$ for some k.

Therefore x solves (N_p^w), too ↯ to uniqueness.

ii) From $w_i > 0$ we have $0 \le w_i(f_i(x) - y_i^0) \le w_i(f_i(\hat{x}) - y_i^0) \quad \forall\, i$ with strict inequality for some k. Taking power p and summing up preserves strict inequality. ↯ to $\hat{x}$ solves (N_p^w).

□

Proposition 4.11. *Let $w \gg 0$*

a) *If $\hat{x}$ solves (N_∞^w) then $\hat{x} \in X_{\text{w-Par}}$.*

b) *(N_∞^w) has at least one Pareto optimal solution.*

c) *If (N_∞^w) has a unique solution $\hat{x}$, then $\hat{x} \in X_{\text{Par}}$.*

Proof:

a) $\hat{x}$ solves (N_∞^w) and $\hat{x} \notin X_{\text{w-Par}}$ implies $\exists\, x \quad f_i(x) < f_i(\hat{x}) \quad \forall\, i = 1, \dots, Q$
$\implies f_i(x) - y_i^0 < f_i(\hat{x}) - y_i^0 \quad \forall\, i = 1, \dots, Q$
$\implies w_i(f_i(x) - y_i^0) < w_i(f_i(\hat{x}) - y_i^0) \quad \forall\, i$
$\implies \max w_i(f_i(x) - y_i^0) < \max w_i(f_i(\hat{x}) - y_i^0)$ ↯ Contradiction

b) Assume that no solution of (N_∞^w) is Pareto optimal.
Suppose $\hat{x}$ solves $(N_\infty^w) \implies \exists\, x \in X_{\text{Par}}$
$f_i(x) \le f_i(\hat{x}) \quad i = 1, \dots, Q$ with strict inequality for one k
$\implies w_i(f_i(x) - y_i^0) \le w_i(f_i(\hat{x}) - y_i^0) \quad i = 1, \dots, Q$
$\implies x$ is optimal for (N_∞^w) ↯ Contradiction

c) follows from b)

□

The problem (N_∞^w) can be used to obtain all (weakly) Pareto optimal points. Let $\varepsilon \gg 0$ and define $y^{00} = y^0 - \varepsilon$. Then $f_i(x) > y_i^{00} \quad \forall\, x \in X,\ i = 1, \dots, Q$.

Theorem 4.12. $\hat{x} \in X_{\text{w-Par}} \iff \exists\, w \gg 0$ *s.t.* $\hat{x}$ *solves*

$$\min_{x \in X} \max_{i=1,\dots,Q} w_i(f_i(x) - y_i^{00}) \tag{4.7}$$

Proof:

„ $\Longleftarrow$ " is the same as b) in Proposition 4.11.

„ $\Longrightarrow$ " Let $w_i = \frac{1}{f_i(\hat{x}) - y_i^{00}} > 0$.
Suppose $\hat{x}$ does not solve (4.7) $\implies \exists\, x \in X$

$$\max_i w_i(f_i(x) - y_i^{00}) < \max_i \frac{1}{f_i(\hat{x}) - y_i^{00}}(f_i(\hat{x}) - y_i^{00}) = 1$$

$\implies f_i(x) - y_i^{00} < f_i(\hat{x}) - y_i^{00} \quad \forall\, i = 1, \dots, Q$
$\implies f_i(x) < f_i(\hat{x}) \quad \forall\, i = 1, \dots, Q$ ↯ to $\hat{x} \in X_{\text{w-Par}}$.

□

To prove the main results about compromise solutions, we introduce some notation.
Let $W := \{w \in \mathbb{R}^Q : w_i \ge 0, \sum w_i = 1\}$ and $W^0 = \operatorname{ri} W = \{w \in \mathbb{R}^Q : w_i > 0, \sum w_i = 1\}$. Furthermore, for $w \in W$ and $y \in Y$: $w \odot y = (w_1 y_1, \dots, w_Q y_Q)$. The set of best approximations of y^{00} is denoted for a certain weight w and norm $\| \, \|_p$ by

$$A(w,p,Y) := \{\hat{y} \in Y : \|w \odot (\hat{y} - y^{00})\|_p = \inf_{y \in Y} \|w \odot (y - y^{00})\|_p\} \tag{4.8}$$

$$A(Y) = \bigcup_{w \in W^0} \bigcup_{1 \le p < \infty} A(w,p,Y). \tag{4.9}$$

We have seen before that

$$A(Y) \subset Y_{\text{eff}} \subset Y_{\text{w-eff}} = \bigcup_{w \in W^0} A(w, \infty, Y). \tag{4.10}$$

This result can be strengthened, as we show below.

Remark. The family of l_p-norms has the following properties:

(P1) $\|y\|_\infty \le \|y\|_p \quad \forall\, 1 \le p < \infty \quad \forall\, y \in \mathbb{R}^Q$

(P2) $\|y\|_p \to \|y\|_\infty \quad p \to \infty \quad \forall\, y \in \mathbb{R}^Q$

(P3) $\|\ \|_p$ is strictly monotone $\forall\, 1 \le p < \infty$.

Theorem 4.13 ([SNT85]). *$A(Y) \subset Y_{\text{p-eff}} \subset Y_{\text{eff}} \subset \text{cl}(A(Y))$ if Y is $\mathbb{R}^Q_+$-closed.*

Proof:

1. $A(Y) \subset Y_{\text{p-eff}}$:

   Let $\hat{y} \in A(Y) \implies \exists\, w \in W^0,\ p \in [1, \infty)$ s.t. $\|w \odot (\hat{y} - y^{00})\|_p \le \|w \odot (y - y^{00})\|_p \quad \forall\, y \in Y$.

   Suppose $\hat{y} \notin Y_{\text{p-eff}}$

   $\implies \exists\, \{\beta_k\} \subset \mathbb{R},\ \{y^k\} \subset Y,\ \{d_k\} \subset \mathbb{R}^Q_+$ s.t. $\beta_k > 0,\ \beta_k(y^k + d^k - \hat{y}) \to -d,\ d \in \mathbb{R}^Q_+ \setminus \{0\}$.

   Distinguish two cases:

   a) β_k bounded:

   Wlog $\beta_k \to \beta_0 \ge 0$

   If $\beta_0 = 0$ from $y^k + d^k - \hat{y} \ge y^{00} - \hat{y}$ we have $\underbrace{\beta_k(y^k + d^k - \hat{y})}_{-d} \ge \underbrace{\beta_k(y^{00} - \hat{y})}_{\to 0}$ and thus $-d \ge 0$ ↯ Contradiction

   If $\beta_0 > 0 \implies y^k + d^k - \hat{y} \longrightarrow \frac{-d}{\beta_0} \ne 0$.

   $Y + \mathbb{R}^Q_+$ is closed $\implies \hat{y} - \frac{d}{\beta_0} \in Y + \mathbb{R}^Q_+ \implies \exists\, y^0 \in Y$ s.t. $\hat{y} > y^0$

   By strict monotonicity $\implies \|w \odot (\hat{y} - y^{00})\|_p > \|w \odot (y^0 - y^{00})\|_p$ ↯ to choice of $\hat{y}$.

   b) β_k unbounded:

   Wlog $\beta_k \to \infty \implies y^k + d^k - \hat{y} \to 0$

   Because $\hat{y}_i > y^{00}_i \quad \forall\, i = 1, \dots, Q \quad \exists\, \overline{\beta} > 0$ s.t. $0 \le \hat{y} - \frac{d}{\beta} - y^{00} < \hat{y} - y^{00} \quad \forall\, \beta > \overline{\beta}$

   From strict monotonicity

   $$\|w \odot (\hat{y} - \frac{d}{\beta} - y^{00})\|_p < \|w \odot (\hat{y} - y^{00})\|_p \quad \forall\, \beta > \overline{\beta}.$$

   and since $\beta_k \to \infty \implies \beta_k > \overline{\beta} \quad \forall\, k \ge k_0$ sufficiently large

   $$\begin{aligned} \implies \|w \odot (y^k + d^k - y^{00})\|_p &= \|w \odot (y^k + d^k - \hat{y} + \tfrac{d}{\beta_k} + \hat{y} - \tfrac{d}{\beta_k} - \overline{y})\|_p \\ &\le \underbrace{\|w \odot (y^k + d^k - \hat{y})\|_p}_{\to 0} + \underbrace{\frac{\|w \odot d\|_p}{\beta_k}}_{\to 0} + \|w \odot (\hat{y} - \tfrac{d}{\beta_k} - y^{00})\|_p \end{aligned}$$

   $$\implies \|w \odot (y^k + d^k - y^{00})\|_p \le \|w \odot (\hat{y} - \tfrac{d}{\beta_k} - y^{00})\|_p < \|w \odot (\hat{y} - y^{00})\|_p$$

   But since $y^k + d^k - y^{00} \ge y^k - y^{00} > 0$ this implies

   $$\|w \odot (y^k - y^{00})\|_p < \|w \odot (\hat{y} - y^{00})\|_p \quad \text{↯ to choice of } \hat{y}$$

2. $Y_{\text{eff}} \subset \text{cl}(A(Y))$:

Let $\hat{y} \in Y_{\text{eff}}$. We show that $\forall\, \varepsilon > 0 \quad \exists\, y^\varepsilon \in A(Y)$ s.t. $\|y^\varepsilon - \hat{y}\|_\infty < \varepsilon$.

Claim: $\exists\, y' \gg \hat{y}$ s.t. $\|y - \hat{y}\|_\infty < \varepsilon \quad \forall\, y \in (y' - \mathbb{R}^Q_+) \cap Y$.

Assume there is no such y'.

Then $\exists\, \{\hat{y}^k\} \subset \mathbb{R}^Q,\ \hat{y}^k > \hat{y},\ \hat{y}^k \to \hat{y}$ and $\forall\, k \quad \exists\, y^k \in (\hat{y}^k - \mathbb{R}^Q_+) \cap Y$ s.t. $\|y^k - \hat{y}\| \geq \varepsilon$.

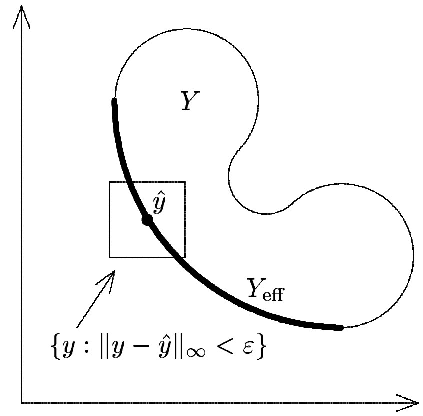


Figure 4.7: $\{y : \|y - \hat{y}\|_\infty < \varepsilon\}$

$Y + \mathbb{R}^Q_+$ is closed and $Y \subset y^{00} + \mathbb{R}^Q_+$ (bounded below). We can assume wlog $y^k \to \overline{y} + \overline{d}$, $\overline{y} \in Y$, $\overline{d} \geq 0$ and $\|\overline{y} + \overline{d} - \hat{y}\|_\infty \geq \varepsilon$. On the other hand $\overline{y} + \overline{d} \in (\hat{y} - \mathbb{R}^Q_+) \cap (Y + \mathbb{R}^Q_+) = \{\hat{y}\}$, since $\hat{y} \in Y_{\text{eff}}$ ⚡ Contradiction.

So we have $y^{00} < \hat{y} \ll y'$ and $\exists\, w \in W^0,\ \beta > 0$ s.t. $y' - y^{00} = \beta(\frac{1}{w_1}, \ldots, \frac{1}{w_Q})$

$\Longrightarrow\ w_i(\hat{y}_i - y_i^{00}) < w_i(y'_i - y_i^{00}) = \beta \qquad \forall\, i = 1, \ldots, Q$

$\Longrightarrow\ \|w \odot (\hat{y} - y^{00})\|_\infty < \beta.$

Let $y(p) \in A(w, p, Y)$, this exists as $Y + \mathbb{R}^Q_+$ is closed.

$\Longrightarrow\ \|w \odot (y(p) - y^{00})\|_\infty \overset{\text{(P1)}}{\leq} \|w \odot (y(p) - y^{00})\|_p \leq \|w \odot (\hat{y} - y^{00})\|_p$

$\quad \overset{\text{(P2)}}{\longrightarrow} \|w \odot (\hat{y} - y^{00})\|_\infty < \beta.$

Thus $\|w \odot (y(p) - y^{00})\|_\infty \leq \beta$ for p sufficiently large.

$\Longrightarrow\ y_i(p) - y_i^{00} \leq \frac{\beta}{w_i} = y'_i - y_i^{00} \quad \forall\, i = 1, \ldots, Q$

$\Longrightarrow\ y(p) \leq y'$

$\Longrightarrow\ y(p) \in (y' - \mathbb{R}^Q_+) \cap Y$

So $y^\varepsilon := y(p)$ for sufficiently large p has the desired properties.

□

The proof of the second inclusion suggests that, if Y is not $\mathbb{R}^Q_+$-convex, p has to be large. The value of p seems to be related to the degree of nonconvexity of Y. Note that if Y is $\mathbb{R}^Q_+$-convex, $p = 1$ is enough, and in general $p = \infty$ works. See also Exercise 34.

We note that the inclusion $\text{cl}\, A(Y) \subset Y_{\text{eff}}$ may not be true.

Example 4.3. Let $Y := \{y \in \mathbb{R}^2 : y_1^2 + (y_2 - 1)^2 \leq 1\} \cup \{y \in \mathbb{R}^2 : y_1 \geq 0,\ y_2 \geq -1\}$

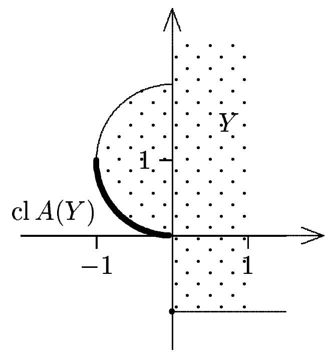


Figure 4.8: $\operatorname{cl} A(Y) \not\subset Y_{\text{eff}}$

Here $0 \notin Y_{\text{eff}}$ but $0 \in \operatorname{cl} A(Y)$.

It should also be noted that if y^{00} is replaced by y^0 in the Theorem 4.12 then not even $Y_{\text{p-eff}} \subset \bigcup_{w \in W^0} A(W^0, \infty, Y)$ is true. See Exercise 35.

Remark.

- The result of Theorem 4.13 is also true, if y^{00} is replaced by y^0, with a modification of the proof.
- Note that only properties (P1)-(P3) have been used. Theorem 4.13 is true for any family of norms that satisfies such properties. This fact has been used by various researchers.

To conclude the section, we present an example.

Example 4.4. We solve Example 4.2 with the compromise solution method:

$$\min(x^2 - 4, (x-1)^4)$$
$$-x - 100 \leq 0$$

Let $w = (\frac{1}{2}, \frac{1}{2})$ and $p = 2$.
The ideal point is $y^0 = (-4, 0)$. We choose $y^{00} = (-5, -1)$.
So (N_p^w) is

$$\min \sqrt{\tfrac{1}{2}(x^2 - 4 + 5)^2 + \tfrac{1}{2}((x-1)^4 + 1)^2}$$
$$\text{s.t.} \ -x - 100 \leq 0$$

Noting that minimization of the term under the root is enough we denote

$$\begin{aligned} g(x) &= \tfrac{1}{2}(x^2+1)^2 + \tfrac{1}{2}((x-1)^4+1)^2 \\ g'(x) &= (x^2+1) \cdot 2x + ((x-1)^4+1) \cdot 4(x-1)^3 \\ &= 2x^3 + 2x + 4(x-1)^7 + 4(x-1)^3 \end{aligned}$$

From $g'(x) = 0$ we obtain $x^* = 0.40563 \in X_{\text{Par}}$ which is Pareto optimal.

4.5 Exercises to Chapter 4

29. Consider $\min_{x \in X} (f_1(x), \dots, f_Q(x))$ and assume

$$0 < \inf_{x \in X} f_i(x) \quad \forall\, i = 1, \dots, Q.$$

Prove that $x \in X_{\text{w-Par}}$ if and only if x solves

$$\min_{x \in X} \max_{i=1,\dots,Q} \lambda_i f_i(x)$$

for some $\lambda \in \operatorname{int} \mathbb{R}^Q_+$.

30. Suppose $\hat{x}$ solves

$$\min_{x \in X} \sum_{i=1}^{Q} \lambda_i f_i(x) \text{ with } \lambda \in \mathbb{R}^Q_+ \setminus \{0\},$$

and that $\hat{x}$ is the unique solution of this problem. Then $\exists\, \hat{\varepsilon}$ s.t. $\hat{x}$ solves $P_k(\hat{\varepsilon})$ for all $k = 1, \dots, Q$.

31. (Corley's method), [Cor80]

Show that $\hat{x} \in X_{\text{Par}}$ if and only if $\exists\, \lambda \in \operatorname{int} \mathbb{R}^Q_+$ and $\exists\, \varepsilon \in \mathbb{R}^Q$ s.t. $\hat{x}$ is an optimal solution of

$$\min_{x \in X} \sum_{i=1}^{Q} \lambda_i f_i(x)$$

subject to $f(x) \leq \varepsilon$.

32. Consider the following problem:

$$\begin{array}{lrcl} \min & -6x_1 - 4x_2 & & \\ \min & -x_1 & & \\ \text{s.t.} & x_1 + x_2 & \leq & 100 \\ & 2x_1 + x_2 & \leq & 150 \\ & x_1, x_2 & \geq & 0 \end{array}$$

Use $\varepsilon = 0$ and solve the ε-constraint problem $P_1(\varepsilon)$. Check if your resulting optimal solution x^* is Pareto optimal by Benson's $P_\varepsilon(x^*)$ test.

33. Solve the problem of Exercise 32 by the compromise solution method. Use $w = (\frac{1}{2}, \frac{1}{2})$ and find the solution of (N_p^w) for $p = 1, 2, \infty$.

34. Let $Y = \{y \in \mathbb{R}^2 : y_1 + y_2 \geq 1,\ 0 \leq y_1 \leq 1\}$.

Show that $\hat{y} = (0, 1) \in Y_{\text{p-eff}}$ (Benson) but $\nexists\, w \in W^\circ$ s.t. $\hat{y} \in A(w, \infty, Y)$, if y^0 is used in the compromise solution method.

35. Let $Y = \left\{(y_1, y_2) \in \mathbb{R}^2_+ : y_1^2 + y_2^2 \geq 1\right\}$. Show that $\exists\, 1 < p < \infty$ s.t.

$$Y_{\text{eff}} = \bigcup_{w \in W^\circ} A(w, p, Y)$$

Choose y^0 in the definition of $A(w, p, Y)$.

36. A function $g : \mathbb{R}^Q \to \mathbb{R}$ is called strictly increasing, if for $a, b \in \mathbb{R}^Q$ with $a < b$ $(a_i \leq b_i,\ i = 1, \dots, Q,\ a \neq b)$ $g(a) < g(b)$ holds.

Consider the following problem, where $\varepsilon \in \mathbb{R}^Q,\ f : \mathbb{R}^n \to \mathbb{R}^Q$.

$$\begin{aligned} \min \quad & g(f(x)) \\ \text{s.t.} \quad & x \in X \\ & f(x) \leq \varepsilon \end{aligned} \qquad (P_{g,\varepsilon})$$

Prove: If g is strictly increasing, then $x \in X_{\text{Par}} \iff \exists\, \varepsilon$ s.t. x solves $P_{g,\varepsilon}$ with finite objective value.

37. Show that Benson's $P_\varepsilon(x^\circ)$ problem, the weighted sum scalarization with $\lambda \in \operatorname{int} \mathbb{R}^Q_+$, the compromise solution method and Corley's method (see Exercise 31) can be seen as special cases of $(P_{g,\varepsilon})$.

Chapter 5

Multicriteria Linear Programming

5.1 Introduction

In this chapter we specifically address multiobjective linear programming programs. Most of the material is based on Steuer, 1985, [Ste85]. I.e. we assume $X = \{x \in \mathbb{R}^n : Ax = b,\ x \geq 0\}$ where A is a $m \times n$ matrix and

$$f_i(x) = c_i x \qquad i = 1, \dots, Q. \tag{5.1}$$

The problem is therefore written as

$$\begin{array}{ll} \min & Cx \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \tag{MCLP}$$

with a $Q \times n$ criteria matrix C. In terms of classification this problem is $(X, C, \mathbb{R}^Q)/\text{id}/(\mathbb{R}^Q, <)$. Since X and $Y = CX$ are closed, convex we can apply all results that hold for convex MCOP, especially

- $S(Y) = Y_{\text{p-eff}} \subset Y_{\text{eff}} \subset \text{cl}\, S(Y)$ (Theorem 3.9)
- If $\exists\, y \in \mathbb{R}^Q$ s.t. $CX \subset y + \mathbb{R}^Q_+$ then Y_{eff} is connected. (Theorem 3.21)
- If X is bounded, $X_{\text{w-Par}}$ is connected. (Theorem 3.24)

In fact, due to linearity, these results can all be strengthened.

Lemma 5.1. *$x^0 \in X$ is Pareto optimal $\iff$ the LP*

$$\begin{array}{lrcl} \max & e^t y & & \\ s.t. & Ax & = & b \\ & Cx + Iy & = & Cx^0 \\ & x, y & \geq & 0 \end{array} \tag{P}$$

($e^t = (1, \dots, 1)$, $I =$ identity matrix) has an optimal solution $\hat{x}, \hat{y}$ with $\hat{y} = 0$.

Proof: This is Theorem 4.5 for (MCLP).

□

Lemma 5.2. *$x^0 \in X$ is Pareto optimal iff the LP*

$$\begin{array}{lrcl} \min & u^t b + w^t C x^0 & & \\ s.t. & u^t A + w^t C & \geq & 0 \\ & w & \geq & e \end{array} \tag{D}$$

has an optimal solution $\hat{u}, \hat{w}$ with $\hat{u}^t b + \hat{w}^t C x^0 = 0$.

Proof: (D) is the dual of (P). Therefore $\hat{x}, \hat{y}$ is optimal in (P) $\iff$ (D) has an optimal solution $\hat{u}, \hat{w}$ and $e^t \hat{y} = \hat{u}^t b + \hat{w}^t C x^0 = 0$.

□

Using these Lemmas we can prove:

Theorem 5.3 (Isermann, 1974, [Ise74]).
$S(Y) = Y_{\text{eff}}$, *i.e.* $x^0 \in X_{\text{Par}} \iff \exists\, \lambda \in \operatorname{int} \mathbb{R}^Q_+$ *s.t.* $\lambda^t C x^0 \leq \lambda^t C x \quad \forall\, x \in X$.

Proof:

„ $\Longleftarrow$ " is always true, see Theorem 3.1.

„ $\Longrightarrow$ " $x^0 \in X_{\text{Par}} \stackrel{\text{Lemma 5.2}}{\Longrightarrow}$ (D) has an optimal solution $\hat{u}, \hat{w}$ s.t. $\hat{u}^t b = -\hat{w}^t C x^0$.
Also $\hat{u}$ is an optimal solution of the problem

$$\min\{u^t b : u^t A \geq -\hat{w}^t C\} \tag{P2}$$

$\Longrightarrow$ an optimal solution of the dual of (P2)

$$\max\{-\hat{w}^t C x : Ax = b,\ x \geq 0\} \tag{D2}$$

exists. Since $u^t b \geq -\hat{w}^t C x \quad \forall\, u$ feasible in (P2) and $\forall\, x$ feasible in (D2) and $\hat{u}^t b = -\hat{w}^t C x^0 \implies x^0$ is optimal in (D2). Because $\hat{w} \geq e \gg 0$ we can use $\lambda = \hat{w}$.
(Note that (D2) is equivalent to $\min \hat{w}^t C x,\ Ax = b,\ x \geq 0$.)

□

Consequently we have $S(Y) = Y_{\text{eff}} = Y_{\text{p-eff}}$ for MCLP and we can find Y_{eff} by weighted sum scalarization with strictly positive weights.

In order to understand the following development of a multicriteria simplex method, we review some results of linear programming. An LP is the following problem

$$\begin{array}{ll} \min & cx \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \tag{LP}$$

We assume that rank $A = m$. A nonsingular submatrix B of A is called **basis**.
We split $A = (B, N)$ and $x = (x_B, x_N)$ and obtain

$$(B, N)(x_B, x_N) = b \tag{5.2}$$

$$\iff \quad x_B = B^{-1}(b - Nx_N) \tag{5.3}$$

Setting $x_N = 0$ we have $x_B = B^{-1}b$. x_B is a **basic solution**, and a **basic feasible solution (bfs)**, if $x_B \geq 0$. The values $\overline{c} = c - c_B B^{-1}A$ are called **reduced costs**.

Linear Programming Theory has the following results:

- If $X \neq \emptyset$ then a basic feasible solution exists.
- If, furthermore, X is bounded then an optimal basic feasible solution exists.
- A bfs is optimal $\iff$ $\overline{c} \geq 0$.

We also need some geometry:
Let $d \in \mathbb{R}^n$ then

$$H_{d,r} := \{x \in \mathbb{R}^n : \langle x, d \rangle = r\} \tag{5.4}$$

is called a **hyperplane**. A hyperplane defines closed and open halfspaces

$$\overline{H}^-_{d,r} := \{x \in \mathbb{R}^n : \langle x, d \rangle \leq r\} \tag{5.5}$$

$$H^-_{d,r} := \{x \in \mathbb{R}^n : \langle x, d \rangle < r\} \tag{5.6}$$

For $X \subset \mathbb{R}^n$ a hyperplane H is called **supporting hyperplane at $\overline{\mathbf{x}}$** (H supports X at $\overline{x}$) if $\overline{x} \in X \cap H$ and $X \subset \overline{H}^-_{d,r}$.

Now let X be the intersection of finitely many closed halfspaces. Then X is called **polyhedron** (e.g. $X = \{x : Ax = b, \ x \geq 0\}$ is a polyhedron).
A polyhedron X is called a **polytope** if X is bounded.
$\overline{x} \in X$ is called an **extreme point** of X if

$$\overline{x} = \alpha x_1 + (1 - \alpha)x_2, \quad x_1, x_2 \in X, \quad 0 \leq \alpha \leq 1 \quad \Longrightarrow \quad x_1 = x_2$$

Assume that $X \neq \emptyset$ is a polyhedron, $X = \{x : Ax \leq b\}$.
Let $r \in \mathbb{R}^n$ be such that $Ar \leq 0$. Then r is called a **ray**.
A ray is called an **extreme ray** if there are no rays r^1, r^2, $\ r^1 \neq \lambda r^2 \quad \forall\, \lambda \in \mathbb{R}_+$ s.t. $r = \frac{1}{2}(r^1 + r^2)$.
The **dimension** of a polyhedron X is the maximal number of affinely independent points of X, minus 1. Let H be a supporting hyperplane of polyhedron X. Then $F = X \cap H$ is called a **face** of X. A face F is itself a polyhedron, thus $\dim F$ is defined. We consider only faces F with $\emptyset \neq F \neq X$. An **extreme point** is a face of dimension 0. A face of dimension 1 is called an **edge** (if it is bounded). A **facet** is a face of dimension $\dim X - 1$.
Finally, a face F is called **maximal**, if there is no face F' of higher dimension s.t. $F \subset F'$, thus facets are maximal faces.

From linear programming it is known that

- Bfs correspond to extreme points of X.
- If $X \neq \emptyset$ and the LP is bounded, the set of optimal solutions of an LP is a face of X or X itself.
- For each extreme point $\hat{x}$ of X, $\exists\, c \in \mathbb{R}^n$ s.t. $\hat{x}$ solves $\min cx,\ x \in X$.

Now we take a look at **parametric linear programming**.

Let c^1, c^2 be two cost rows and consider a combined (parametric) objective

$$c^* = \lambda c^1 + (1-\lambda)c^2 \tag{5.7}$$

A parametric LP is

$$\begin{aligned} \min \quad & c^* x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{PLP}$$

with the objective to find an optimal solution for each value of λ.

The algorithm to do so is as follows:

Phase I: Determine, if possible, an initial bfs (extreme point of X)

Phase II: Solve the problem with $c^1 x$, to obtain an optimal bfs (extreme point)

Phase III: Vary λ from 1 to 0, solve the corresponding problem to obtain optimal bfs for all λ.

From $c^* = \lambda c^1 + (1-\lambda)c^2$ we get

$\overline{c}^* = \lambda \overline{c}^1 + (1-\lambda)\overline{c}^2$.

Now suppose $\hat{B}$ is an optimal basis for some $\hat{\lambda}$. Then $\overline{c}^* \geq 0$ and we distinguish 2 cases

1.) $\overline{c}^2 \geq 0$. Then, for all $\lambda < \hat{\lambda}$ $\overline{c}^* \geq 0$. Thus $\hat{B}$ is an optimal basis for all $0 \leq \lambda \leq \hat{\lambda}$.

2.) $\exists\, j$ s.t. $\overline{c}^2_j < 0$. Then $\exists\, \lambda < \hat{\lambda}$ s.t. $\overline{c}^*_j < 0$.

We determine the critical value, where the first $\overline{c}^*_j$ becomes negative:

$$\begin{aligned} \overline{c}^*_j = \lambda \overline{c}^1_j + (1-\lambda)\overline{c}^2_j = 0 \quad & \text{for } j \text{ s.t. } \overline{c}^2_j < 0,\ \overline{c}^1_j \geq 0 \\ \Longrightarrow \lambda(\overline{c}^1_j - \overline{c}^2_j) + \overline{c}^2_j = 0 & \\ \lambda = \tfrac{-\overline{c}^2_j}{\overline{c}^1_j - \overline{c}^2_j} & \end{aligned}$$

So let $\lambda' := \max_{j \in J} \frac{-\overline{c}^2_j}{\overline{c}^1_j - \overline{c}^2_j}$ where $J = \{j : \overline{c}^2_j < 0,\ \overline{c}^1_j \geq 0\}$ be the critical value. Then $\hat{B}$ is optimal for $\min c^* x,\ Ax = b,\ x \geq 0$ for all $\lambda \in [\lambda', \hat{\lambda}]$ and at λ' new bases become optimal.

Let j' be the index at which the critical value λ' is attained. Then we choose j' as pivot column, and pivot it into the basis. Proceeding in this way, we generate a sequence of critical values $1 = \lambda^1 > \ldots > \lambda^p = 0$ and optimal bases $B^1, \ldots, B^p$ which define optimal solutions of (PLP) for all λ. Essentially, we have solved a bicriteria linear program.

Example 5.1. Consider the LP with $c^1 = (3,1)$, $c^2 = (-1,-2)$ and

$$\begin{array}{lrcl} \min & c^* x & & \\ \text{s.t.} & x_2 & \leq & 3 \\ & 3x_1 - x_2 & \leq & 6 \\ & x_1, x_2 & \geq & 0 \end{array}$$

The initial Simplex Tableau of this problem, with c^1 optimized is

-1	-2	0	0	0	$\overline{c}^2$
3	1	0	0	0	$\overline{c}^1$
0	$\boxed{1}$	1	0	3	
3	-1	0	1	6	

$\overline{c}^* = (3,1,0,0)$

Therefore $J = \{1,2\}$ and $\lambda' = \max\{\frac{1}{3+1}, \frac{2}{3}\} = \frac{2}{3}$, $j' = 2$. We pivot x_2 into the basis and get

-1	0	2	0	6	$\overline{c}^2$
3	0	-1	0	-3	$\overline{c}^1$
0	1	1	0	3	
$\boxed{3}$	0	1	1	9	

$\overline{c}^* = (1,0,\frac{5}{3},0)$

Now $J = \{1\}$ and $\lambda' = \frac{1}{4}$. We pivot x_1 into the basis to get

0	0	$\frac{7}{3}$	$\frac{1}{3}$	9	$\overline{c}^2$
0	0	-2	-1	-12	$\overline{c}^1$
0	1	1	0	3	
1	0	$\frac{1}{3}$	$\frac{1}{3}$	3	

$\overline{c}^* = (0,0,\frac{11}{6},\frac{1}{12})$

Now $J = \emptyset$ and the algorithm stops.

The result is:

$B = (a_3, a_4)$, $x = (0,0)$ is optimal for $\lambda \in [\frac{2}{3}, 1]$

$B = (a_2, a_4)$, $x = (0,3)$ is optimal for $\lambda \in [\frac{1}{4}, \frac{2}{3}]$, and

$B = (a_1, a_2)$, $x = (3,3)$ is optimal for $\lambda \in [0, \frac{1}{4}]$.

Graphically:

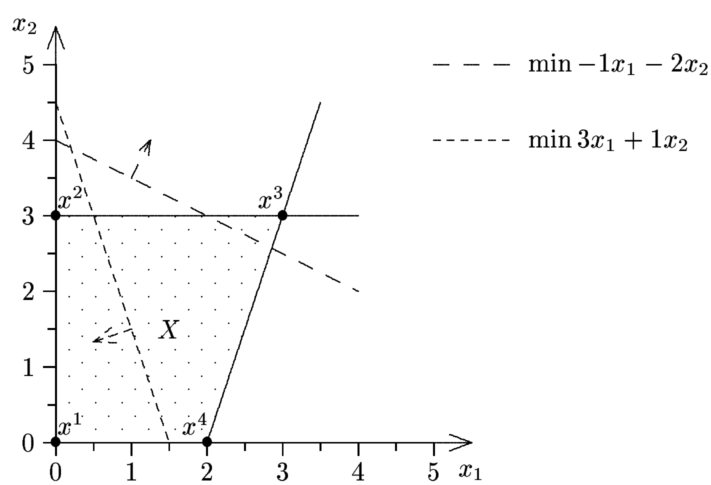


Figure 5.1: Feasible Set and Objectives in Example 5.1

In objective space: $C = \begin{pmatrix} c^1 \\ c^2 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -1 & -2 \end{pmatrix}$

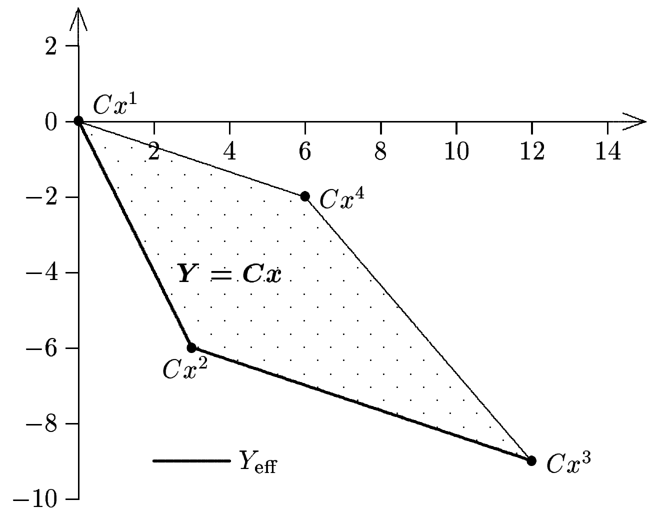


Figure 5.2: Criterion Space in Example 5.1

Note that in the sequence $1 = \lambda^1 > \lambda^2 > \ldots > \lambda^p = 0$ an optimal bfs x_{B^i} is always optimal for all $\lambda \in [\lambda^i, \lambda^{i+1}]$. Therefore, for each λ^i, $2 \le i \le p-1$ we have **two** optimal bfs x_{B^i} and $x_{B^{i-1}}$.
$\Longrightarrow$ $\text{conv}(x_{B^i, B^{i-1}})$ is optimal for $\lambda = \lambda^i$.
Because $Y = Cx$ is a polyhedron, and here $Y \subset \mathbb{R}^2$ and because $Y_{\text{eff}} \subset \delta Y$, we know that Y_{eff} must consist of efficient edges (and possibly extreme rays).
Therefore

1.) $Y_{\text{eff}} = \bigcup_{i=2}^{p-1} \text{conv}(x_{B^{i-1}}, x_{B^i})$

2.) Y_{eff} is connected.

The general case of Q criteria will be investigated now.

5.2 Theory of MCLP

We consider

$$\begin{aligned} \min \quad & Cx \\ \text{s.t.} \quad & Ax = b \\ & x \ge 0 \end{aligned} \qquad \text{(MCLP)}$$

For $\lambda \in \text{int}\, \mathbb{R}^Q_+$ we denote by LP(λ) the LP $\min_{x \in X} \lambda^t Cx$. We denote by $\overline{C} = C - C_B B^{-1} A$ the reduced cost matrix with respect to basis B, and $R = \overline{C}_N$ $\quad(\overline{C}_B = 0$ always). Proofs will often use Theorem 5.3 in this section.

Lemma 5.4. *If $X_{\text{Par}} \neq \emptyset$ then X has a Pareto optimal extreme point.*

Proof: From Theorem 5.3 $X_{\text{Par}} = S(Y),\ X_{\text{Par}} \neq \emptyset$

$\implies \exists\ \lambda \in \operatorname{int} \mathbb{R}_+$ s.t. $\min\limits_{x \in X} \lambda^t Cx$ has an optimal solution

$\implies \min\limits_{x \in X} \lambda^t Cx$ has an optimal extreme point solution, which is Pareto optimal, by Theorem 5.3.

□

Definition 5.1. B is called **efficient basis** iff $\exists\ \lambda \in \mathbb{R}^Q_+$ s.t. B is an optimal basis of LP(λ) (in particular, B defines a bfs x_B).

Lemma 5.5.

a) Let B be an efficient basis and x_B be the extreme point of X associated with B, then $x_B \in X_{\text{Par}}$.

b) Let $x \in X_{\text{Par}}$ be an extreme point. Then $\exists$ efficient basis B associated with x.

Proof:

a) B efficient basis $\implies \exists\ \lambda \in \operatorname{int} \mathbb{R}^Q_+$ s.t. B is optimal basis for LP(λ) $\implies x_B$ is an extreme point optimal solution of LP(λ) $\overset{\text{Theorem 5.3}}{\implies} x_B \in X_{\text{Par}}$

b) Theorem 5.3 $\implies \exists\ \lambda \in \operatorname{int} \mathbb{R}^Q_+$ s.t. x is optimal for LP(λ).
Since x is an extreme point $\implies \exists$ optimal basis of LP(λ) associated with x.

□

Definition 5.2. Two bases $\overline{B}$ and $\hat{B}$ are called **adjacent**, if one can be obtained from the other by a single pivot step.

Definition 5.3.

a) Let B be an efficient basis. x_j is called **efficient nonbasic variable** if $\exists\ \lambda \in \operatorname{int} \mathbb{R}^Q_+$ s.t. $\lambda^t R \geq 0,\ \lambda^t r^j = 0$, where r^j is the j-th column of R.

b) Let B be an efficient basis and x_j an efficient nonbasic variable. Then a feasible pivot from B with x_j entering the basis is called an **efficient pivot w.r.t. B and x_j**.

Lemma 5.6. *Let B be an efficient basis and x_j be an efficient nonbasic variable. Then any efficient pivot from B leads to an adjacent efficient basis $\hat{B}$.*

Proof: Let x_j be the entering variable

$\implies \exists\ \lambda \in \operatorname{int} \mathbb{R}^Q_+$ s.t. $\lambda^t R_B \geq 0,\ \lambda^t r^j_B = 0$. Thus x_j is a nonbasic variable with reduced cost 0.

$\implies$ Reduced costs do not change after a pivot with x_j entering.

$\implies \lambda^t R_{\hat{B}} \geq 0$ and $\lambda^t r^j_{\hat{B}} = 0$ i.e. $\hat{B}$ is optimal for LP(λ) and therefore an adjacent efficient basis.

□

If x_B and $x_{\hat{B}}$ are the Pareto optimal extreme points associated to adjacent efficient bases $B, \hat{B}$, we see from the proof of Lemma 5.6 that both $x_B, x_{\hat{B}}$ are optimal for the same LP(λ). Therefore the edge $\text{conv}(x_B, x_{\hat{B}}) \subset X_{\text{Par}}$.

To check, whether a nonbasic variable x_j at efficient basis B is efficient, we can perform a test.

Theorem 5.7. *Let B be an efficient basis and x_j be nonbasic. All feasible pivots (even with negative pivot elements) with x_j entering are efficient pivots iff*

$$\begin{array}{lrcl} \max & e^t v & & \qquad e = (1, \dots, 1) \\ s.t. & Ry - r^j \delta + Iv & = & 0 \\ & 0 & \leq & y \\ & 0 & \leq & \delta \\ & 0 & \leq & v \end{array}$$

has an optimal value of 0.

Proof: By Definition 5.3 a) x_j is efficient nonbasic variable, iff

$$\begin{array}{lrcl} \min & 0^t \lambda & = & 0 \\ \text{s.t.} & R^t \lambda & \geq & 0 \\ & (r^j)^t \lambda & = & 0 \iff (r^j)^t \lambda \leq 0 \iff (-r^j)^t \lambda \geq 0 \\ & I\lambda & \geq & e \\ & \lambda & \geq & 0 \end{array}$$

has an optimal objective value of 0 (i.e. is feasible).

The dual of this is

$$\begin{array}{lrcl} \max & e^t v & & \\ \text{s.t.} & Ry - r^j \delta + Iv + It & = & 0 \\ & 0 & \leq & y \\ & 0 & \leq & \delta \\ & 0 & \leq & v, t \end{array}$$

But since in an optimal solution of this, t will always be zero, this is exactly

$$\begin{array}{lrcl} \max & e^t v & & \\ \text{s.t.} & Ry - r^j \delta + Iv & = & 0 \\ & y, \delta, v & \geq & 0 \end{array} \qquad \text{(SP)}$$

□

Note that (SP) is always feasible $(y, d, v, t = 0)$.

Therefore we have

- x_j is an efficient nonbasic variable $\iff$ (SP) is bounded
- x_j is an inefficient nonbasic variable $\iff$ (SP) is unbounded

Definition 5.4. Two efficient bases $\overline{B}$ and $\hat{B}$ are called **connected**, if one can be obtained from the other by performing only efficient pivots.

We prove that all efficient bases are connected using parametric programming. Note that a single objective optimal pivot is an efficient pivot.

Theorem 5.8. *All efficient bases are connected.*

Proof: Let $\overline{B}$ and $\hat{B}$ be efficient bases. Let $\overline{\lambda}, \hat{\lambda} \in \operatorname{int} \mathbb{R}^Q_+$ be the weights for which $\overline{B}, \hat{B}$ are optimal for LP($\overline{\lambda}$), LP($\hat{\lambda}$).

Consider the parametric LP with objective

$$\lambda^{*t} C = \Phi \hat{\lambda}^t C + (1 - \Phi) \overline{\lambda}^t C, \quad \Phi \in [0, 1].$$

Let $\hat{B}$ be the starting basis (optimal for $\Phi = 1$). After several parametric programming pivots, we get a basis $\tilde{B}$ optimal for LP($\overline{\lambda}$). (Note that $\lambda^* = \Phi \hat{\lambda} + (1 - \Phi)\overline{\lambda} \in \operatorname{int} \mathbb{R}^Q_+ \quad \forall\, \Phi$.) All intermediate bases are thus optimal for some λ^*, i.e. efficient. All pivots are efficient (see the parametric programming description in Section 5.1).

If $\tilde{B} = \overline{B}$ we are done. Otherwise $\overline{B}$ can be obtained from $\tilde{B}$ by optimal (efficient) pivots.

□

X_{Par} may contain some unbounded edges $U = \{x : x = x^i + \mu r^j,\ \mu \geq 0\}$ where r^j is an extreme ray and x^i is an extreme point of X.

An unbounded edge always starts at an extreme point, which must therefore be Pareto optimal. Let B be the efficient bases associated with that extreme point. Then the unbounded Pareto optimal edge is detected by an efficient nonbasic variable, in which the column contains only nonpositive elements.

We conclude that the set of all Pareto optimal extreme points and unbounded edges can be found by efficient pivots from efficient bases !

This observation is the basis of the multicriteria simplex algorithm.

After the algebra, let's have a look at the geometry:

Definition 5.5. Let $F \subset X$ be a face of X. F is called a **Pareto face**, if $F \subset X_{\text{Par}}$. It is called **maximal Pareto face**, if there is no Pareto face F' of higher dimension s.t. $F \subset F'$.

We now look at the structure of X_{Par}.

Lemma 5.9.

a) Suppose $\exists\, \lambda \in \operatorname{int} \mathbb{R}^Q_+$ *s.t.* $\lambda^t C x = const \quad \forall\, x \in X$ *then* $X_{\text{Par}} = X$.

b) Otherwise $X_{\text{Par}} \subset \bigcup_{t=1}^{T} F_t$, *where* F_t *is a face of* X *and* T *is the number of faces of* X.

Proof:

a) obvious, because $\lambda^t C x = \text{const} \quad \forall\, x \in X$

b) follows from the fact that $X_{\text{Par}} \subset \delta X$ (because $Y_{\text{eff}} \subset \delta Y$ and $C : X \to Y$ is linear) and the fact that $\delta X = \bigcup_{t=1}^{T} F_t$.

□

Now let F be a face of X. Then any $x \in F$ can be written as a convex combination of its extreme points plus a nonnegative combination of extreme rays.

Let $x \in F$ and $x^1, \dots, x^k$ be the extreme points of F, $r^1, \dots, r^p$ the extreme rays of F, then

$$x = \sum_{i=1}^{k} \alpha_i x^i + \sum_{i=1}^{p} \mu_i r^i \qquad 0 \le \alpha_i \le 1, \ \sum \alpha_i = 1, \ 0 \le \mu_i. \tag{5.8}$$

A point in the relative interior of F can be written as

$$x \in \mathrm{ri}(F) \iff x = \sum_{i=1}^{k} \alpha_i x^i + \sum_{i=1}^{p} \mu_i r^i \qquad 0 < \alpha_i < 1, \ \sum \alpha_i = 1, \ 0 \le \mu_i \tag{5.9}$$

(see e.g. [NW88] Chapter I.4, Theorem 4.8)

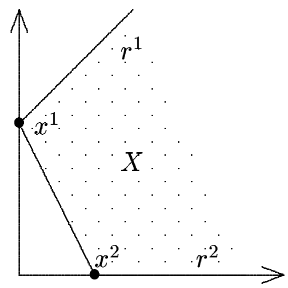


Figure 5.3: A Polyhedron with Extreme Points and Extreme Rays

Suppose that $\emptyset \neq X_{\mathrm{Par}} \neq X$. Then we obtain

Theorem 5.10. *A face $F \subset X$ is a Pareto face iff $\exists\, x^0 \in \mathrm{ri}(F)$ s.t. $x^0 \in X_{\mathrm{Par}}$.*

Proof:

„ $\Longrightarrow$ " is by definition

„ $\Longleftarrow$ " Let $x^0 \in X_{\mathrm{Par}}$. We show that $\exists\, \lambda^0 \in \mathrm{int}\,\mathbb{R}^Q_+$ s.t. F is optimal for $\mathrm{LP}(\lambda^0)$.

First, by 5.3 $\exists\, \lambda^0$ s.t. x^0 solves $\mathrm{LP}(\lambda^0)$, in particular $\mathrm{LP}(\lambda^0)$ is bounded

$\Longrightarrow$ ${\lambda^0}^t C x^i \ge {\lambda^0}^t C x^0 \quad \forall$ extreme points x^i and

${\lambda^0}^t C r^j \ge 0 \quad \forall$ extreme rays r^j

(Note that: $\exists\, j : \lambda^0 C r^j < 0 \iff \mathrm{LP}(\lambda^0)$ is unbounded)

Suppose $\exists\, x^i,\ i \in \{1, \dots, r\}$ s.t. ${\lambda^0}^t C x^i > {\lambda^0}^t C x^0$

$$\begin{aligned} \Longrightarrow {\lambda^0}^t C x^0 &= \sum_{i=1}^{k} \underbrace{\alpha_i}_{>0} \underbrace{{\lambda^0}^t C x^i}_{\ge {\lambda^0}^t C x^0} + \sum_{j=1}^{p} \mu_i \underbrace{{\lambda^0}^t C r^j}_{\ge 0} \\ &> \sum_{i=1}^{k} \alpha_i {\lambda^0}^t C x^0 = {\lambda^0}^t C x^0 \quad \text{↯ Contradiction} \\ \Longrightarrow {\lambda^0}^t C x^i &= {\lambda^0}^t C x^0. \end{aligned}$$

Therefore all extreme points of F are optimal for $\mathrm{LP}(\lambda^0)$. Now, since ${\lambda^0}^t C x^i = {\lambda^0}^t C x^0$ we get $\forall\, r^j$ either $\mu_i = 0$ or ${\lambda^0}^t C r^j = 0$.

$\Longrightarrow$ F is optimal for $\mathrm{LP}(\lambda^0)$.

□

Therefore, X_{Par} is the union of maximally efficient faces, each of which is the set of optimal solutions of LP(λ), for some $\lambda \in \operatorname{int} \mathbb{R}^Q_+$.

If we combine this with the fact, that the set of Pareto optimal extreme points is connected by Pareto optimal edges, we get:

Theorem 5.11. *X_{Par} is connected (therefore Y_{eff} is connected, too).*

Proof: Theorem 5.8 and Theorem 5.10 for X_{Par}. Y_{eff} is connected because X_{Par} is and C is linear, thus continuous.

□

Example 5.2.

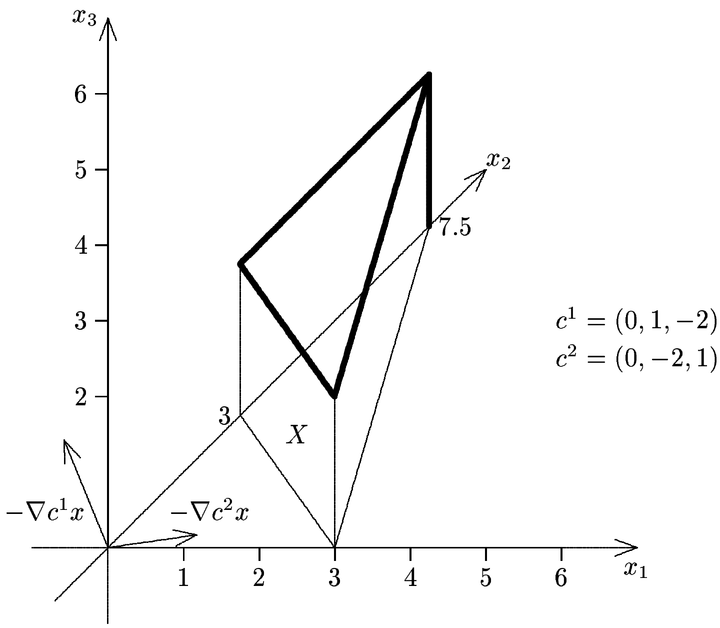


Figure 5.4: X_{Par} for Example 5.2

We use the fact that $x \in X_{\text{Par}} \iff \exists\, \lambda \in \operatorname{int} \mathbb{R}^2_+$ s.t. x solves LP(λ), i.e. $\exists\, c^* = \lambda c^1 + (1-\lambda)c^2$ s.t. x solves LP(λ) with objective c^*. We use the negative gradient of the objective c^* to determine the optimal facet.

So X_{Par} has a 2-dimensional face and a 1-dimensional face, which are maximal Pareto faces.

5.3 A Multicriteria Simplex Algorithm

Consider MCLP:

$$\begin{aligned} \min \quad & Cx \\ \text{s.t.} \quad & Ax = b \qquad\qquad \text{(MCLP)} \\ & x \geq 0 \end{aligned}$$

Then one and only one of the following cases can occur:

1. $X = \emptyset$
2. $X \neq \emptyset$ but $X_{\text{Par}} = \emptyset$
3. $X_{\text{Par}} \neq \emptyset$

Thus, our algorithm will have three phases:

Phase I: Determine an initial extreme point (bfs) or stop with the conclusion that $X = \emptyset$. (This can be done by the usual Phase I simplex method).

Phase II: Determine an initial Pareto optimal extreme point (efficient basis) or stop with the conclusion $X_{\text{Par}} = \emptyset$.

Phase III: Pivot among efficient bases to determine all Pareto optimal extreme points and extreme rays.

<u>Phase II:</u>

After Phase I, we have a feasible point $x^0 \in X$. Then we proceed in two steps:

First, one LP is solved to check if $X_{\text{Par}} = \emptyset$ or the MCLP has a Pareto optimal extreme point. Then a weighted sum LP(λ) is solved, for an appropriate λ, to obtain a Pareto optimal extreme point.

We solve

$$\begin{array}{llcl} \max & e^t y & & \\ \text{s.t.} & Ax & = & b \\ & Cx + Iy & = & Cx^0 \\ & x, y & \geq & 0 \end{array} \qquad \text{(P1)}$$

This problem is always feasible ($x = x^0, y = 0$), so there are two possibilities:

1. The objective is unbounded. Then from Theorem 4.7 of Benson's method $Y_{\text{p-eff}} = \emptyset$. Theorem 5.3 $\Longrightarrow$ $X_{\text{Par}} = \emptyset$.
2. Otherwise the objective is bounded. Let (x^*, y^*) be an optimal solution. From Proposition 4.6 x^* is Pareto optimal. However, we do not know if it is an extreme point of the original MCLP.

So far we know:

(MCLP) has a Pareto optimal solution $\Longleftrightarrow$ (P1) has an optimal solution.

By the duality, (P1) has an optimal solution if and only if

$$\begin{array}{llcl} \min & u^t b + w^t C x^0 & & \\ \text{s.t.} & u^t A + w^t C & \geq & 0 \\ & w & \geq & e \\ & u & \gtrless & 0 \end{array} \qquad \text{(P2)}$$

has an optimal solution u^*, w^* with $u^{*t} b + w^{*t} C x^0 = e^t y^*$ (see Lemma 5.2).

Therefore u^* is an optimal solution of

$$\begin{aligned} \min \quad & u^t b \\ \text{s.t.} \quad & u^t A \geq -w^{*t} C \end{aligned} \tag{P3}$$

which is just (P2) for $w = w^*$ fixed.

Therefore the dual of (P3)

$$\begin{aligned} \max \quad & -w^{*t} C x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{P4}$$

has an optimal solution, and therefore an optimal extreme point, which by Theorem 5.3 is Pareto optimal.

So we have (in addition to Lemma 5.2):

The MCLP has an efficient solution if and only if (P2) has an optimal solution. So in Phase II

- we solve (P2), if (P2) is unbounded or infeasible, $X_{\text{Par}} = \emptyset$.
- otherwise we use the optimal solution w^* of (P2) and solve (P4) to obtain an initial Pareto optimal extreme point.

We can now summarize the multicriteria Simplex algorithm, where we use the following notation:

LB is a list of bases to be processed

LPX is a list of Pareto optimal extreme points

LPU is a list of Pareto optimal unbounded edges

Multicriteria Simplex Algorithm

(1) a) Solve the problem

$$\begin{aligned} \min \quad & e^t \hat{x} \\ \text{s.t.} \quad & Ax + I\hat{x} = b \\ & x, \hat{x} \geq 0 \end{aligned}$$

b) If the optimal solution is nonzero, STOP, $X = \emptyset$.
Otherwise go to (2) with a feasible solution x^0 of MCLP.

(2) a) Solve the problem

$$\begin{aligned} \min \quad & u^t b + w^t C x^0 \\ \text{s.t.} \quad & u^t A + w^t C \geq 0 \\ & w \geq e \end{aligned}$$

b) If the optimal solution is unbounded, STOP, $X_{\text{Par}} = \emptyset$.
Otherwise let (u^*, w^*) be an optimal solution, go to c)

c) Solve

$$\begin{aligned} \min \quad & {w^*}^t Cx \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

Add the optimal basis to LB, the optimal extreme point in LPX, go to ③.

③ a) If LB = $\emptyset$ STOP: all Pareto optimal extreme points and unbounded edges are found. Otherwise choose a basis B in LB, remove it from LB, go to b)

b) For all nonbasic variables x_j for basis B solve

$$\begin{aligned} \max \quad & e^t v \\ \text{s.t.} \quad & +Ry - r^j\delta + Iv = 0 \\ & 0 \leq y, \delta, v \end{aligned}$$

and do the following steps:

i) Add all efficient bases adjacent to B to LB, if they are new.

ii) Add all extreme points corresponding to adjacent efficient bases to LPX, if new.

iii) Add all unbounded Pareto optimal edges emanating from x_B to LPU (unbounded edges are characterized by an (x_B, r^j) pair).

c) Go to ③ a)

Remark.

1. The list LPX can be determined after termination of the algorithm from LB, if a copy is kept till the end.
2. Because the Simplex algorithm may require an exponential number of steps (in terms of problem size m, n, Q), the same is true for a multicriteria Simplex algorithm.
3. The test for nonbasic variable efficiency can be replaced by several other more efficient, but more complicated methods (see Steuer, 1985, [Ste85] for a survey).
4. The question, whether a polynomial time MCLP algorithm is possible depends on the number of Pareto optimal extreme points. There may exist exponentially many. Two results recently published are interesting:

   Benson, 1997, [Ben97]; numerical tests (10 random examples with inequality constraints)

n	m	Q	#Pareto optimal points	
30	25	4	average	7245.9
50	50	4	average	83780.6
60	50	4	$\geq$	200000

Küfer, 1998, [Küf98]

The expected number of Pareto optimal extreme points for a certain family of randomly generated MCLP is polynomial in n, m, Q.

However, examples with all (i.e. exponentially many) extreme points Pareto optimal can be constructed for all (n, m, Q)-choices.

We close this section with an example for the multicriteria Simplex algorithm.

Example 5.3.

$$\begin{array}{llllll}
\min & -x_1 & -2x_2 & & & \\
\min & -x_1 & & +2x_3 & & \\
\min & x_1 & & -x_3 & & \\
\text{s.t.} & x_1 & +x_2 & & \leq & 1 \\
 & & x_2 & & \leq & 2 \\
 & x_1 & -x_2 & +x_3 & \leq & 4
\end{array}$$

① $x_1 = x_2 = x_3 = 0$ is an initial feasible extreme point

② a) Solve

$$\begin{array}{ll}
\min & u_1 + 2u_2 + 4u_3 \\
\text{s.t.} & u^t \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 1 & 0 & 0 & 1 \end{pmatrix} + w^t \begin{pmatrix} -1 & -2 & 0 & 0 & 0 & 0 \\ -1 & 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 \end{pmatrix} \geq 0 \\
 & \hspace{22em} w \geq e
\end{array}$$

the constraints are equivalent to

$$\begin{array}{rcl}
-u^t \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} - w^t \begin{pmatrix} -1 & -2 & 0 \\ -1 & 0 & 2 \\ 1 & 0 & -1 \end{pmatrix} + Is & = & 0 \\
w - Iz & = & e \\
u, w & \geq & 0
\end{array}$$

Phase I (artificial variables)

0	0	0	−1	−1	−1	0	0	0	1	1	1	0	0	0	−3
−1	0	−1	1	1	−1	1	0	0	0	0	0	0	0	0	0
−1	−1	1	2	0	0	0	1	0	0	0	0	0	0	0	0
0	0	−1	0	−2	[1]	0	0	1	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	−1	0	0	1	0	0	1
0	0	0	0	1	0	0	0	0	0	−1	0	0	1	0	1
0	0	0	0	0	1	0	0	0	0	0	−1	0	0	1	1

After 5 Simplex operations obtain

0	1	5	0	0	0	0	1	0	1	0	0	original
1	2	4	0	0	0	0	0	0	0	0	0	objective

0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0
0	1	−2	0	0	0	1	−1	0	−1	1	−1	1	−1	1	1
0	0	0	1	0	0	0	0	0	−1	0	0	1	0	0	1
0	0	0	0	0	1	0	0	0	0	0	−1	0	0	−1	1
[1]	1	−1	0	0	0	0	−1	0	−1	0	0	1	0	0	2
0	0	−1	0	0	0	0	0	1	0	−2	1	0	1	−1	1
0	0	0	0	1	0	0	0	0	0	−1	1	0	1	0	1

We delete the artificial variables, because the LP is feasible, replace it with the original objective and make 1 Pivot step (Pivot element indicated) to get an optimal solution.

b) An optimal solution is $w^* = (1, 1, 1)$

c) Solve $\min {w^*}^t Cx, \ x \in X$

−1	−2	0	0	0	0	0
−1	0	2	0	0	0	0
1	0	−1	0	0	0	0
1	[1]	0	1	0	0	1
0	1	0	0	1	0	2
1	−1	1	0	0	1	4

$\lambda R = (-1, \overset{\downarrow}{-2}, 1)$

1	0	0	2	0	0	2
−1	0	2	0	0	0	0
1	0	−1	0	0	0	0
1	1	0	1	0	0	1
−1	0	0	−1	1	0	1
2	0	1	1	0	1	5

$\lambda R = (1, 1, 2) \geq 0$

LB $= \{(2, 5, 6)\}$ LPX $= \{x^1 = (0, 1, 0)\}$

(3) a) $B = (2, 5, 6)$ LB $= \emptyset$

Nonbasic variable x_1. Solve the LP of the following tableau.

1	1	2	−1	0	0	0	0
1	0	2	−1	1	0	0	0
−1	2	0	[1]	0	1	0	0
1	−1	0	−1	0	0	1	0

1	3	2	0	0	1	0	0
0	2	2	0	1	1	0	0
−1	2	0	1	0	1	0	0
0	1	0	0	0	1	1	0

The LP has an optimal solution, x_1 is efficient. From now on we will not display the right hand side column, it is always 0.

Nonbasic variable x_3.

$$\begin{array}{rrrrrrr}
1 & 1 & 2 & -1 & 0 & 0 & 0 \\
\hline
1 & 0 & 2 & 0 & 1 & 0 & 0 \\
-1 & 2 & 0 & -2 & 0 & 1 & 0 \\
1 & -1 & 0 & \boxed{1} & 0 & 0 & 1
\end{array}$$

x_4 is efficient.

Nonbasic variable x_4.

$$\begin{array}{rrrrrrr}
1 & 1 & 2 & -2 & 0 & 0 & 0 \\
\hline
1 & 0 & 2 & -2 & 1 & 0 & 0 \\
-1 & 2 & 0 & 0 & 0 & 1 & 0 \\
1 & -1 & 0 & 0 & 0 & 0 & 1 \\
 & & & \uparrow & & &
\end{array}$$

LP unbounded, x_4 not efficient.

i) x_1 entering $\implies$ x_2 leaving, basis $(1, 5, 6)$.

x_3 entering $\implies$ x_6 leaving, basis $(2, 3, 5)$.

$\text{LB} = \{(1, 5, 6), (2, 3, 5)\}$.

ii)

$$T(1,5,6) \quad \begin{array}{rrrrrr|r}
0 & -1 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 2 & 1 & 0 & 0 & 1 \\
0 & -1 & -1 & -1 & 0 & 0 & -1 \\
\hline
1 & 1 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 2 \\
0 & 2 & 1 & -1 & 0 & 1 & 3
\end{array} \qquad x^2 = (1, 0, 0)$$

$$T(2,3,5) \quad \begin{array}{rrrrrr|r}
1 & 0 & 0 & 2 & 0 & 0 & 2 \\
-5 & 0 & 0 & -2 & 0 & -2 & -10 \\
3 & 0 & 0 & 1 & 0 & 1 & 5 \\
\hline
1 & 1 & 0 & 1 & 0 & 0 & 1 \\
-1 & 0 & 0 & -1 & 1 & 0 & 1 \\
2 & 0 & 1 & 1 & 0 & 1 & 5
\end{array} \qquad x^3 = (0, 1, 5)$$

$\text{LPX} = \{x^1, x^2, x^3\}$.

a) $B = (1, 5, 6) \qquad \text{LB} = \{(2, 3, 5)\}$

Nonbasic variable x_2 $\implies$ x_1 leaves, leads to $(2, 5, 6)$, not new.

Nonbasic variable x_3.

−1	1	1	−1	0	0	0		0	3	2	−3	0	1	0
−1	0	1	0	1	0	0		0	2	2	−2	1	1	0
$\boxed{1}$	2	1	−2	0	1	0		1	2	1	−2	0	1	0
−1	−1	−1	1	0	0	1		0	1	0	−1	0	0	1
											↑			

LP unbounded, x_3 is not efficient.

Nonbasic variable x_4.

−1	1	1	−1	0	0	0		0	3	1	−2	0	1	0
−1	0	1	−1	1	0	0		0	2	2	−2	1	1	0
$\boxed{1}$	2	1	−1	0	1	0		1	2	1	−1	0	1	0
−1	−1	−1	1	0	0	1		0	1	0	0	0	1	1
											↑			

LP unbounded, x_4 is not efficient.

No new basis to add, go to a)

a) $B = (2, 3, 5)$ LB $= \emptyset$

Nonbasic variable x_1.

−1	1	−1	1	0	0	0		0	3	−1	0	1	0	0
$\boxed{1}$	2	0	−1	1	0	0		1	2	0	−1	1	0	0
−5	−2	−2	5	0	1	0		0	8	−2	0	5	1	0
3	1	1	−3	0	0	1		0	−5	0	9	−3	0	1
										↑				

x_1 is not efficient.

Nonbasic variable x_4.

−1	1	−1	−1	0	0	0		2	2	0	−2	0	0	1
1	2	0	−2	1	0	0		1	2	0	−2	1	0	0
−5	−2	−2	2	0	1	0		1	0	0	0	0	1	2
3	1	$\boxed{1}$	−1	0	0	1		3	1	1	−1	0	0	1

x_4 is not efficient.

Nonbasic variable $x_6 \Longrightarrow x_3$ leaves, leads back to $(2, 5, 6)$.

No new basis to add.

a) LB $= \emptyset$

STOP

We found 3 efficient bases and 3 Pareto optimal extreme points, with the following structure:

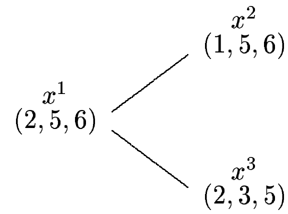


Figure 5.5: Efficient Bases and Corresponding Extreme Points

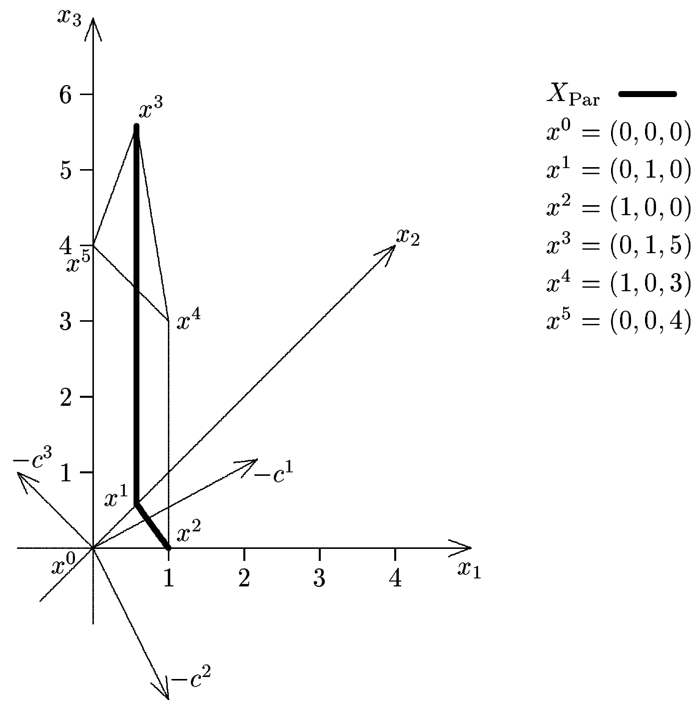


Figure 5.6: Feasible and Pareto Set in Example 5.3

5.4 Identifying Scalarization Vectors and Pareto Faces

The set $\Lambda = \{\lambda \in \text{int}\,\mathbb{R}^Q_+ : \sum \lambda_i = 1\}$ can be subdivided into regions, which correspond to those weighting vectors λ, which make a certain face Pareto optimal. I.e. for each Pareto face $F \quad \exists\, \Lambda_F \subset \Lambda$ s.t. F is optimal for LP(λ) for all $\lambda \in \Lambda_F$.

First assume $\lambda C x$ is bounded over $X \quad \forall\, \lambda \in \Lambda$.

Let F be a Pareto face, and $x^i,\ i = 1, \ldots, t$ be the set of all extreme points of F. Because F is Pareto, from the proof of Theorem 5.10 $\exists\, \lambda_F \in \Lambda$ s.t. F solves LP(λ_F). Thus $x^1, \ldots, x^t$ solve LP(λ_F).

Hence we can apply the optimality conditions. Let R^i be the reduced cost matrix of a basis associated to x^i. Then x^i is optimal $\iff \lambda^t R^i \geq 0$.

Therefore the face is optimal iff $\lambda^t R^i \geq 0,\ i = 1, \ldots, t$.

Proposition 5.12. *The set of all λ s.t. Pareto face F solves LP(λ) is defined by the system* $\sum_{i=1}^{Q} \lambda_i = 1,\ \lambda_i \geq 0,\ \lambda^t R^i \geq 0 \quad \forall\, x^i$ *extreme points of F.*

Example 5.4. In Example 5.3 let us consider the Pareto face $\text{conv}(x^1, x^2)$.

x^1 corresponds to basis $(2,5,6)$, $R^1 = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 2 & 0 \\ 1 & -1 & 0 \end{pmatrix}$

x^2 corresponds to basis $(1,5,6)$, $R^2 = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 2 & 1 \\ -1 & -1 & -1 \end{pmatrix}$

So we get the system $\lambda^t R^1 \geq 0,\ \lambda^t R^2 \geq 0,\ e^t\lambda = 1,\ \lambda \gg 0$

$$\begin{array}{rrrcl} \lambda_1 & -\lambda_2 & +\lambda_3 & \geq & 0 \\ & 2\lambda_2 & -\lambda_3 & \geq & 0 \\ 2\lambda_1 & & & \geq & 0 \\ -\lambda_1 & +\lambda_2 & -\lambda_3 & \geq & 0 \\ & 2\lambda_2 & -\lambda_3 & \geq & 0 \\ \lambda_1 & +\lambda_2 & -\lambda_3 & \geq & 0 \\ \lambda_1 & +\lambda_2 & +\lambda_3 & = & 1 \\ & & \lambda_1,\lambda_2,\lambda_3 & \geq & 0 \end{array} \quad \text{or} \quad \begin{array}{rrrcl} \lambda_1 & -\lambda_2 & +\lambda_3 & = & 0 \\ & 2\lambda_2 & -\lambda_3 & \geq & 0 \\ \lambda_1 & +\lambda_2 & -\lambda_3 & \geq & 0 \\ \lambda_1 & +\lambda_2 & +\lambda_3 & = & 1 \\ & & \lambda_1,\lambda_2,\lambda_3 & \geq & 0 \end{array}$$

or, eliminating λ_3, $\lambda_2 = \frac{1}{2},\ 0 < \lambda_1 < \frac{1}{2}$.

In total λ's corresponding to Pareto faces are as shown in Figure 5.7.

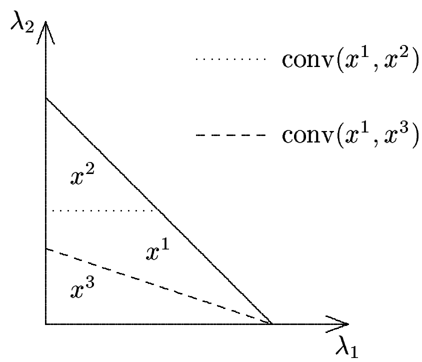


Figure 5.7: Weights to obtain Pareto Faces in Example 5.3

If X is not bounded, it may happen that X_{Par} contains unbounded faces, i.e. $\exists\ \lambda \in \Lambda$ such that $\text{LP}(\lambda)$ is unbounded.

In this case there exists, additionally to $\Lambda_F \subset \Lambda$ for all bounded Pareto faces F, a subset $\Lambda_\circ \subset \Lambda$ with $\Lambda_\circ = \{\lambda \in \Lambda : \text{LP}(\lambda) \text{ is unbounded}\}$.

Let us finally turn to the determination of maximal Pareto faces.

Let B be an efficient basis and N^p be the nonbasic variables, which allow feasible pivots. Let $J \subset N^p$ then we have:

Proposition 5.13. *All variables in J are nonbasic efficient variables* $\iff$

$$\begin{array}{lrcl} \max & e^t v & & \\ s.t. & Ry - R^J\delta + Iv & = & e \\ & y, \delta, v & \geq & 0 \end{array} \qquad \text{P}(J)$$

has an optimal solution, where R^J denotes the columns of R pertaining to variables in J.

Proof: Exercise 40.

□

Let us call $J \subset N^p$ a **maximal set of efficient nonbasic variables**, if $\nexists\ J' \subset N^p,\ J \subset J'$ s.t. P(J') has an optimal solution.

Now let $B^i,\ i = 1, \ldots, k$ be all efficient bases and $J^{i,j},\ i = 1, \ldots, k,\ j = 1, \ldots, l$ be all maximal index sets of efficient nonbasic variables for basis B^i. Furthermore let $E^t = (B^i, r^t),\ t = 1, \ldots, k'$ denote unbounded Pareto edges, where r^t is an extreme ray. Let $Q^{i,j} = B^i \cup J^{i,j}$ and select a minimal number of index sets representing all $Q^{i,j}$.
I.e. choose $U^1, \ldots, U^o$ s.t.

1) For each $Q^{i,j}\quad \exists\ U^s$ s.t. $Q^{i,j} \subset U^s$
2) For each $U^s\quad \exists\ Q^{i,j}$ s.t. $U^s = Q^{i,j}$
3) $\nexists\ U^s, U^{s'}, s \neq s'$ s.t. $U^s \subset U^{s'}$

Then for $s \in \{1, \ldots, o\}$ let

$$
\begin{aligned}
I_b^s &= \{i \in \{1, \ldots k\} \mid B^i \subset U^s\} && (5.10)\\
I_u^s &= \{t \in \{1, \ldots k'\} \mid B^i \subset U^s\} && (5.11)
\end{aligned}
$$

and define

$$X^s = \{x \mid x = \sum_{i \in I_b^s} \alpha_i x^i + \sum_{t \in I_u^s} \mu_j r^j, \sum \alpha_i = 1,\ \alpha_i \geq 0,\ \mu_j \geq 0\} \qquad (5.12)$$

Then we have

Theorem 5.14. $X^s \subset X_{\text{Par}} \quad \forall\ s = 1, \ldots, o.$

Proof: By definition $\exists\ Q^{i,j}$ s.t. $Q^{i,j} = U^s$.
$\implies P(Q^{i,j} \setminus B^i)$ has an optimal solution
$\implies$ Its dual

$$
\begin{aligned}
\min\quad & e^t \lambda \\
\text{s.t.}\quad & R\lambda \geq 0 \\
& -R^J \lambda \geq 0 \\
& \lambda \geq e
\end{aligned}
\qquad \text{DP}(J)
$$

has an optimal solution λ^*
$\implies$ all $x \in X^s$ are optimal solutions of LP(λ^*)
$\implies X^s \subset X_{\text{Par}}$.

□

Theorem 5.15. *If* $x \in X_{\text{Par}} \implies \exists\ s \in \{1, \ldots, o\}$ *s.t.* $x \in X^s$.

Proof: Let $x \in X_{\text{Par}}$.
$\implies \exists$ maximal Pareto face F s.t $x \in F$.
Choose an extreme point x^i of F and let B^i be a basis associated with x^i.

Let I be the index set of efficient bases adjacent to B^i and $J^\circ := \left\{ \bigcup_{l \in I} B^l \right\} \setminus B^i$.

Because all B^l are efficient and adjacent to B^i, J° is a set of efficient nonbasic variables at B^i.

$\Longrightarrow$ P(J°) has an optimal solution.

$\Longrightarrow$ $\exists$ maximal index set of efficient nonbasic variables J s.t. $J^\circ \subset J$.

Then by the further construction of index sets

$\Longrightarrow$ $x \in X^s$ for some $s \in \{1, \dots, o\}$.

$\square$

If all efficient bases are nondegenerate, X^s are exactly the maximal Pareto faces of X. Otherwise some X^s may not be maximal.

This method is from Isermann, 1977, [Ise77].

Example 5.5. In Example 5.3:

$$\begin{array}{lll}
B^1 = \{2,5,6\} & J^{1,1} = \{1\} & J^{1,2} = \{3\} \\
B^2 = \{1,5,6\} & J^{2,1} = \{2\} & \\
B^3 = \{2,3,5\} & J^{3,1} = \{6\} & \\
Q^{1,1} = \{1,2,5,6\} & Q^{1,2} = \{3,2,5,6\} & \\
Q^{2,1} = \{1,2,5,6\} & & \\
Q^{3,1} = \{2,3,5,6\} & & \\
U^1 = \{1,2,5,6\} & U^2 = \{2,3,5,6\} & \\
I_b^1 = \{1,2\} & I_b^2 = \{1,3\} &
\end{array}$$

There are no unbounded edges.

$$\begin{aligned}
X^1 &= \{x \mid \alpha_1 x^1 + \alpha_2 x^2 : \alpha_1 + \alpha_2 = 1,\ \alpha_i \geq 0\} = \text{conv}(x^1, x^2) \\
X^2 &= \{x \mid \alpha_1 x^1 + \alpha_2 x^3 : \alpha_1 + \alpha_2 = 1,\ \alpha_i \geq 0\} = \text{conv}(x^1, x^3)
\end{aligned}$$

Thus $X_{\text{Par}} = X^1 \cup X^2$, as expected.

5.5 Exercises to Chapter 5

38. Consider the parametric LP

$$\begin{array}{lrcl}
\min & \lambda(-2x_1 + x_2) + (1-\lambda)(-4x_1 - 3x_2) & & \\
\text{s.t.} & x_1 + 2x_2 & \leq & 10 \\
 & x_1 & \leq & 5 \\
 & x_1, x_2 & \geq & 0
\end{array}$$

Solve the problem with the three phase algorithm of Section 5.1. Determine X_{Par}, Y_{eff}. Illustrate the results graphically.

39. a) Give an example of an MCLP s.t. X_{Par} is a singleton, although X is full dimensional.

b) It is possible that some objectives are unbounded, yet $X_{\text{Par}} \neq \emptyset$. Show this behaviour for the MCLP

$$\begin{array}{lrcl} \min & x_1 + 2x_2 & & \\ \min & -2x_2 & & \\ \text{s.t.} & -x_1 + x_2 & \leq & 3 \\ & x_1 + x_2 & \geq & 3 \\ & x_1, x_2 & \geq & 0. \end{array}$$

What can you say about X_{Par} in this case ?

40. Let $J \subset N$ be an index set of nonbasic variables at efficient basis B.

Show that each variable x_j, $j \in J$ is efficient if and only if the problem

$$\begin{array}{lrcl} \max & e^t v & & \\ \text{s.t.} & Ry - R^J \delta + Iv & = & e \\ & y, \delta, v & \geq & 0 \end{array}$$

has an optimal solution. Here R^J is the part of R pertaining to variables x_j, $j \in J$.

(Hint: Take the definition of nonbasic variable efficiency and look at the dual of the above LP.)

41. A basis B is called weakly efficient, iff B is an optimal basis of LP(λ) for some $\lambda \in \mathbb{R}^Q_+ \backslash \{0\}$. A feasible pivot with entering nonbasic variable x_j is called weakly efficient if the basis obtained is weakly efficient. Prove the following theorem:

Let x_j be nonbasic at weakly efficient basis B. Then all feasible pivots with x_j entering are weakly efficient $\iff$ the subproblem

$$\begin{array}{lrcl} \max & v & & \\ \text{s.t.} & Ry - r^j \delta + ev & \geq & 0 \\ & y, \delta, v & \geq & 0 \end{array}$$

has an optimal objective value of 0.

42. Solve the MCLP

$$\begin{array}{lrcl} \min & -3x_1 - x_2 & & \\ \min & x_1 - 2x_2 & & \\ \text{s.t.} & 3x_1 + 2x_2 & \geq & 6 \\ & x_1 & \leq & 10 \\ & x_2 & \leq & 3 \\ & x_1, x_2 & \geq & 0 \end{array}$$

using the multicriteria simplex algorithm.

43. Determine, for each Pareto optimal extreme point of the MCLP in Exercise 42, the set of all λ, s.t. the extreme point solves LP(λ).

Chapter 6

Other Optimality Concepts

In this chapter we study some optimality concepts different from (strict, weak, proper) Pareto optimality.

6.1 Lexicographic Optimization

Here we consider problems of the type $(X, f, \mathbb{R}^Q)/\text{id}/(\mathbb{R}^Q, <_{\text{lex}})$ or, in other words

$$\underset{x \in X}{\text{lexmin}}(f_1(x), \dots, f_Q(x)) \tag{6.1}$$

Recall that $y^1 <_{\text{lex}} y^2$ iff $y^1_q < y^2_q$ where $q = \min\{i : y^1_i \neq y^2_i\}$.

Lemma 6.1. *Let $x \in X$ be such that $f(x) \leq_{\text{lex}} f(y) \quad \forall\, y \in X$. Then $x \in X_{\text{Par}}$.*

Proof: Suppose $x \notin X_{\text{Par}} \implies \exists\, y \in X$ s.t. $f(y) < f(x)$.
Let $q := \min\{i : f_i(y) < f_i(x)\}$. Then $f_i(x) = f_i(y) \quad \forall\, i = 1, \dots, q-1$ and $f_q(x) < f_q(y)$.
Therefore $f(x) <_{\text{lex}} f(y)$ ↯ Contradiction

□

Because of the ranking of the objectives, we can solve a lexicographic program sequentially, with one objective at a time and using optimal values as constraints.

① Define $X_1 := X,\ i := 1$

② Solve $\min_{x \in X_i} f_i(x)$ (P_i)

③ a) If P_i has a unique solution x^*_i, STOP,
x^*_i is the optimal solution of the lexicographic problem

b) If (P_i) is unbounded, STOP, the lexicographic problem is unbounded.

c) If $i = Q$ and (P_Q) has an optimal solution, STOP.
The set of optimal solutions is $\{x \in X_Q : f_Q(x) = \min_{x \in X_Q} f_Q(x)\}$.

d) Otherwise let $X_{i+1} := \{x \in X_i : f_i(x) = \min_{x \in X_i} f_i(x)\},\ i := i + 1$ and go to ②

Note that, if all f_i are continuous, (P_k) unbounded implies that all objectives are unbounded, and all problems (P_i) are unbounded for $i < k$, too. Furthermore (P_{k+1}) is not defined.

Proposition 6.2. *If x is a unique solution of (P_k), $k < Q$, or if x is a solution of (P_Q) then $f(x) \leq_{\text{lex}} f(y) \quad \forall\, y \in X$.*

Proof: Suppose $\exists\, y \in X$ s.t. $f(y) <_{\text{lex}} f(x)$.

Because x is a solution of (P_i) $\quad i = 1, \ldots, k \implies f_i(y) = f_i(x) \quad \forall\, i = 1, \ldots, k$.

Therefore, if $k < Q$ $\quad f_{k+1}(y) < f_{k+1}(x)$ must hold, contradicting uniqueness of x, or, if $k = Q$ we have $f(x) = f(y)$ contradicting the choice of y.

□

Note also, that if x is a unique solution of a problem (P_i) then $x \in X_{\text{s-Par}}$. Otherwise there would exist $y \in X$ s.t. $f_i(y) \leq f_i(x) \quad \forall\, i = 1, \ldots, Q$, which by Pareto optimality of X could only hold with $f(y) = f(x)$, and thus by uniqueness $y = x$.

Proposition 6.3. *If x is a unique solution of (P_k) for some $k \in \{1, \ldots, Q\}$, then $x \in X_{\text{s-Par}}$.*

We may also choose an arbitrary order of the objectives and apply lexicographic optimization. Let $\pi : \{1, \ldots, Q\} \to \{1, \ldots, Q\}$ be a permutation and consider the permutation of the objective function $(f_{\pi(1)}, \ldots, f_{\pi(Q)})$.

As in Lemma 6.1 we can show that a solution of $(X, f, \mathbb{R}^Q)/\pi/(\mathbb{R}^Q, <_{\text{lex}})$ is Pareto optimal. We denote by $\Pi(Q)$ the set of all permutations of $\{1, \ldots, Q\}$ and by $X_{\Pi(Q)}$ the set of all solutions of permuted lexicographic problems.

Proposition 6.4. $X_{\Pi(Q)} \subset X_{\text{Par}}$.

Example 6.1. The inclusion in Proposition 6.4 is strict in general.

Let $X = [0, 1]$, $f_1(x) = x$, $f_2(x) = 1 - x$

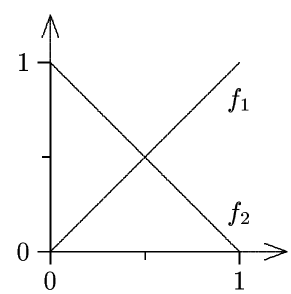


Figure 6.1: Illustration of Example 6.1

Here $X_{\text{Par}} = X$.

The solution of $([0, 1], f, \mathbb{R}^2)/\text{id}/(\mathbb{R}^2, <_{\text{lex}})$ is $x = 0$.

The solution of $([0, 1], f, \mathbb{R}^2)/\pi = (2, 1)/(\mathbb{R}^2, <_{\text{lex}})$ is $x = 1$.

Therefore $X_{\Pi(Q)} = \{0, 1\}$, and $X_{\Pi(Q)} \subset X_{\text{Par}}$.

Also, because of uniqueness, $X_{\Pi(Q)} \subset X_{\text{s-Par}}$, and again the inclusion is strict, as $X_{\text{Par}} = X_{\text{s-Par}}$.

Note that finding $X_{\Pi(Q)}$ is usually not a good approach. It involves solving $|\Pi(Q)| = Q!$ lexicographic problems. But if X is a finite set, finding $X_{\Pi(Q)}$ can be done in time polynomial in $|X|$ and Q.

6.2 Max-Ordering Optimization

The second problem type we consider is $(X, f, \mathbb{R}^Q)/\max/(\mathbb{R}, \leq)$ or

$$\min_{x \in X} \max_{i=1,\dots,Q} f_i(x) \tag{6.2}$$

Let X_{MO} denote the solution set of this problem.

What are the relations to Pareto optimality?

Proposition 6.5. *A solution of* $\min_{x \in X} \max_{i=1,\dots,Q} f_i(x)$ *is weakly Pareto optimal but not necessarily Pareto optimal.*

Proof: Exercise 45.

□

From Proposition 6.5 $X_{\text{MO}} \subset X_{\text{w-Par}}$.

Let us now assume that $\inf_{x \in X} f_i(x) > -\infty \quad \forall\, i = 1, \dots, Q$. Then let $y_i^{00} < \inf_{x \in X} f_i(x)$, and note that X_{Par} is the same for objectives $(f_1, \dots, f_Q)$ and $(f_1 - y_1^{00}, \dots, f_Q - y_Q^{00})$, because this is just a translation of Y.

We have already shown (see Theorem 4.12):

Proposition 6.6. $x^* \in X_{\text{w-Par}} \iff \exists\, \lambda \in \operatorname{int} \mathbb{R}^Q_+$ *s.t.* x^* *solves* $\min_{x \in X} \max_{i=1,\dots,Q} \lambda_i (f_i(x) - y_i^{00})$.

Therefore $X_{\text{w-Par}}$ can be determined through the solution of Max-ordering problems.

Concerning Pareto points we obtain:

Proposition 6.7. $X_{\text{MO}} \cap X_{\text{Par}} \neq \emptyset$ *and* $X_{\text{MO}} \subset X_{\text{Par}}$ *if* $|X_{\text{MO}}| \leq 1$.

Proof: Let $x \in X_{\text{MO}}$ and suppose $x \notin X_{\text{Par}}$.

$\Longrightarrow \exists\, y \in X_{\text{Par}}$ s.t. $f_i(y) \leq f_i(x) \quad \forall\, i = 1, \dots, Q$ and $f_k(y) < f_k(x)$ for some k.

$\Longrightarrow \max_{i=1,\dots,Q} f_i(y) \leq \max_{i=1,\dots,Q} f_i(x)$

From optimality of x, equality holds, and $y \in X_{\text{MO}}$.

□

More about this intersection in Section 6.3.

Next, we show that the Max-ordering problem can be solved as a single objective problem, and optimal solutions have a geometric characterization like Theorem 2.16.

$$\min_{x \in X} \max_{i=1,\dots,Q} f_i(x) \tag{6.2}$$

$$\Longleftrightarrow$$

$$\begin{aligned} \min \quad & z \\ \text{s.t.} \quad & f_i(x) \le z \qquad i = 1, \dots, Q \\ & x \in X \end{aligned} \tag{6.3}$$

Using level sets $L^i_{\le}(z) = \{x \in X : f_i(x) \le z\}$ we obtain:

Proposition 6.8. $x \in X_{\text{MO}} \iff \bigcap_{i=1}^{Q} L^i_{\le}(\max_i f_i(x)) \neq \emptyset$ *and for all* $z < \max_i f_i(x) \quad \bigcap_{i=1}^{Q} L^i_{\le}(z) = \emptyset$.

Since we consider only the worst objective for minimization, it may happen, that this is the same for all $x \in X$, i.e. the objective is considerably worse than all others.
We use the ideal point y^0 again. Let $x^i,\ i = 1, \dots, Q$ be such that $y^0_i = f_i(x^i)$.
Then

$$f_q(x^q) = y^0_q \le \min_{x \in X} \max_{i=1,\dots,Q} f_i(x) \le \max_{i=1,\dots,Q} f_i(x^q) \tag{6.4}$$

Proposition 6.9. *If* $\exists\ x^q$ *with* $f_q(x^q) = y^0_q$ *such that* $f_i(x^q) \le y^0_q \quad \forall\ i = 1, \dots, Q$ *then* $x^q \in X_{\text{MO}}$ *and the optimal objective value is* y^0_q.

Proof: $\quad f_i(x^q) \le y^0_q \quad \forall\ i = 1, \dots, Q$
$\implies \max_{i=1,\dots,Q} f_i(x^q) \le y^0_q$. This implies that (6.4) holds with equalities, i.e.

$$f_q(x^q) = y^0_q = \min_{x \in X} \max_{i=1,\dots,Q} f_i(x).$$

□

In this case, the minimum of one objective is worse than the value of all others for at least one minimizer of this objective.
(6.4) also implies

$$\max_{q=1,\dots,Q} f_q(x^q) \le \min_{x \in X} \max_{i=1,\dots,Q} f_i(x) \le \min_{q=1,\dots,Q} \min_{x^q \in X_q} \max_{i=1,\dots,Q} f_i(x^q) \tag{6.5}$$

where $X_q = \{x \in X : f_q(x) = \min_{x \in X} f_q(x)\}$.
This yields lower and upper bounds.
Now let

$$\Lambda = \{\lambda : \sum_{i=1}^{Q} \lambda_q = 1,\ \lambda_q \ge 0\}. \tag{6.6}$$

Proposition 6.10. $\max_{\lambda \in \Lambda} \min_{x \in X} \sum_{i=1}^{Q} \lambda_i f_i(x) \le \min_{x \in X} \max_{i=1,\dots,Q} f_i(x).$

Proof: $\sum_{i=1}^{Q} \lambda_i f_i(x) \le \sum_{i=1}^{Q} \lambda_i \max_{i=1,\dots,Q} f_i(x) \le \max_{i=1,\dots,Q} f_i(x)$ holds for each x and λ.

$$\Longrightarrow \min_{x\in X} \sum_{i=1}^{Q} \lambda_i f_i(x) \le \min_{x\in X} \max_{i=1,\dots,Q} f_i(x)$$

□

Note that the $\max_{\lambda\in\Lambda} \min_{x\in X}$ can also be reversed.

6.3 Lexicographic Max-Ordering

Lexicographic Max-ordering is a combination of max-ordering (worst objective to be minimized) and lexicographic optimization.

Definition 6.1.

a) For $y \in \mathbb{R}^Q$ let $\text{sort}(y) := (\text{sort}_1(y), \dots, \text{sort}_Q(y))$ with $\text{sort}_1(y) \ge \dots \ge \text{sort}_Q(y)$.

b) $x^* \in X$ is called a **lexicographic Max-ordering solution** (lex-MO solution) if

$$\text{sort}(f(x^*)) \le_{\text{lex}} \text{sort}(f(x)) \quad \forall\, x \in X. \tag{6.7}$$

A lexicographic max-ordering problem is denoted, in the classification, by

$$(F, f, \mathbb{R}^Q) / \text{sort} / (\mathbb{R}^Q, <_{\text{lex}}).$$

We denote by $Y_{\text{lex-MO}} = f(X_{\text{lex-MO}})$ its image in objective space.

Theorem 6.11.

a) $|\,\text{sort}(Y_{\text{lex-MO}})| = |\{\text{sort}(f(x)) : x \in X_{\text{lex-MO}}\}| = 1.$

b) $X_{\text{lex-MO}} \subset X_{\text{Par}} \cap X_{\text{MO}}$ *and* $X_{\text{lex-MO}} = X_{\text{Par}} \cap X_{\text{MO}}$ *if* $|X_{\text{MO}}| \le 1.$

Proof:

a) Follows because $<_{\text{lex}}$ is a total order.

b) Let $x \in X_{\text{lex-MO}}$.
First, assume $x \notin X_{\text{Par}} \Longrightarrow \exists\, x' \in X$ s.t. $f(x') < f(x)$
$\Longrightarrow \text{sort}(f(x')) \le_{\text{lex}} \text{sort}(f(x))$ and $\text{sort}(f(x)) \ne \text{sort}(f(x))$ ↯ Contradiction
Second, assume $x \notin X_{\text{MO}} \Longrightarrow \exists\, x' \in X$ s.t. $\max_{i=1,\dots,Q} f_i(x') < \max_{i=1,\dots,Q} f_i(x)$
$\Longrightarrow \text{sort}_1(f(x')) < \text{sort}_1(f(x))$
$\Longrightarrow \text{sort}(f(x')) <_{\text{lex}} \text{sort}(f(x))$
The rest follows from Proposition 6.7.

□

Therefore, there is a unique sorted objective value vector, and a lex-MO solution is both Pareto and max-ordering optimal.

Example 6.2. The inclusion $X_{\text{lex-MO}} \subset X_{\text{MO}} \cap X_{\text{Par}}$ may be strict.

Let $X = \{a, b, c, d, e\}$ with the values

x	$f(x)$	sort $f(x)$
a	$(1, 3, 8, 2, 4)$	$(8, 4, 3, 2, 1)$
b	$(4, 3, 8, 1, 1)$	$(8, 4, 3, 1, 1)$
c	$(7, 5, 4, 6, 1)$	$(7, 6, 5, 4, 1)$
d	$(3, 7, 4, 6, 5)$	$(7, 6, 5, 4, 3)$
e	$(4, 7, 5, 6, 5)$	$(7, 6, 5, 5, 4)$

Then
$$\begin{aligned} X_{\text{MO}} &= \{c, d, e\} \\ X_{\text{Par}} &= \{a, b, c, d\} \\ X_{\text{lex-MO}} &= \{c\} \end{aligned}$$

Since sort is a permutation of the objective functions (depending on x) we see that for each $x^* \in X_{\text{lex-MO}}$ $\exists\, \pi \in \Pi(Q)$ s.t. x^* solves $\operatorname{lexmin}_{x \in X}(f_{\pi(1)}(x), \dots, f_{\pi(Q)}(x))$. Using Proposition 6.3 we obtain:

Corollary 6.12. *All $x^* \in X_{\text{lex-MO}}$ s.t. $\{x : f(x) = f(x^*)\}$ is a singleton are strictly Pareto optimal.*

Next, we show that $X_{\text{lex-MO}}$ is invariant under permutations and strictly monotone increasing mappings.

Proposition 6.13.

a) $X_{\text{lex-MO}}$ is the same for $(f_1, \dots, f_Q)$ and $(f_{\pi(1)}, \dots, f_{\pi(Q)})$ for all permuations $\pi \in \Pi(Q)$.

b) Let $\tau : \mathbb{R} \to \mathbb{R}$ be strictly increasing. Then $X_{\text{lex-MO}}$ is the same for $(f_1, \dots, f_Q)$ and $(\tau \circ f_1, \dots, \tau \circ f_Q)$.

Proof:

a) obvious: $\operatorname{sort}(f_1(x), \dots, f_Q(x)) = \operatorname{sort}(f_{\pi(1)}(x), \dots, f_{\pi(Q)}(x))$.

b) By the strict monotonicity

$$f_i(x) < f_i(x') \iff \tau(f_i(x)) < \tau(f_i(x'))$$

Therefore $\operatorname{sort}(f(x)) <_{\text{lex}} \operatorname{sort}(f(x')) \iff \operatorname{sort}(\tau(f(x))) <_{\text{lex}} \operatorname{sort}(\tau(f(x')))$

□

Beside the fact that $X_{\text{lex-MO}} \subset X_{\text{Par}}$, we can strengthen the result of Proposition 6.6 for $X_{\text{lex-MO}}$.

Suppose that $\inf_{x \in X} f_i(x) > -\infty \quad \forall\, i = 1, \dots, Q$.

Theorem 6.14. $x \in X_{\text{Par}} \iff \exists\, \lambda \in \operatorname{int} \mathbb{R}^Q_+$ *s.t.* $x \in X_{\text{lex-MO}}$ *for* $(\lambda_1(f_1 - y_1^{00}), \dots, \lambda_Q(f_Q - y_Q^{00}))$.

Proof:

„ $\Longleftarrow$ " Let $x^* \in X_{\text{lex-MO}}$ for the given functions and assume $x^* \notin X_{\text{Par}} \quad \Longrightarrow \quad \exists\, x \in X$ s.t. $f(x) < f(x^*)$

$\Longrightarrow \ \lambda_i(f_i(x) - y_i^{00}) \leq \lambda_i(f_i(x^*) - y_i^{00}) \quad \forall\, i = 1, \dots, Q$

and strict inequality for some k.

$\Longrightarrow \ \text{sort}(\lambda_i(f_i(x) - y_i^{00})) <_{\text{lex}} \text{sort}(\lambda_i(f_i(x^*) - y_i^{00}))$ ↯ Contradiction

„ $\Longrightarrow$ " Let $x^* \in X_{\text{Par}}$. Define $\lambda_i := \frac{1}{f_i(x^*) - y_i^{00}}$.

Then $\lambda_i(f_i(x^*) - y_i^{00}) = 1 \quad \forall\, i = 1, \dots, Q$.

$x^* \in X_{\text{Par}} \implies \forall\, x \in X \quad f(x) \neq f(x^*) \quad \exists\, k \in \{1, \dots, Q\}$ s.t. $f_k(x) > f_k(x^*)$

$\Longrightarrow \ \lambda_k(f_k(x) - y_k^{00}) > 1$

$\Longrightarrow \ \text{sort}(\lambda_i(f_i(x) - y_i^{00})) >_{\text{lex}} (1, \dots, 1) = \text{sort}(\lambda_i(f_i(x^*) - y_i^{00}))$ ↯ Contradiction

□

Let us discuss the solution of lex-MO problems now. Could we apply a procedure like the lexicographic method ?

First we would have to solve the max-ordering problem. Then fix the value of the worst objective, solve the max-ordering problem for the remaining $Q - 1$ objectives. Unfortunately, we do not know which objective will be the worst, and there may be several x with the worst value obtained for different objectives, but both MO solutions. See e.g. c, $f_1(c) = 7$, and d, $f_2(d) = 7$ in Example 6.2.

Under additional assumptions on f_i, we can show that there is one objective f_q s.t.

$$f_q(x) = \min_{x \in X} \max_{i=1,\dots,Q} f_i(x) \quad \forall\, x \in X_{\text{MO}} \tag{6.8}$$

The following is from Behringer, 1977, [Beh77].

Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be convex. We use X_{MO} to denote the set of all optimal solutions of the max-ordering problem and $X_{\text{lex-MO}}$ for the optimal solutions of the lex-MO problem.

Furthermore:

$$\begin{aligned} z_{\text{MO}} &:= \min_{x \in X} \max_{i=1,\dots,Q} f_i(x) && (6.9)\\ A_i &:= \{x \in X : f_i(x) = \max_{j=1,\dots,Q} f_j(x)\} && (6.10)\\ L_i &:= \{x \in A_i : f_i(x) = \min_{x \in A_i} f_i(x)\} && (6.11) \end{aligned}$$

Note that $\max_{i=1,\dots,Q} f_i(x)$ is a convex function. If X is compact, f_i are continuous on X and hence $X_{\text{MO}} \neq \emptyset$ and compact. Iteratively we get that $X_{\text{lex-MO}} \neq \emptyset$ and compact.

(For this and all following results it is enough that f_i are lower semicontinuous and strictly quasiconvex.)

Lemma 6.15. *If f_i are convex, X is convex then X_{MO} is convex.*

Proof: Assume $X_{\text{MO}} \neq \emptyset$. Because f_i are convex $\max\limits_{i=1,\dots,Q} f_i(x)$ is convex.

$$\begin{aligned} X_{\text{MO}} &= \{x \in X : \text{sort}_1(f(x)) = z_{\text{MO}}\} = \{x \in X : \text{sort}_1(f(x)) \leq z_{\text{MO}}\} \\ &= \bigcap_{i=1}^{Q} \{x \in X : f_i(x) \leq z_{\text{MO}}\} = \bigcap_{i=1}^{Q} L^i_{\leq}(z_{\text{MO}}) \end{aligned}$$

is convex as intersection of convex sets.

□

Theorem 6.16. *Assume X is convex, f_i are convex. Then $\exists\, k \in \{1,\dots,Q\}$ s.t. $f_k(x) = z_{\text{MO}} \quad \forall\, x \in X_{\text{MO}}$.*

Proof: Let $\hat{x} \in X_{\text{MO}} \implies \exists\, j$ s.t. $f_j(\hat{x}) = z_{\text{MO}} \implies f_j(\hat{x}) \geq f_i(\hat{x}) \quad \forall\, i = 1,\dots,Q$.

Suppose $\nexists\, k \in \{1,\dots,Q\}$ s.t. $f_k(x) = f_j(x^0) \quad \forall\, x \in X_{\text{MO}}$.

$\implies \forall\, k \in \{1,\dots,Q\} \quad \exists\, x^k \in X_{\text{MO}}$ s.t. $f_k(x^k) < f_j(x^0)$ and $f_i(x^k) \leq f_j(x^0) \quad \forall\, i = 1,\dots,Q$. (Note that $x^k \in X_{\text{MO}}$ does not allow $f_i(x^k) > f_j(x^0)$.)

Let $x^* := \sum\limits_{k=1}^{Q} \alpha_k x^k$ with $\alpha_k > 0$, $\sum \alpha_k = 1$. Then $x^* \in X_{\text{MO}}$, because of convexity (Lemma 6.15).

$\implies f_i(x^*) \leq \sum\limits_{k=1}^{Q} \alpha_k \underbrace{f_i(x^k)}_{\text{strict inequality for } i = k} < f_j(x^0)$ contradicting $x^0 \in X_{\text{MO}}$.

□

Theorem 6.16 says that z_{MO} is attained for all $x \in X_{\text{MO}}$ for at least one objective. The index k in 6.16 is called a **common index**.

Theorem 6.17. *Suppose X is convex, f_i are convex, $X_{\text{MO}} \neq \emptyset$.*
Then k is a common index $\iff X_{\text{MO}} = L_k$.

Proof:

„$\implies$“ Let k be a common index.

First suppose $x \in X_{\text{MO}} \implies f_k(x) = z_{\text{MO}} \implies x \in L_k \implies X_{\text{MO}} \subset L_k$

Now suppose $x \in L_k \implies f_k(y) \geq f_k(x) \quad \forall\, y \in A_k$ (6.12)

Assume $x \notin X_{\text{MO}}$.

$\implies \max\limits_{i=1,\dots,Q} f_i(x) > z_{\text{MO}}$

Because $X_{\text{MO}} \neq \emptyset \implies \exists\, \hat{x} \in X_{\text{MO}}$ and because k is a common index.

$\implies f_k(\hat{x}) = \max\limits_{i=1,\dots,Q} f_i(\hat{x}) = z_{\text{MO}}$ and $\hat{x} \in A_k$

$\implies \max\limits_{i=1,\dots,Q} f_i(x) > z_{\text{MO}} = f_k(\hat{x})$

With (6.12) $\implies f_k(\hat{x}) \geq f_k(x) = \max\limits_{i=1,\dots,Q} f_i(x) > f_k(\hat{x})$ ↯ Contradiction

$\implies x \in X_{\text{MO}} \implies L_k \subset X_{\text{MO}}$

„$\impliedby$“ Let $x \in L_k = X_{\text{MO}}$.

$\implies f_k(x) = \max\limits_{i=1,\dots,Q} f_i(x)$ by definition of L_k and $\max\limits_{i=1,\dots,Q} f_i(x) = \min\limits_{x \in X} \max\limits_{i=1,\dots,Q} f_i(x)$ by definition of $X_{\text{MO}} \implies k$ is a common index.

□

The following theorem gives criteria for k to be a common index.

Theorem 6.18. *Suppose X is convex, f_i are convex and $X_{\text{MO}} \neq \emptyset$. Then*

a) $L_i = \emptyset \implies i$ *is not a common index*

b) *Let* $J := \{i \in \{1, \dots, Q\} \mid L_i \neq \emptyset\}$ *and* $m_i := \min_{x \in A_i} f_i(x)$.
Define $\overline{m} := \min_{i \in J} m_i$. *Then if* $m_i > \overline{m}$ $\quad i$ *is not a common index.*

c) *Let* $\overline{J} := \{i \in J : m_i = \overline{m}\}$. *Then* $L_k = \bigcup_{j \in \overline{J}} L_j \iff k \in \overline{J}$ *is a common index.*

Proof:

a) i is a common index $\iff \emptyset \neq X_{\text{MO}} = L_i = \emptyset$ ↯ Contradiction

b) Suppose $m_i > m_j$ and that i is a common index. Then $L_i = X_{\text{MO}} \neq \emptyset$.
Let $x^0 \in X_{\text{MO}}$ and $\hat{x} \in L_j \neq \emptyset$.
$\implies z_{\text{MO}} = \max_{l=1,\dots,Q} f_l(x^0) = f_i(x^0) = m_i > m_j = f_j(\hat{x}) = \max_{i=1,\dots,Q} f_j(\hat{x})$ ↯ Contradiction

c) „ $\Longleftarrow$ “

Let $k \in \overline{J}$ be a common index.
First, $L_k \subseteq \bigcup_{j \in \overline{J}} L_j$ is clear.
Let $x \in \bigcup_{i \in \overline{J}} L_i \implies x \in L_j$ for some $j \in \overline{J}$.
$\implies f_j(x) = \max_{i=1,\dots,Q} f_i(x) = \min_{y \in A_j} f_j(y) = m_j = \overline{m}$.
By Theorem 6.17 $L_k = X_{\text{MO}}$
$\implies f_k(\hat{x}) = m_k = \overline{m} = \max_{i=1,\dots,Q} f_i(\hat{x}) = z_{\text{MO}} \quad \forall\, \hat{x} \in L_k$
$\implies f_j(x) = z_{\text{MO}} \implies x \in X_{\text{MO}} \implies x \in L_k$

„ $\Longrightarrow$ “

$L_k = \bigcup_{i \in \overline{J}} L_i$ for some $k \in \overline{J}$.
Since $X_{\text{MO}} \neq \emptyset \quad \exists$ common index $\hat{k}$ and $X_{\text{MO}} = L_{\hat{k}}$
From a) and b) $\implies \hat{k} \in \overline{J}$, then from „ $\Longleftarrow$ “ $L_{\hat{k}} = \bigcup_{i \in \overline{J}} L_i$
$\implies X_{\text{MO}} = L_{\hat{k}} = \bigcup_{i \in \overline{J}} L_i = L_k$
By Theorem 6.17 k is a common index.

□

Therefore, we can find a common index with the following procedure:

(1) Find $J := \{i \in \{1, \dots, Q\} \ : \ L_i \neq \emptyset\}$

(2) Find one $x^i \in L_i \quad \forall\, i \in J \qquad \overline{J} := \{i \in J : f_i(x^i) \leq f_j(x^j) \quad \forall\, j \in J\}$

(3) $J^* := \{i \in \overline{J} \ : \ L_j \subset L_i \quad \forall\, j \in \overline{J}\}$ is the set of all common indices.

An algorithm for solving lex-MO problems is as follows (for X convex, compact, nonempty and f_i convex):

① $\overline{X} := X \qquad \overline{Q} := \{1, \dots, Q\}$

② Solve the Max-ordering problem for $\overline{f} = f$ and find $\overline{X}_{\text{MO}} \neq \emptyset$.

③ If $|\overline{Q}| = 1 \quad X_{\text{lex-MO}} = \overline{X}_{\text{MO}}, \quad$ STOP

Otherwise determine a common index k, let $\overline{X} = \overline{X}_{\text{MO}}$, $\overline{Q} = \overline{Q} \setminus \{k\}$, $\overline{f} = \overline{f} \setminus f_k$, go to ②

To conclude this chapter, we study some characteristic properties of lex-MO solutions.
Recall that a multicriteria optimization class is the set of all MCOP with the same model map and ordered set. We dicuss properties of the MCO class $\cdot / \text{sort} / (\mathbb{R}^Q, <_{\text{lex}})$, lexicographic max-ordering.
This part is from Ehrgott, 1997, [Ehr97].

Definition 6.2.

a) An MCO class $\cdot/\theta/(\mathbb{R}^p, \preceq)$ satisfies the normalization property, if when $Q = 1$ (i.e. $f : \mathbb{R}^n \to \mathbb{R}$), it coincides with single objective optimization.

$$\cdot/\theta/(\mathbb{R}^p, \preceq) = \cdot/\text{id}/(\mathbb{R}, <) \tag{6.13}$$

which means for any X, for any $f : \mathbb{R} \to \mathbb{R}$ the set of optimal solution is equal to $\{x \in X : f(x) \leq f(y) \quad \forall\, y \in X\}$.

b) An MCO class $\cdot/\theta/(\mathbb{R}^p, \preceq)$ satisfies the regularity property, if for any choice of X, f, and Q, the set of optimal solutions is contained in X_{MO}.

Let us denote the **set of optimal solutions of an MCOP** $(X, f, \mathbb{R}^Q)/\theta/(\mathbb{R}^p, \preceq)$ by

$$\text{Opt}((X, f, \mathbb{R}^Q)/\theta/(\mathbb{R}^p, \preceq)).$$

Now let $(X, f, \mathbb{R}^Q)$ be data for an MCOP, and $y \in \mathbb{R}^Q$ s.t. $\exists\, x \in \text{Opt}((X, f, \mathbb{R}^Q)/\theta/(\mathbb{R}^p, \preceq))$ with $f(x) = y$. Let $K = \{i_1, \dots, i_k\} \subset \{1, \dots, Q\}$. The **reduced problem** RP(K) is $(X^K, f^K, \mathbb{R}^K)/\theta/(\mathbb{R}^p, \preceq)$ where $X^K = \{x \in X : f_i(x) = y_i \quad \forall\, i \in \{1, \dots, Q\} \setminus K\}$ and $f^K = (f_{i_1}, \dots, f_{i_k})$.

Definition 6.3. An MCO class satisfies the reduction property, if for all data $(X, f, \mathbb{R}^Q)$ and $K \subset \{1, \dots, Q\}$ and y as above

$$\begin{aligned} &\text{Opt}((X^K, f^K, \mathbb{R}^K)/\theta/(\mathbb{R}^p, \preceq)) = \\ &\qquad \{x \in \text{Opt}((X, f, \mathbb{R}^Q)/\theta/(\mathbb{R}^p, \preceq)),\ f_i(x) = y_i \quad \forall\, i \notin K\}. \end{aligned} \tag{6.14}$$

Proposition 6.19. *The* lex-MO *class satisfies normalization, regularity and reduction property.*

Proof:

1) Normalization is obvious because $\operatorname{sort}(f(x)) = f(x)$ and $<_{\text{lex}} = <$ if $Q = 1$.

2) From Theorem 6.11 b) regularity is clear ($X_{\text{lex-MO}} \subset X_{\text{MO}}$).

3) We write Opt and Opt(RP(K)) for short for the optimal solutions of the original and reduced problem, respectively.

   Let $\text{Opt}^* := \{x \in \text{Opt} : f_i(x) = y_i \quad \forall\, i \notin K\}$. We have to show $\text{Opt}(\text{RP}(K)) = \text{Opt}^*$.

   We note that $\forall\, x^* \in \text{Opt}^*, \quad \forall\, x \in \text{Opt}(\text{RP}(K))$

   $$f_i(x^*) = f_i(x) = y_i \quad \forall\, i \in \{1, \dots, Q\} \setminus K \tag{6.15}$$

   $$\text{and} \;\; \operatorname{sort}(f(x^*)) \leq_{\text{lex}} \operatorname{sort}(f(x)) \tag{6.16}$$

   First let $\hat{x} \in \text{Opt}(\text{RP}(K))$.

   (6.15) and (6.16) $\implies \quad \operatorname{sort}(f^K(x^*)) \leq_{\text{lex}} \operatorname{sort}(f^K(\hat{x}))$

   By the choice of $x^*, \hat{x}$ this holds with equality $\implies \operatorname{sort}(f(\hat{x})) = \operatorname{sort}(f(x^*)) \implies \hat{x} \in \text{Opt}^*$.

   Second let $\hat{x} \in \text{Opt}^*$.

   Then $\hat{x}$ is feasible for RP(K).

   Assume $\exists\, x \in \text{Opt}(\text{RP}(K))$ s.t. $\operatorname{sort}(f^K(x)) <_{\text{lex}} \operatorname{sort}(f^K(\hat{x}))$

   with (6.15) $\implies \quad \operatorname{sort}(f(x)) <_{\text{lex}} \operatorname{sort}(f(\hat{x}))$ $\lightning$ Contradiction

   $\implies \operatorname{sort}(f(\hat{x})) \leq_{\text{lex}} \operatorname{sort}(f(x))$

   with (6.16) $\implies \quad \operatorname{sort} f(\hat{x}) = \operatorname{sort}(f(x)) \implies \hat{x} \in \text{Opt}(\text{RP}(K))$

□

Theorem 6.20. *An MCO class satisfies normalization, reduction and regularity property* $\iff \quad \cdot/\theta/(\mathbb{R}^p, \preceq) = \cdot/\operatorname{sort}/(\mathbb{R}^Q, <_{\text{lex}})$.

Proof: We only need to show „ $\implies$ ".

We have to show $\text{Opt}((X, f, \mathbb{R}^Q)/\theta/(\mathbb{R}^p, \preceq)) = \text{Opt}((X, f, \mathbb{R}^Q)/\operatorname{sort}/(\mathbb{R}^Q, <_{\text{lex}})$ for any choice of X and f.

We proceed by induction on Q.

$Q = 1$ is obvious from normalization.

Suppose the result is true for not more than $Q - 1$ criteria.

Let $\hat{x} \in \text{Opt}((X, f, \mathbb{R}^Q)/\theta/(\mathbb{R}^p, \preceq))$ and let

$$y := \min_{x \in X} \max_{i=1,\dots,Q} f_i(x) = z_{\text{MO}}.$$

By the regularity property $\exists\, k \in \{1, \dots, Q\}$ s.t. $f_k(\hat{x}) = y$ and $f_i(\hat{x}) \leq y \quad \forall\, i = 1, \dots, Q$

$\implies \hat{x} \in \{x \in \text{Opt}((X, f, \mathbb{R}^Q)/\theta/(\mathbb{R}^p, \preceq)) : f_k(x) = y\}$.

Let $K = \{1, \dots, Q\} \setminus \{k\}$.

$$\begin{aligned}
\implies \{x \in \text{Opt}((X, f, \mathbb{R}^Q)/\theta/(\mathbb{R}^p, \preceq)) : f_k(x) = y\}& \\
&\overset{\text{red.}}{=} \text{Opt}((X^K, f^K, \mathbb{R}^{Q-1})/\theta/(\mathbb{R}^p, \preceq)) \\
&\overset{\text{ind.}}{=} \text{Opt}((X^K, f^K, \mathbb{R}^{Q-1})/\operatorname{sort}/(\mathbb{R}^{Q-1}, <_{\text{lex}})) \\
&\overset{\text{red.}}{=} \{x \in \text{Opt}((X, f, \mathbb{R}^Q)/\operatorname{sort}/(\mathbb{R}^Q, <_{\text{lex}})) : f_k(x) = y\} \\
&\subseteq \text{Opt}((X, f, \mathbb{R}^Q)/\operatorname{sort}/(\mathbb{R}^Q, <_{\text{lex}}).
\end{aligned}$$

The reverse inclusion is proved in the same way.

□

6.4 Exercises to Chapter 6

44. Solve the lexicographic problem

$$\begin{array}{lr} \min & -x_1 + x_2 - x_3 \\ \min & x_2 \\ \min & -x_1 - 2x_2 \\ \text{s.t.} & x_1 + x_2 \leq 1 \\ & x_1 - x_2 + x_3 \leq 4 \end{array}$$

What happens if you reverse the order of objective functions ?

45. Prove, or give counterexamples, to the following conjectures.

If x^* is a solution of $\min_{x \in X} \max_{i=1,\dots,Q} f_i(x)$, then $x^* \in X_{\text{Par}}$ ($X_{\text{w-Par}}$).

46. Solve the problem

$$\begin{array}{lr} \min \max\{-x_1 - 2x_2,\ -x_1 + 2x_3,\ x_1 - x_3\} & \\ \text{s.t.} \qquad x_1 + x_2 \leq 1 & \\ x_1 - x_2 + x_3 \leq 4 & \end{array}$$

Is the optimal solution Pareto optimal ?

(Hint: Write the problem as an LP.)

Bibliography

[ABB53] K.J. Arrow, C.W. Barankin, and D. Blackwell. Admissible points of convex sets. In H.W. Kuhn and A.W. Tucker, editors, *Contributions to the Theory of Games*, volume 2. Princeton University Press, Princeton New Jersey, 1953.

[Beh77] F.A. Behringer. Lexicographic quasiconcave multiobjective programming. *Zeitschrift für Operations Research*, 21:103–116, 1977.

[Ben78] H.P. Benson. Existene of efficient solutions for vector maximization problems. *Journal of Optimization Theory and Applications*, 26(4):569–580, 1978.

[Ben79] H.P. Benson. An improved definition of proper efficiency for vector maximization with respect to cones. *Journal of Optimization Theory and Applications*, 71:232–241, 1979.

[Ben97] H.P. Benson. An outer approximation algorithm for generating all efficient extreme points in the outcome set of a multiple objective linear programming problem. Technical report, University of Florida, 1997. To appear in Journal of Global Optimization.

[Bor77] J. Borwein. Proper efficient points for maximization with respect to cones. *SIAM Journal on Control and Optimization*, 15(1):57–63, 1977.

[BSS93] M.S. Bazaara, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming – Theory and Algorithms.* John Wiley & Sons, New York, 1993.

[Cor80] H.W. Corley. An existence result for maximization with respect to cones. *Journal of Optimization Theory and Applications*, 31(2):277–281, 1980.

[EHK+97] M. Ehrgott, H.W. Hamacher, K. Klamroth, S. Nickel, A. Schöbel, and M.M. Wiecek. A note on the equivalence of balance points and Pareto solutions in multiple-objective programming. *Journal of Optimization Theory and Applications*, 92(1):209–212, 1997.

[Ehr97] M. Ehrgott. *Multiple Criteria Optimization – Classification and Methodology.* Shaker Verlag, Aachen, 1997.

[Geo68] A.M. Geoffrion. Proper efficiency and the theory of vector maximization. *Journal of Mathematical Analysis and Applications*, 22:618–630, 1968.

[GN90] A. Göpfert and R. Nehse. *Vector Optimization (in German)*. B.G. Teubner, Leipzig, 1990.

[Hel23] E. Helly. On sets of convex bodies with common points (in German). *Jahrbuch der Deutschen Mathematiker Vereinigung*, 32:175–176, 1923.

[HLW71] Y.Y. Haimes, L.S. Lasdon, and D.A. Wismer. On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Transactions on Systems, Man, and Cybernetics*, 1:296–297, 1971.

[Ise74] H. Isermann. Proper efficiency and the linear vector maximum problem. *Operations Research*, 22:189–191, 1974.

[Ise77] H. Isermann. The enumeration of the set of all efficient solutions for a linear multiple objective program. *Operational Research Quarterly*, 28(3):711–725, 1977.

[Küf98] K.-H. Küfer. On the asymptotic number of efficient vertices in multiple objective linear programming. *Journal of Complexity*, 14:333–377, 1998.

[KSS97] P. Korhonen, S. Salo, and R.E. Steuer. A heuristic for estimating Nadir criterion values in multiple objective linear programming. *Operations Research*, 45(5):751–757, 1997.

[KT51] H.W. Kuhn and A.W. Tucker. Nonlinear programming. In J. Neyman, editor, *Proceedings of the second Berkeley Symposium on on Mathematical Statistics and Probability*, pages 481–492. University of California Press, Berkeley, 1951.

[Man69] O.L. Mangasarian. *Nonlinear Programming*. McGraw Hill, New York, 1969.

[MB94] C. Molivert and N. Boissard. Structure of efficient sets for strictly quasi convex objectives. *Journal of Convex Analysis*, 1(2):143–150, 1994.

[Nac78] P.H. Naccache. Connectedness of the set of nondominated outcomes in multicriteria optimization. *Journal of Optimization Theory and Applications*, 25:459–467, 1978.

[NW88] Nemhauser and Wolsey. *Integer and Combinatorial Optimization*. John Wiley & Sons, New York, 1988.

[SNT85] Y. Sawaragi, H. Nakayama, and T. Tanino. *Theory of Multiobjective Optimization*. Academic Press, Orlando, 1985.

[Ste85] R.E. Steuer. *Multiple Criteria Optimization: Theory, Computation and Application*. John Wiley & Sons, New York, 1985.

[SW70] J. Stoer and C. Witzgall. *Convexity and Optimization in Finite Dimensional Spaces I.* 1970.

[War83] A.R. Warburton. Quasiconcave vector maximization: Connectedness of the sets of Pareto-optimal and weak Pareto-optimal alternatives. *Journal of Optimization Theory and Applications*, 40:537–557, 1983.

List of Figures

1.1 Objective Functions of Example 1.3 2
1.2 Criterion Space in Example 1.1 3
1.3 Criterion Space in Example 1.3 3
1.4 Efficient Points in Example 1.3 4
1.5 Min Max Solution for Example 1.3 5
1.6 Illustration of Two Cones 7

2.1 Illustration of Definition 2.1 14
2.2 Feasible Set of Example 2.1 14
2.3 Feasible Set of Example 2.1 15
2.4 „lower left sector“ of Y 16
2.5 $(y^0 - \mathbb{R}^Q_+) \cap Y,\ y^0 - \mathbb{R}^Q_+$ 17
2.6 $(y^\alpha - \mathbb{R}^Q_+)^c$ 18
2.7 Empty Efficient Set 21
2.8 Efficient and Weakly Efficient Points 21
2.9 Level Set and Level Curve in Example 2.3 22
2.10 Level Sets 22
2.11 Location Problem of Example 2.4 25
2.12 Properly Efficient Point 27
2.13 Properly Pareto Optimal $y \in Y_{\text{eff}}$ 29
2.14 Conical Hull and Tangent Cone 30
2.15 Benson's Proper Efficiency 32

3.1 $\text{Opt}(\lambda, Y)$ 39
3.2 Properly Separated S_1 and S_2 40
3.3 Y_{eff} and $Y_{\text{p-eff}}$ 43
3.4 Illustration of the Proof of Theorem 3.9 44
3.5 Illustration of the Proof of Theorem 3.9 45
3.6 The Set Y in Example 3.2 45
3.7 Objective Functions of Example 3.3 48
3.8 Connectedness of Y_{eff} 49
3.9 Illustration of the Proof of Theorem 3.21 50

3.10 Illustration of the Proof of Theorem 3.21 51

4.1 Y, Y_{eff}, y^0 and y^N 55
4.2 Feasible Set of Example 4.1 58
4.3 Solution of $P_k(\varepsilon)$ 59
4.4 Illustration of Benson's Problem 61
4.5 Objective Functions of Example 4.2 63
4.6 Sets $\{y : \|y - y^0\| \leq c\}$ for Different Norms l_1, l_2 and γ 64
4.7 $\{y : \|y - \hat{y}\|_\infty < \varepsilon\}$ 68
4.8 $\operatorname{cl} A(Y) \not\subset Y_{\text{eff}}$ 69

5.1 Feasible Set and Objectives in Example 5.1 77
5.2 Criterion Space in Example 5.1 78
5.3 A Polyhedron with Extreme Points and Extreme Rays 82
5.4 X_{Par} for Example 5.2 83
5.5 Efficient Bases and Corresponding Extreme Points 91
5.6 Feasible and Pareto Set in Example 5.3 91
5.7 Weights to obtain Pareto Faces in Example 5.3 92

6.1 Illustration of Example 6.1 98

Index

$<_{\text{lex}}$, 10
$P_\varepsilon(x^0)$, 60
$P_k(\varepsilon)$, 58
$T_Y(y)$, 30
X_{Par}, 13
$X_{\text{lex-MO}}$, 101
$Y_{\text{p-eff}}$, 42
Y_{eff}, 13
$Y_{\text{lex-MO}}$, 101
cone(Y), 30
lex-MO solution, 101
$\mathbb{R}^Q_+$-compact, 18
$\mathbb{R}^Q_+$-semicompact, 18
$\mathbb{R}^Q_+$-semicontinuous, 19
$\preceq_K$, 8
sort(y), 101
$\varepsilon-$constraint method, 58
l_p-norm, 65

adjacent, 79
antisymmetric, 5
asymmetric, 5

bases
  connected efficient bases, 81
basic feasible solution, 75
basic solution, 75
basis, 75
Benson, 31
Benson's Method, 60
bfs, 75
binary relation, 5
  antisymmetric, 5
  asymmetric, 5
  connected, 5
  irreflexive, 5
  negatively transitive, 5
  reflexive, 5
  strongly connected, 5
  symmetric, 5
  transitive, 5
Borwein, 31

chain, 17
common index, 104
componentwise order, 7
cone, 7
  tangent cone, 30
conical hull, 30
connected, 5, 49
connected efficient bases, 81
connectedness of X_{Par}, 49
connectedness of Y_{eff}, 49
continuous optimization problem, 1
convex cone, 8
Corley's method, 70

decision space, 2
dimension, 75
discrete optimization problem, 1
dominate, 13

edge, 75
efficient, 13
  properly efficient, 27
efficient basis, 79
efficient nonbasic variable, 79
efficient pivot w.r.t. B and x_j, 79
efficient point, 4

elements of a MCOP, 10
equivalence relation, 6
extreme point, 75
extreme ray, 75

face, 75
facet, 75

Geoffrion, 27

Helly, 25
hyperplane, 75

ideal point, 55
inductively ordered set, 16
irreflexive, 5
Isermann, 74

Kuhn
  Kuhn-Tucker, 34

level curve, 21
level curves, 21
level set, 21
lexicographic Max-ordering solution, 101
lower semicontinuous, 36

max-ordering optimization, 99
maximal face, 75
maximal Pareto face, 81
maximal set of efficient nonbasic variables, 93
MCO class, 11
MCOP, 10
Minimax Theorem of Sion-Kakutani, 44
model map, 9
monotone norm, 64
multicriteria optimization class, 11
Multicriteria Simplex Algorithm, 85

Naccache, 50
Nadir point, 56
negatively transitive, 5
nondominated, 13
nontrivial cone, 8
norm, 64
not connected, 49

objective (criterion) space, 2
optimal solution, 10
optimal value, 10
Optimization with Multiple Criteria, 1

parametric linear programming, 76
Pareto face, 81
  maximal Pareto face, 81
Pareto optimal, 13, 22
  properly Pareto optimal, 27, 31
Pareto optimal point, 13
Pareto optimal solutions, 2
partial order, 7
  strict partial order, 7
pointed cone, 8
polyhedron, 75
polytope, 75
preorder, 6
preordered set, 6
properly efficient, 27
properly Pareto optimal, 27, 31, 34

quasiorder, 6

ray, 75
reduced costs, 75
reduced problem, 106
reflexive, 5
RP(K), 106

set of all efficient points, 13
set of all Pareto optimal, 13
set of optimal solutions of an MCOP, 106
Simplex Algorithm
  Multicriteria Simplex Algorithm, 85
Sion-Kakutani, 44
strict componentwise order, 7

strict level sets, 22
strict partial order, 7
strictly efficient, 20
strictly monotone norm, 64
strictly Pareto optimal, 20, 22
strongly connected, 5
supporting hyperplane, 75
symmetric, 5

tangent cone, 30
transitive, 5
Tucker
    Kuhn-Tucker, 34
Tucker's Theorem of the alternative, 34

upper bound, 56

weak componentwise order, 7
weakly efficient, 20
weakly Pareto optimal, 22
weakly Pareto optimal point, 20

Zorn's Lemma, 17