Åmund Engmark

Multicriteria optimization problems

Bachelor's project in Mathematical Sciences
Supervisor: Elisabeth Anna Sophia Köbis
May 2020

Bachelor's project

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

NTNU
Kunnskap for en bedre verden

Åmund Engmark

Multicriteria optimization problems

Bachelor's project in Mathematical Sciences
Supervisor: Elisabeth Anna Sophia Köbis
May 2020

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

NTNU
Kunnskap for en bedre verden

Multicriteria optimization problems

Åmund Engmark

May 2021

Abstract

In this bachelor thesis, we will consider various themes in multi-objective optimization. In particular, we will look at some optimality notions, existence results, scalarization results, and some problems in uncertain multiobjective optimization.

Contents

1 Introduction

In single-objective optimization one wishes to solve the problem

$$\begin{aligned} &\min f(x) \\ &\text{s.t.} \quad x \in \Omega, \end{aligned}$$

where $\Omega \subseteq X$ for a linear space X is the feasible set of the problem and $f : X \to \mathbb{R}$ is some scalar-valued function. In multiobjective optimization problems (MOP) we wish to solve the problem when f is a vector-valued function $f : X \to \mathbb{R}^n$. Since there exists no total order on $\mathbb{R}^n$, there is no obvious meaning of "$\min f(x)$" in the multiobjective case. This complicates the problem.

The absence of a total order on $\mathbb{R}^n$ leads to the concept of *notions of optimality*, ways in which a solution is minimal. Popular notions of optimality include lexicographic optimality, max-ordering optimality, and nondomination/efficiency - optimality. In this thesis we will focus on the nondomination/efficiency notion of optimality only. In short, a vector is said to be non-dominated if it is impossible to improve one of its components without worsening another (see Definition 1.2. An $x \in \mathit{\Omega}$ is said to be efficient if $f(x)$ is non-dominated. In the following, we define this more precisely and then derive many useful results in multiobjective optimization using efficiency. The results in this chapter relay mainly on [1].

1.1 Efficiency and nondominance

In order to define efficiency, and the related concepts weak and strict efficiency, we must first define the ordering relations $\leqq, \leq, <$ on $\mathbb{R}^n$.

Definition 1.1. *Let y, $\hat{y} \in \mathbb{R}^n$ be vectors. Then we write*

- $y \leq \hat{y}$ *if* $y_i \leq \hat{y}_i \quad$ *for* $i = 1, \ldots, n,$ *and* $y \neq \hat{y}$,
- $y < \hat{y}$ *if* $y_i < \hat{y}_i \quad$ *for* $i = 1, \ldots, n,$
- $y \leqq \hat{y}$ *if* $y_i \leq \hat{y}_i \quad$ *for* $i = 1, \ldots, n.$

Definition 1.2. *A feasible solution $\hat{x} \in \Omega$ is said to be*

- *efficient if there exist no other $x \in \Omega$ such that $f(x) \leq f(\hat{x})$,*
- *weakly efficient if there exist no other $x \in \Omega$ such that $f(x) < f(\hat{x})$,*
- *strictly efficient if there exist no other $x \in \Omega \setminus \{\hat{x}\}$ such that $f(x) \leqq f(\hat{x})$.*

If $\hat{x}$ is (weakly) efficient, then $f(\hat{x})$ is said to be (weakly) nondominated.

Efficiency is also called Pareto-optimality.

We see that a function value $f(\hat{x})$ being nondominated means that there exist no other function values $f(x)$ which are better than or equal to that value in every component, and lower than the value in at least one component. Thus $f(\hat{x})$ is "minimal" in some sense. There may still be other function values $f(x)$ so that $f_i(x) < f_i(\hat{x})$ for several $i \in \{1, \ldots, n\}$ but then there will be at least one k such that $f(x)$ fares worse than $f(\hat{x})$ in the k-th component, $f_k(x) > f_k(\hat{x})$. Of course, there will be many nondominated function values for a MOP and none of these "minimas" can be regarded as better than another unless some additional criteria are specified.

$\hat{x}$ is efficient if $f(\hat{x})$ is nondominated, so an efficient solution can be thought of as an analogue to a minimizer in single-objective optimization. Weak/strict efficiency are weaker/stronger conditions than efficiency. In the former case we require only that there are no $x \in \Omega$ which are better than $\hat{x}$ in all components, while in the latter case we require that there are no x's such that $f(x)$ either dominates $f(\hat{x})$ or is equal to $f(\hat{x})$ for $x \neq \hat{x}$

Notation 1. *The set of (weakly) nondominated points in a codomain-set Y is denoted as Y_N (Y_{wN}). The set of (weakly/strictly) efficient solutions in a feasible set X is denoted as X_E (X_{wE}/X_{sE}).*

Note that it doesn't make sense to talk about a "strict nondominance set" Y_{sN} since strict efficiency refers to solutions whose function values are unique.

Figure 1.1[1] illustrates nondominated values of $f : X \to \mathbb{R}^2$ in red. Taking efficient $\hat{x}$ as an example, we see that $f(\hat{x}) = [f_1(\hat{x}), f_2(\hat{x})]$ is nondominated since we cannot lower $f_1(x)$ without increasing $f_2(x)$ and vice versa.

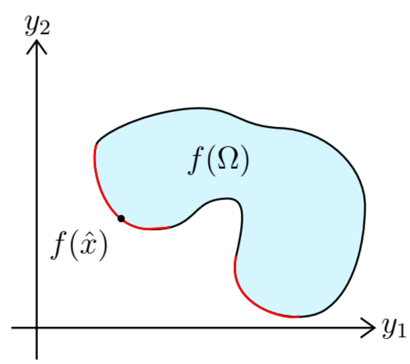

Figure 1: Nondominated points

1.2 Existence results

We give conditions ensuring that given a set Y, we have $Y_N \neq \emptyset$, i.e. ensuring there exists nondominated points. Then it is easy to give conditions ensuring $X_E \neq \emptyset$, i.e. there exists efficient solutions.

Theorem 1.1 ([3, Borwein])**.** *Suppose there exists some $y_0 \in Y$ s.t. $Y_0 = \{y \in Y \mid y \leqq y_0\}$ is compact. Then Y_N is nonempty.*

Borweins theorem says that if Y is compact at the 'bottom', so that intuitively there is a solid lower border, then there will exist nondominated points. Another way of showing Y_N is nonempty is Corleys theorem, which uses the concept of $\mathbb{R}^p_{\geqq}$-semicompactness.

Definition 1.3. *A set $Y \subseteq \mathbb{R}^n$ is called* $\mathbb{R}^p_{\geqq}$-semicompact *if every open cover of Y of the form*

$$C = \{(y^i - \mathbb{R}^p_{\geqq})^c : y_i \in Y, i \in I\}$$

has a finite subcover.

Theorem 1.2 ([4, Corley])**.** *Suppose Y is $\mathbb{R}^p_{\geqq}$-semicompact. Then Y_N is nonempty.*

Corleys theorem only requires semicompactness, but it is often easier to verify that a set $Y \in \mathbb{R}^p$ is $\mathbb{R}^p_{\geqq}$ - compact.

Definition 1.4. *A set Y is $\mathbb{R}^p_{\geqq}$-compact if for all $y \in Y$, $(y - \mathbb{R}^p_{\geqq}) \cap Y$ is compact.*

It can be shown that if Y is $\mathbb{R}^p_{\geqq}$- compact then it is $\mathbb{R}^p_{\geqq}$-semicompact, so Corleys theorem can be used when we know that Y is $\mathbb{R}^p_{\geqq}$- compact. We also have the following intuitive result for positive $\mathbb{R}^p_{\geqq}$-compact sets

Theorem 1.3 ([1, Theorem 2.21])**.** *Let $Y \subseteq \mathbb{R}^p_{\geqq}$ be $\mathbb{R}^p_{\geqq}$ - compact. Then Y_N is externally stable, meaning $Y \subseteq Y_N + \mathbb{R}^p_{\geqq}$.*

Proof. We want to show that for any $y \in Y$ and the set $U = (y - \mathbb{R}^p_{\geqq}) \cap Y$, we have $U \cap Y_N \neq \emptyset$. Then we would know that $y = y_n + a$ for some $y_n \in Y_n, a \in \mathbb{R}^p_{\geqq}$. To show $U \cap Y_N \neq \emptyset$, we show $U_N \neq \emptyset$ and $U_N \subset Y_N$. By Definition 1.4, U is compact, so that by Theorem 1.2, $U_N \neq \emptyset$. Assume for some $y' \in U$, that $y' \notin Y_N$ (we need not consider $y' \notin U$ since then certainly $y' \notin U_N$). Then since y' is not nondominated, there will be some $y'' \in Y$ so that $y'' \leq y'$. Since then $y'' \leq y$ so $y'' \in U$, we must have $y' \notin U_N$. □

Corley's theorem means that if we can show that f(X) is $\mathbb{R}^p_{\geq}$- semicompact, then $X_E \neq \emptyset$, so there exists efficient points. To derive sufficient conditions, we define $\mathbb{R}^p_{\geq}$ – semicontinuity.

Definition 1.5. *$f : \mathbb{R}^n \to \mathbb{R}^p$ is said to be $\mathbb{R}^p_{\geq}$ – semicontinuous if $f^{-1}(y - \mathbb{R}^p_{\geq}) = \{x \mid f(x) \leq y\}$ is closed $\forall y \in \mathbb{R}^p$.*

Henceforth we will not bother to distinguish the feasible set Ω from the domain X, and assume that the domain X is always feasible.

The next theorem gives a sufficient condition for knowing there exist efficient points.

Theorem 1.4 ([1, Theorem 2.19])**.** *Let $X \in \mathbb{R}^n$ be compact, and $f : \mathbb{R}^n \to \mathbb{R}^p$ be $\mathbb{R}^p_{\geq}$-semicontinuous. Then $Y = f(X)$ is $\mathbb{R}^p_{\geq}$-semicompact, so that $X_E \neq \emptyset$.*

Proof. We want to show that an open cover of Y of form $\{(y^i - \mathbb{R}^p_{\geqq})^c : y_i \in Y, i \in I\}$ has a finite subcover. Since $Y = f(X)$, we'll get that

$$\{f^{-1}((y^i - \mathbb{R}^p_{\geqq})^c) : y_i \in Y, i \in I\}$$

is a cover of X, and since f is $\mathbb{R}^p_{\geqq}$-semicontinuous, it is an open cover. Since X is compact, we know this open cover has a finite subcover S. Then taking the image again, $f(S)$ is a finite subcover of Y, proving Y is $\mathbb{R}^p_{\geq}$-semicompact. Now we can use Theorem 1.2 (Corleys theorem) to get that $Y_N \neq \emptyset$ and so $X_E \neq \emptyset$.

□

There are analogues to theorems 1.2 and 1.4 for weak nondominance and efficiency:

Theorem 1.5 ([1, Theorem 2.25])**.** *If $Y \subseteq \mathbb{R}^p$ is compact, then Y_{wN} is nonempty.*

Proof. If Y_{wN} is empty, then for every $y \in Y$ there would be some other $\bar{y} \in Y$ so that $\bar{y} < y$. This means

$$Y \subseteq \bigcup_{y \in Y} (y + \mathbb{R}^p_{>}).$$

Since Y is compact, this open cover will have a finite subcover $S = \bigcup_{y \in F} (y + \mathbb{R}^p)$, where F is a finite subset of Y. But then we could choose the lowest of the y's in F, $y_{\min}$, and this would be in Y yet not in S since $0 \notin \mathbb{R}^p_>$. This contradicts $Y \subseteq S$. $\square$

Lemma 1.6. *If $X \in \mathbb{R}^n$ is compact, and $f : \mathbb{R}^n \to \mathbb{R}^p$ is continuous, then $Y = f(X)$ is compact.*

Proof. Let

$$U = \bigcup_{i \in I} Y_i$$

be an open cover of Y. Then $f^{-1}(U)$ is open since f is continuous, and can be written as an open cover of X,

$$C = \bigcup_{i \in I} X_i.$$

This has a finite subcover, C_F, and then $U_F = f(C_F)$ will give a finite subcover of f of elements in U. $\square$

Corollary 1.6.1 ([1, Corrolary 2.26])**.** *Let $X \subseteq \mathbb{R}^n$ be compact and f be continuous. Then $X_{wE} \neq \emptyset$.*

Proof. Since $Y = f(X)$ is compact by Lemma 1.6, we have $Y_{wN} \neq 0$ by Theorem 1.5. Then $X_{wE} \neq 0$ follows. $\square$

We can characterize X_E, X_{wE}, X_{sE} in terms of so-called *level sets* and *level curves* to provide further intuition on efficient solutions.

Definition 1.6. *For function $f : X \to R$ and $\hat{x} \in X$,*

$$\mathcal{L}_{\leq}(f(\hat{x})) = \{x \in X \mid f(x) \leq f(\hat{x})$$

is called the level set *of f at x,*

$$\mathcal{L}_{=}(f(\hat{x})) = \{x \in X \mid f(x) = f(\hat{x})$$

is called the level curve *of f at $\hat{x}$,*

$$\mathcal{L}_<(f(\hat{x})) = \{x \in X \mid f(x) < f(\hat{x})$$

is called the strict level set *of f at x.*

Theorem 1.7 ([1, Theorem 2.30]). *For function $f : X \to \mathbb{R}^p$ and $\hat{x} \in X$, we have*

1. *$\hat{x}$ is efficient if and only if*

$$\cap_{i=1}^{p} \mathcal{L}_{\leq}(f_i(\hat{x})) = \cap_{i=1}^{p} \mathcal{L}_{=} f_i(\hat{x})).$$

2. *$\hat{x}$ is strictly efficient if and only if*

$$\cap_{i=1}^{p} \mathcal{L}_{\leq}(f_i(\hat{x})) = \hat{x}.$$

3. *$\hat{x}$ is weakly efficient if and only if*

$$\cap_{i=1}^{p} \mathcal{L}_{<}(f_i(\hat{x})) = \emptyset.$$

These results are fairly clear to see. For example, the first statement says that in the case that there are values of f that gives better or equal results than $f(\hat{x})$ in every component, then those values are definitely all equal and not better than $f(\hat{x})$, which is the definition of efficiency.

We illustrate the use of level sets for finding efficient solutions geometrically with an example.

Example 1. *Suppose the decision variable is the location of a building $x \in \mathbb{R}^2$ and we simultaneously want to 1) minimize the cost of constructing the building (represented by the function $f_1(x)$) and 2) keep it close to a number of posts $A = \{a_1, a_2, a_3\}$ (represented by the function $f_2(x)$). This can be interpreted as wanting to find efficient ("minimal") solutions of $f(x) = [f_1(x), f_2(x)]$. The construction site is an aproximately elliptical valley with minor axis $= 1$ and major axis $= 2$, centered at the point $(1, 1)$. Building on* *the hillsides quadratically increases construction costs. This is assumed to be the only factor which varies construction cost in our model. We can represent this as* $f_1(x) = (x_1 - 1)^2 + 2(x_2 - 1)^2$. *Then* $0 = f(1, 1)$ *is the base cost, and moving away from that center is penalized. One interpretation of minimizing the total distance to the posts in* A *is to minimize the sum of the euclidean distances from each post (see Section 5 for more about such "location problems"), so we let* $f_2(x) = \sum_{i=1}^{3} d_2(x, a_i)$. *Now let* $A = \{(2, 2), (-1, 3), (0, 0)\}$. *Now* $\hat{x} = (0.5, 1)$ *looks like it's reasonably close to the posts in* A *and to the center of the valley, so we want to find out if* $\hat{x} = (0.5, 1)$ *might be an efficient solution.* $f_1(0.5, 1) = 0.25$ *and* $f_2(0.5, 1) = 5.42$. *In Figure 2 the two level sets are plotted. We see that the intersections of their interiors are nonempty, i.e.* $\cap_{i=1}^{2} \mathcal{L}_<(f_i(\hat{x})) \neq \emptyset$. *So by Theorem 1.7,* $\hat{x}$ *isn't even weakly efficient. Every point in the intersection of the sets is superior. If the two ellipses had coincided only at the point* $\hat{x}$, *then* $\hat{x}$ *would have been strictly efficient. If they had coincided at several points along the edges, then* $\hat{x}$ *would have been only efficient (impossible in this case because of convexity).*

We see from this example that level sets are good for visualizing efficient solutions.

2 Scalarization methods

A scalarization method is a way of finding efficient solutions of a MOP by solving a related single-objective optimization problem. We do this since finding efficient solutions directly is complicated, while minimizing a single objective function is straightforward. An important scalarization method is the weighted-sum scalarization method (WSM). Here we solve the problem of minimizing a weighted sum of the components of the multi-dimensional objective function. If for example the objective function is

$$f(x) = [f_1(x), f_2(x), f_3(x)]$$

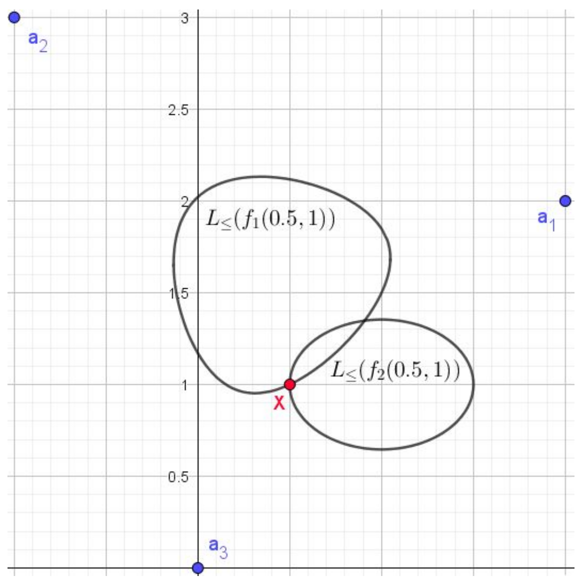

Figure 2: Level sets used to determine that $\hat{x}$ is not efficient

then we solve the problem of minimizing

$$\lambda_1 f_1(x) + \lambda_2 f_2(x) + \lambda_3 f_3(x).$$

It can be shown that doing this when all weights are nonnegative will always give efficient solutions. The other scalarization methods we will consider are the $\epsilon-$constraint method and the OWC-method, the latter of which is only relevant for uncertain optimization problems (Section 3).

Example 2. *Suppose the construction of an oil platform in the nordic sea is being planned. We are concerned with minimizing the cost of building the platform, represented as the function $f_1(x)$, and maximizing the yearly amounts of raw oil that can be drawn from the ocean on the platform, represented as the function $p(x)$, i.e. we want to minimize $f_2(x) := -p(x)$. Let the decision variable x be the model of the oil platform. Then we can find an optimal (efficient) platform-model $\hat{x}$ by minimizing*

$$g(x) = \lambda_1 f_1(x) + \lambda_2 f_2(x)$$

where λ_1 and λ_2 are weights representing the importance of each of the two objectives. Below we will prove that a minimizer of $g(x)$ is indeed an efficient solution.

2.1 The weighted sum scalarization method (WSM)

Generally for a WSM-scalarization of a MOP, the objective function we want to minimize is the sum of products $\langle \lambda, y \rangle = \sum_{i=1}^{p} \lambda_i y_i$. We write the problem as

$$\mathcal{W}(\lambda) : \min_{y \in Y} \langle \lambda, y \rangle$$

or

$$\mathcal{W}(\lambda) : \min_{x \in X} \langle \lambda, f(x) \rangle$$

depending on the context. It is common to let $\sum_{i=1}^{p} \lambda_i = 1$.

Definition 2.1. *For $\lambda \in \mathbb{R}^p_{\geq}$, the set*

$$S(\lambda, Y) := \{\hat{y} \in Y \mid \hat{y} \text{ solves } \mathcal{W}(\lambda)\}$$

is the set of optimal points for a WSM problem on the set Y.

Definition 2.2. *Let the sets $S(Y)$ and $S_0(Y)$ be defined as*

$$S(y) = \bigcup_{\lambda \in \mathbb{R}^p_{>}} S(\lambda, Y),$$

$$S_0(Y) := \bigcup_{\lambda \in \mathbb{R}^p_{\geqq}} S(\lambda, Y).$$

We see that $S_0(Y)$ is the set of solutions of all weighted sum problems on the set Y. Clearly $S(Y) \subseteq S_0(Y)$. To what extent does solving a weighted sum scalarization of a MOP give nondominated points of a set $Y \subseteq \mathbb{R}^p$? When can all efficient solutions be provided by the weighted sum method? In other words, for which λ is it true that

$$S(\lambda, Y) \subseteq Y_N,$$

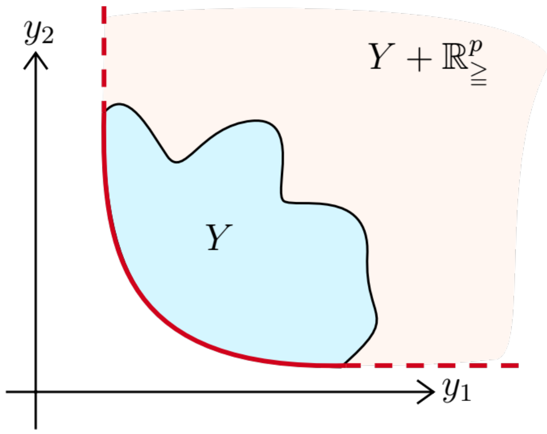

Figure 3: $\mathbb{R}^p_{\geqq}$-convex set

and when does the reverse set inclusion hold as well? To answer these questions, we first define $\mathbb{R}^p_{\geq}$-convexity, a notion which is enough for proving many useful relations on WSM, and is less restrictive than requiring convexity:

Definition 2.3. *$Y \subseteq \mathbb{R}^p$ is said to be $\mathbb{R}^p_{\geq}$- convex if $Y + \mathbb{R}^p_{\geq}$ is convex.*

$\mathbb{R}^p_{\geq}$-convexity is illustrated in Figure 3, where the set $Y + R^p_{\geq}$ defined by the red line is convex, while Y is not.

Theorem 2.1 ([1, Theorem 3.4 and Theorem 3.5])**.** *For any $Y \subseteq \mathbb{R}^p$, we have $S_0(Y) \subseteq Y_{wN}$. If also Y is $\mathbb{R}^p_{\geq}$-convex, then $Y_{wN} = S_0(Y)$.*

Proof. If some $\hat{y} \in S_0(Y)$ was not weakly nondominated, there would be some $\bar{y} \in Y$ with $\bar{y} < \hat{y} \Rightarrow \langle \lambda, y \rangle < \langle \lambda, \hat{y} \rangle$ for any $\lambda \in \mathbb{R}^p_{\geq}$. This contradicts that $\hat{y}$ would solve $\min_{y \in Y} \langle \lambda, y \rangle$ for some $\lambda \in \mathbb{R}^p_{\geq}$. For proof that $Y_{wN} \subseteq S_0(Y)$ in the case of $\mathbb{R}^p_{\geq}$-convexity, see Theorem 3.5 in [1]. □

Theorem 2.2 ([1, Theorem 3.6 and Corrolary 3.7])**.** *For any $Y \subseteq \mathbb{R}^p$, we have $S(Y) \subseteq Y_N$. If also Y is $\mathbb{R}^p_{\geq}$-convex, we have $Y_N \subseteq S_0(Y)$.*

Proof. If some $\hat{y} \in S(Y)$ was not nondominated, there would be some $\bar{y} \leq \hat{y} \Rightarrow \langle \lambda, y \rangle < \langle \lambda, y \rangle$ for any $\lambda \in \mathbb{R}^p_{>}$ contradicting that $\hat{y}$ would solve

$\min_{y\in Y}\langle\lambda, y\rangle$ for some $\lambda \in \mathbb{R}^p_>$. That $Y_N \subseteq S_0(Y)$ in the case of $\mathbb{R}^p_\geq$-convexity, follows from Theorem 2.1 and $Y_N \subseteq Y_{wN}$. □

From these theorems we directly obtain efficient solutions for a MOP $\min_{x\in X} f(x)$:

Theorem 2.3 ([1, Proposition 3.9])**.** *Let $\hat{x}$ solve the WSM problem $\mathcal{W}(\lambda)$. Then*

1. *If $\lambda \in \mathbb{R}_\geq$, then $\hat{x} \in X_{wE}$,*
2. *If $\lambda \in \mathbb{R}_>$, then $\hat{x} \in X_E$,*
3. *If $\lambda \in \mathbb{R}_\geq$ and $\hat{x}$ is the unique solution, then $\hat{x} \in X_{sE}$.*

Proof. The first point follows from Theorem 2.1, and $f(\hat{x}) \in S_0(f(X))$. The second point follows similarly from Theorem 2.2. The third point holds since $y = f(\hat{x})$ will be the unique element in $Y = f(X)$ which solves $\mathcal{W}(\lambda)$. Then $y \in Y_N$, otherwise y wouldn't be the unique element solving the minimization problem. Furthermore, since $\hat{x}$ is the only element $x \in X$ giving $f(x) = y$, it must be strictly efficient. □

We can go the other way, showing that effective solutions will solve WSM problems. This is where we use the part of Theorem 2.1 which deals with convexity.

Theorem 2.4 ([1, Proposition 3.10])**.** *Let X be convex and f be such that f_i is convex for all $i \in \{1,\dots,p\}$. Then the following is true:*

$$\hat{x} \in X_{wE} \Rightarrow \exists\lambda \in \mathbb{R}^p_\geq \textit{ s.t. } \hat{x} \textit{ solves } \mathcal{W}(\lambda).$$

Proof. Then $Y = f(X)$ will be convex, so that by 2.1,

$$\begin{aligned} Y_{wN} &= S_0(Y) \\ \Rightarrow f(\hat{x}) \in S_0(Y) &= \bigcup_{\lambda\in\mathbb{R}^p_{\geqq}} S(\lambda, Y) \end{aligned}$$

meaning that $f(\hat{x})$ solves $\mathcal{W}(\lambda)$ for some $\lambda \geq 0$. □

Note that in this theorem there is no distinction between X_E and X_{wE}. So even if we know $\hat{x} \in X_E$, we can only guarantee there exists a $\lambda \in \mathbb{R}^p_{\geq}$ and not a $\lambda \in \mathbb{R}^p_{>}$. To guarantee there exists a $\mathbb{R}^p_{>}$, we need that $\hat{x}$ is *properly efficient* as well. In short, if $\hat{x}$ is properly efficient, then any x which might improve $f(x)$ in at least one of the components of f, will give a *non-negligible* worsening of f in some other component, whereas if $\hat{x}$ is merely efficient, then there might be a x which improves f in at least one component and only *marginally* worsens f in some other component. We will not go in depth into properly efficiency here.

Example 3. *Implementing an algorithm (Algorithm 1) for solving a MOP with the weighted sum method, we obtain examples of efficient solutions for the function* $f(x) = (x^2, (x-3)^2)$. *With weights* $\lambda = [1,1], [1,2], [1,3]$, *we respectively get the efficient solutions* $x = 1.5, 2.0, 2.25$.

2.2 The ϵ-constraint method

Our next scalarization method is the ϵ-constraint method, which transforms the MOP into a constrained single-objective optimization problem. The idea behind the method is that we minimize one of the function components $f_j(x)$ under the constraint that all other components $f_k(x)$ satisfy $f_k(x) \leq \epsilon_k$, $k \neq j$, where ϵ_k is some predetermined number. Thus we want to guarantee that all components except j is within some tolerance vector $\epsilon_{-j} = (\epsilon_1, \ldots, \epsilon_{j-1}, \epsilon_{j+1}, \ldots, \epsilon_p)$, and then minimize f_j under this constraint (In general we use the subscript $-j$ of a vector to indicate that we are removing the element of index j).

Definition 2.4. *The ϵ-contraint problem, denoted ϵC_P, is*

$$\epsilon C_P(\epsilon, j) : \min_{x \in X(\epsilon, j)} f_j(x) \text{ where } X(\epsilon, j) := \{x \in X \mid f(x) \leqq \epsilon_{-j}\}.$$

In this definition, we see that the component ϵ_j of the p-dimensional vector ϵ is not used at all. This component may then be ignored or set to

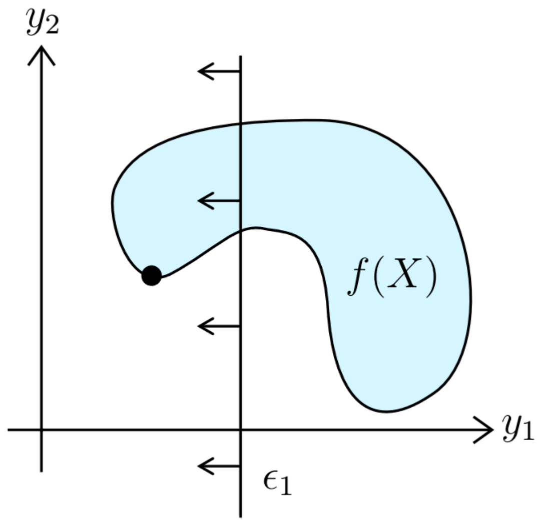

Figure 4: An efficient solution found by the ϵ-constraint method

∞. We include it mainly to simplify notation when we want to use the same ϵ for several different components $j \in \{1, \ldots, p\}$.

We'll show that a solution is efficient if and only if it solves an ϵ-constraint problem, but first we give an example.

Example 4. *Solving $\epsilon C_P(\epsilon, 2)$ for the function $f(x) = [f_1(x), f_2(x)]$ and $\epsilon = [\epsilon_1, \epsilon_2]$, we get the solution shown as the black dot in Figure 4.*

Theorem 2.5 ([1, Proposition 4.3])**.** *If $\hat{x}$ is an optimal solution of problem $\epsilon C_P(\epsilon, i)$ for any i, then $\hat{x} \in X_{wE}$.*

Proof. If $\hat{x} \notin X_{wE}$, there would be some $\bar{x} \in X$ so that $f(\bar{x}) < f(\hat{x})$. But then $f_i(\bar{x}) < f_i(\hat{x})$ and also $f_i(\bar{x}) \leq \epsilon_{-i}$, contradicting the premise. $\square$

Theorem 2.6 ([1, Proposition 4.4])**.** *If $\hat{x}$ uniquely solves problem $\epsilon C_P(\epsilon, i)$: $\min_{x \in X(\epsilon, i)} f_i(x)$ for any i and any ϵ, then $\hat{x} \in X_{sE}$.*

Proof. If $\hat{x} \notin X_{sE}$, there would be some $\bar{x} \in X$ so that $f(\bar{x}) \leqq f(\hat{x})$. But then $f_i(\bar{x}) \leq f_i(\hat{x})$ and also $f_i(\bar{x}) \leq \epsilon_{-i}$, which means that there is either a better solution to $\epsilon C_P(\epsilon, i)$ or that $\hat{x}$ is not the unique solution, contradicting the premise. $\square$

Theorem 2.7 ([1, Theorem 4.5])**.** *for a feasible $\hat{x} \in X$, we have $\hat{x} \in X_E$ iff there exists an $\hat{\epsilon} \in \mathbb{R}^p$ s.t. $\hat{x}$ an optimal solution of $\epsilon C_P(\hat{\epsilon}, i)$ for all $i = 1, 2, \ldots, p$.*

Proof. ($\Rightarrow$) when $x \in X_E$, for each j, let $\epsilon_{-j} = f_{-j}(\hat{x})$. In other words, let $\epsilon = f(\hat{x})$. Then $\hat{x}$ will be an optimal solution of $\epsilon C_P(\epsilon, j)$, for otherwise there would be some $\bar{x}$ so that $f_j(\bar{x}) < f_j(\hat{x})$ and $f_{-j}(\bar{x}) \leq \epsilon_{-j} = f_{-j}(\hat{x})$, contradicting $x \in X_E$.

($\Leftarrow$) If there is such an ϵ, then if $\hat{x}$ is not efficient, there is a $\bar{x}$ so that $f_k(\bar{x}) < f_k(\hat{x})$ for some $k \in 1, \ldots, p$ and $f_{-k}(\bar{x}) \leqq f_{-k}(\hat{x}) \leqq \epsilon_{-k}$ for some $k = \{1, \ldots, p\}$. But this contradicts the existence of ϵ since $\bar{x}$ is better than $\hat{x}$ for $\epsilon C_P(\epsilon, k)$.

□

3 Uncertain multi-objective problems and robustness

We now get to the main body of this thesis, which is showing results for uncertain multiobjective optimization problems. Many real-world problems in optimization have uncertain parameters, denoted $\theta \in U$, where U is the *uncertainty set* of possible outcomes of the uncertain parameters. This means θ is a factor that determines the outcome of the objective function f which is unknown and can be any value in U. We describe such problems as

$$P(U) : \min_{x \in X} f(x, \theta),\ \theta \in U.$$

It might be impractical to estimate the θ-parameters or use a probability-based approach to deal with them. In this case, we use the concept of robustness to solve the minimization problem. To understand robustness, consider first a single objective uncertain problem

$$P(U) : \min_{x \in X} f(x, \theta),$$

with $f : X \times U \to \mathbb{R}$. Intuitively, a robust minimizer $\hat{x} \in X$ of $P(U)$ is a point which gives tolerably small function values $f(\hat{x}, \theta)$ across the *entire* spectrum U of unknown parameters θ. When we make a robust solution, we do not consider the probability distributions of the unknown parameters, but instead try to make sure that every possibility of a parameter outcome gives "acceptable" values of the objective function f. There exist several notions of robustness. For this thesis, we will only consider the minimax concept of robustness. In the single objective case minimaxing means that we minimize the objective function f in the case of the worst possible scenario of a parameter realization $f(x, \theta)$. Mathematically, this means that we solve the problem $P_{mm}(U) : \min_{x \in X} \sup_{\theta \in U} f(x, \theta)$. Of course, in the multiobjective case, $P_{mm}(U)$ is not unambiguous due to the lack of a total order, so in the next section we expand on the analogy of minimax robustness for multidimensional functions.

Example 5. *Suppose we have the following outcome of a function for different realizations of the parameter θ:*

Values of $f(x, \theta)$			
x/θ	$\theta = \theta_1$	$\theta = \theta_2$	$\theta = \theta_3$
$x =$ *blue*	*10*	*9*	*8*
$x =$ *red*	*3*	*5*	*12*
$x =$ *green*	*14*	*7*	*6*

Following the minimax-principle, we should choose $x = blue$ since this minimizes the worst case scenario, $f(blue, \theta_1) = 10$ vs. $f(green, \theta_1) = 14$ and $f(red, \theta_3) = 12$.

3.1 Minimax-robustness for multiobjective problems

We will henceforth assume $X \subseteq \mathbb{R}^n$ and $U \subseteq \mathbb{R}^m$. We have an objective function $f : X \times U \to \mathbb{R}^p$ for which we want to find efficient solutions in the worst-case realizations of the unknown parameters. We name such a solution a robust efficient solution. It is then implicit that we talk about robustness in the minimax sense. As we mentioned in the previous section, the notions "worst-case realizations" and thus "robust efficient solution" is not unambiguous for this objective function, since the worst-case θ realizations will likely be different across the components of the function-values. To define robust efficiency, we first define the set $f_U(X)$:

Definition 3.1. *For a $x \in X$, we write $f_U(x)$ to denote the set*

$$\{f(x,\theta) \mid \theta \in U\}$$

of all values $f(x)$ can take under the possible outcomes from U.

Now, recall from Section (1.1) that a feasible point $\hat{x} \in X$ is [weakly/$\cdot$/ strictly] efficient if and only if there are no other $\bar{x} \in X \setminus \{\hat{x}\}$ such that $f(\bar{x}) \in f(\hat{x}) - \mathbb{R}^p_{[\geqq/\geq/>]}$. This can be extended to the uncertain case as follows: If for a feasible $\hat{x}$ there are no $\bar{x} \in X$ such that

$$f_U(\bar{x}) \subseteq f_U(\hat{x}) - \mathbb{R}^p_{\geq},$$

then we call $\hat{x}$ a robust efficient solution of the problem $P(U)$. In other words, if there are no $\bar{x}$ such that

$$\forall \theta \in U \exists \hat{\theta} \in U \text{ s.t. } f(\bar{x},\theta) \leq f(\hat{x},\hat{\theta}),$$

then $\hat{x}$ is robust efficient.

Definition 3.2 (Robust efficiency)**.** *a point $\hat{x} \in X$ is said to be* robust [weakly/$\cdot$/ strictly] *efficient (often shortened to* [rwe/re/rse]*) if there is no $x \in X \setminus \{\hat{x}\}$ such that*

$$f_U(x) \subseteq f_U(\hat{x}) - \mathbb{R}^p_{[>/\geq/\geqq]}.$$

The following equivalences implies alternative, intuitive notions of robust efficiency

Theorem 3.1 ([2, Lemma 3.4])**.** *For an uncertain MOP $P(U)$, and any $x, \hat{x} \in X$, the following are equivalent:*

1. $f_U(x) \subseteq f_U(\hat{x}) - \mathbb{R}^p_{[\geqq/\geq/>]}$,

2. $f_U(x) - \mathbb{R}_{\geqq} \subseteq f_U(\hat{x}) - \mathbb{R}^p_{[\geqq/\geq/>]}$,

3. $\forall \theta \in U, \exists \hat{\theta} \in U$ *s.t.* $f(x, \theta)[\leqq / \leq / <] f(\hat{x}, \hat{\theta})$.

Proof.

(1 ⇒ 2) For any $f(x,\theta) - a$, $a \geqq 0$, since $f(x,\theta) = f_U(\hat{x}, \hat{\theta}) - b$ with $b \in \mathbb{R}^p_{[\geqq/\geq/>]}$, we see that for $c = (b + a) \in \mathbb{R}^p_{[\geqq/\geq/>]} f(x,\theta) - a = f_U(\hat{x}, \hat{\theta}) - c$, proving the implication.

(2 ⇒ 1) Clear since $f_U(x) \in f_U(x) - \mathbb{R}^p_{[\geqq/\geq/>]}$.

(1 ⇒ 3) means that every element of $f_U(x)$, which can be written as $f(x,\theta)$ for some $\theta \in U$, we have that the element can also be written as an element in $f_U(\hat{x})$ (a point of the form $f(\hat{x}, \hat{\theta})$) minus some number $a\, [\geqq / \geq / >]\, 0$. This in turn means that we have $f(x,\theta)\, [\leqq / \leq / <]\, f(\hat{x}, \hat{\theta})$.

(3 ⇒ 1) For every $\theta \in U$ we get that $f(x,\theta) = f(x, \hat{\theta} - a, a \in \mathbb{R}^p_{[\geqq/\geq/>]}$ which means that every element in $f_U(x)$ can be written as an element in $f_U(\hat{x}) - \mathbb{R}^p_{[\geqq/\geq/>]}$.

□

Intuitively, when for a θ we are looking for a potential $\hat{\theta}$ as mentioned in the discussion before Definition 3.2, we will look at the θ's which give "upper-right" corner of the parameter-outcome set $f(\hat{x}, \theta) : \theta \in U$ (the red outline of $f_U(\hat{x})$ in Figure 5, where it touches $f_U(\hat{x})$), and if this upper right corner serve as an upper bound of all of $f(\bar{x}, \theta) : \theta \in U$ (i.e. it is worse than that set), then we would know $\hat{x}$ is not robust efficient. In Figure 5 we see that there does indeed exist a $\bar{x}$ such that $f_U(\bar{x})$ is completely covered by $f_U(\hat{x})$, so $\hat{x}$ is not re. In Figure 6 however, there is no such $\bar{x}$, meaning that $\hat{x}$

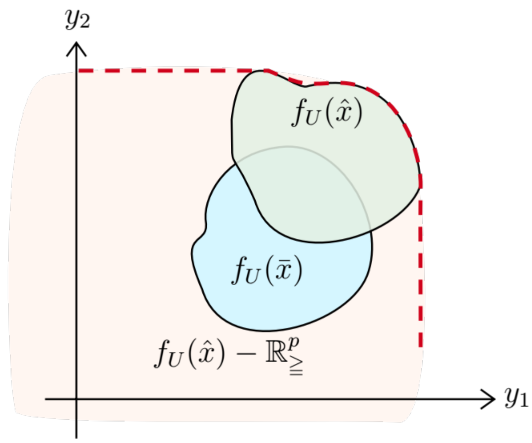

Figure 5: $\hat{x}$ is not efficient since $f_U(\bar{x}) \subset f_U(\hat{x}) - \mathbb{R}^p_{\geq}$.

is re. For the problem in Figure 6, x_2 is re as well, but not x_1 (it is dominated by x_2). We summarize the discussion in the above paragraph, and give more alternative definitions of (minimax) robust efficiency:

For strict and weak efficiency we also have the following two statements:

Theorem 3.2 ([2, Lemma 3.4]). *For an uncertain MOP $P(U)$, and any $x, \hat{x} \in X$,*

1. $f_U(x) \subseteq f_U(\hat{x}) - \mathbb{R}^p_{\geqq} \Rightarrow \sup_{\theta \in U} f_i(x, \theta) \leq \sup_{\hat{\theta} \in U} f_i(\hat{x}, \hat{\theta}) \forall i \in 1, \dots, p$
2. $f_U(x) \subseteq f_U(\hat{x}) - \mathbb{R}^p_{>} \Rightarrow \max_{\theta \in U} f_i(x, \theta) < \max_{\hat{\theta} \in U} f_i(\hat{x}, \hat{\theta}) \forall i \in 1, \dots, p$

If the two maxima in (2) exists.

Proof. 1. Since for any $\theta \in U$, $f(x, \theta)$ is an element in $f_U(\hat{x}) - \mathbb{R}^p_{\geqq}$, we have for any i that $f_i(x, \theta) \leqq f_i(\hat{x}, \hat{\theta})$ for some other $\hat{\theta} \in U$. Now we take the suprema on both sides to get the result

2. As above, for any $\theta \in U$ there is another $\hat{\theta} \in U$ so that $f(x, \theta) < f(\hat{x}, \hat{\theta})$. Now we take maximum on both sides. Since we take maximum and not supremum, we know that the strict inequality is preserved.

□

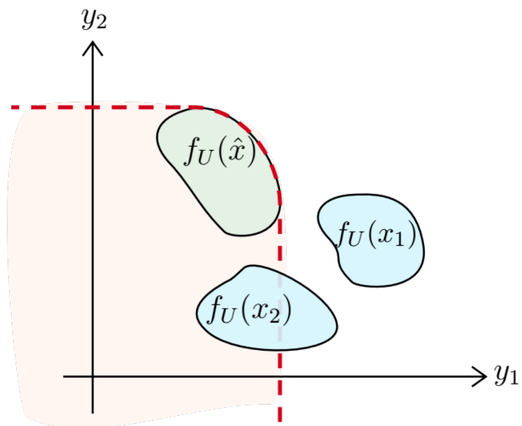

Figure 6: $\hat{x}$ is efficient since $f_U(x_i) \not\subset f_U(\hat{x}) - \mathbb{R}^p_{\geqq}$ for $i = 1, 2$.

We should show that our concepts of robust efficiency are good analogues to 1) minimax robustness for uncertain single-objective optimization problems, and 2) to the efficiency concept of deterministic MOP's. A natural way to do this is to show that for 1), a solution to an uncertain MOP $P(U)$ where the dimension of f is equal to 1 (so that it is really a single objective problem), is *re* if and only if it is a solution to the minimax problem $P_{mm}(U)$, and for 2) that for an uncertain MOP $P(U)$ with $|U| = 1$ (so that it is really a deterministic MOP), a solution is re if and only if it is an element in X_E as defined in section (1.1).

Theorem 3.3 ([2, Lemma 3.6])**.** *For a point $\hat{x} \in X$ and a function $f : X \to \mathbb{R}$ we have:*

$$\hat{x} \text{ is re} \iff \hat{x} \text{ is minimax robust.}$$

Proof.

$$
\begin{aligned}
& \hat{x} \text{ is not minimax robust} \\
\iff & \exists \bar{x} \in X \text{ s.t. } \sup_{\theta \in U} f(\bar{x}, \theta) < \sup_{\theta \in U} f(\hat{x}, \theta) \\
\iff & f(\bar{x}, \bar{\theta}) < f(\hat{x}, \hat{\theta}) \text{ for some } \hat{\theta}, \bar{\theta} \in U \\
\iff & f_U(\bar{x}) \subseteq f_U(\hat{x}) - \mathbb{R}^p_{\leq} \\
\iff & \hat{x} \text{ is not re.}
\end{aligned}
$$

□

Theorem 3.4 ([2, Lemma 3.5])**.** *For a point $\hat{x} \in X$ and a function $f : X \to \mathbb{R}^p$ and an uncertainty set U such that $|U| = 1$, we have:*

$$\hat{x} \textit{ is } \text{re} \iff \hat{x} \textit{ is efficient.}$$

Proof. Let θ be the single element in U. Then:

$$
\begin{aligned}
& \hat{x} \in X_E \\
\iff & \exists \bar{x} \in X \text{ s.t. } f(\bar{x}, \theta) \leq f(\hat{x}, \theta) \\
\iff & f_U(\bar{x}) \subseteq f_U(\hat{x}) - \mathbb{R}^p_{\leq} \text{ (since } f_U(x) = \{f(x, \theta)\}) \\
\iff & \hat{x} \text{ is not } \textit{re}.
\end{aligned}
$$

□

There exists of course completely analogous propositions to 3.3 and 3.4 for *rse* and *rwe* points. So robust efficiency is a good extension of efficiency on the one hand and robust optimality on the other hand.

4 Scalarization

Scalarization methods as introduced in section (2) can be used to obtain re/rwe/rse solutions for uncertain MOP's. We must then obtain a robust version of the method for each of the methods.

4.1 Weighted sum scalarization

Recall that for a deterministic MOP the WSM problem is written as

$$\mathcal{W}_P(\lambda) : \min_{x \in X} \langle \lambda, f(x) \rangle$$

(assume $\lambda \geq 0$ as always). The analogue to this for an uncertain MOP $P(U)$ is then to find efficient solutions of the function

$$\langle \lambda, f(x, \theta) \rangle,$$

which we know can be found using minimax robustness, leading to the deterministic single objective problem

$$\mathcal{W}_{P(U)}(\lambda) : \min_{x \in X} \sup_{\theta \in U} \langle \lambda, f(x, \theta) \rangle,$$

How useful is solving $\mathcal{W}_{P(U)}(\lambda)$ for finding re/rwe/rse solutions? We recall from section (2.1) that solving $\mathcal{W}_P(\lambda)$, $\lambda > 0$ for deterministic MOP's always gives efficient solution (Theorem 2.2). We also recall that if the feasible set X is convex, then we can obtain all efficient solutions with the WSM. For uncertain MOP's, we do have a theorem that says solving the problem $\mathcal{W}_{P(U)}$ will give re/rse/rwe solutions. But unfortunately, there are no similar theorems to Theorem 2.4 so we cannot guarantee that all re/rse/rwe solutions can be found by solving WSM-problems. We will show in example 6 a MOP where the feasible set is convex but we cannot obtain all re solutions with WSM.

Theorem 4.1 ([2, Theorem 4.3]). *For an uncertain MOP $P(U)$ the following hold:*

1. *If $\hat{x} \in X$ is the unique optimal solution to $\mathcal{W}_{P(U)}(\lambda)$, then $\hat{x}$ is rse.*
2. *If $\hat{x}$ is an optimal solution to $\mathcal{W}_{P(U)}(\lambda)$ and $\max_{\theta \in U} \langle \lambda, f(x, \theta) \rangle$ exists for all $x \in X$, then $\hat{x}$ is re.*
3. *If the conditions in (2) hold and $\lambda > 0$, then $\hat{x}$ is rwe.*

Proof. Assume $\hat{x}$ is not rse/re/rwe under the conditions described. Fixate a weight $\lambda \in \mathbb{R}^p_{[\geq/\geq/>]}$. Then:

$$\begin{aligned}
&\exists \bar{x} \in X \text{ s.t. } f_U(\bar{x}) \subseteq f_U(\hat{x}) - \mathbb{R}^p_{[\geqq/\geq/>]} \\
\Rightarrow &\forall \bar{\theta} \in U\, \exists \hat{\theta} \in U \text{ s.t. } \langle \lambda, f(\bar{x}, \bar{\theta}) \rangle \; [\leq / < / <] \; \langle \lambda, f(\hat{x}, \hat{\theta}) \rangle \\
\Rightarrow &\sup_{\theta \in U} \langle \lambda, f(\bar{x}, \theta) \rangle \; [\leq / < / <] \; \sup_{\theta \in U} \langle \lambda, f(\hat{x}, \theta) \rangle,
\end{aligned}$$

where strict inequality is preserved since $\max_{\theta \in U} \langle \lambda, f(x, \theta) \rangle$ exists in the re/rwe cases. So $\hat{x}$ does not [uniquely/$\cdot$/$\cdot$] solve $\mathcal{W}_{P(U)}(\lambda)$. $\square$

Example 6. *In Figure 7 we see that x_2 solves the robust weighted sum problem*

$$\mathcal{W}_P(U)(\lambda = [2, 1]) : \min_{x \in X} \max_{\theta \in U} 2 f_1(x, \theta) + f_2(x, \theta),$$

since by maximizing the values of $2f_1(x) + f_2(x)$ over all θ for the three possibilities of x, x_2 gives the smallest value $f(x_2) = 6$. Thus x_2 must be efficient. In Figure 8 we see a MOP where, although all solutions are efficient, only x_1 and x_3 will be obtained from the robust WSM for any value of $\lambda \geq 0$. This shows that we cannot guarantee that WSM will give all efficient solutions.

We solve a practical problem using the WSM-method

Example 7. *We wish to place a refinery near industrial sites $a_1 = 5, a_2 = 6$ and $a_3 = 8$ (positions along a one-dimensional line). Additionally, around some unknown spot between $x = 4$ and $x = 6$, represented as θ, we know there is unsteady building ground, so we would prefer to not build there. We don't have any information on theta so we let it be an uncertain parameter in the uncertainty set $U = [4, 6]$. We want to let the building be in the interval $[0, 10]$. This location problem can be modelled as the problem of finding efficient solutions of*

$$f(x) = [5e^{-2(x-\theta)^2}, (x-5)^2, (x-6)^2, (x-8)^2].$$

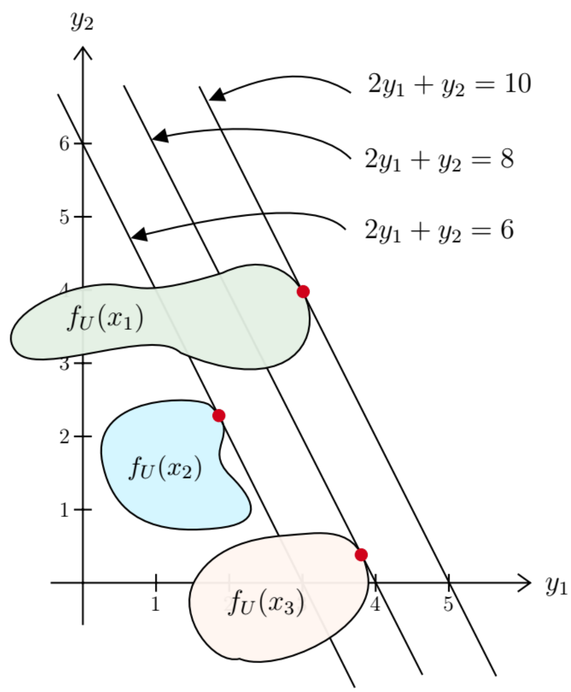

Figure 7: WSM used to obtain that x_2 is efficient

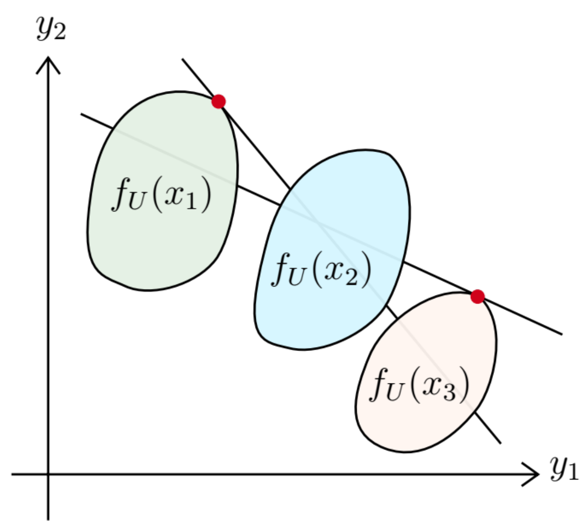

Figure 8: WSM not obtaining the efficient solution x_2

```
[6.505383015004064, 5.799999999997697, 6.910922011761761]
```

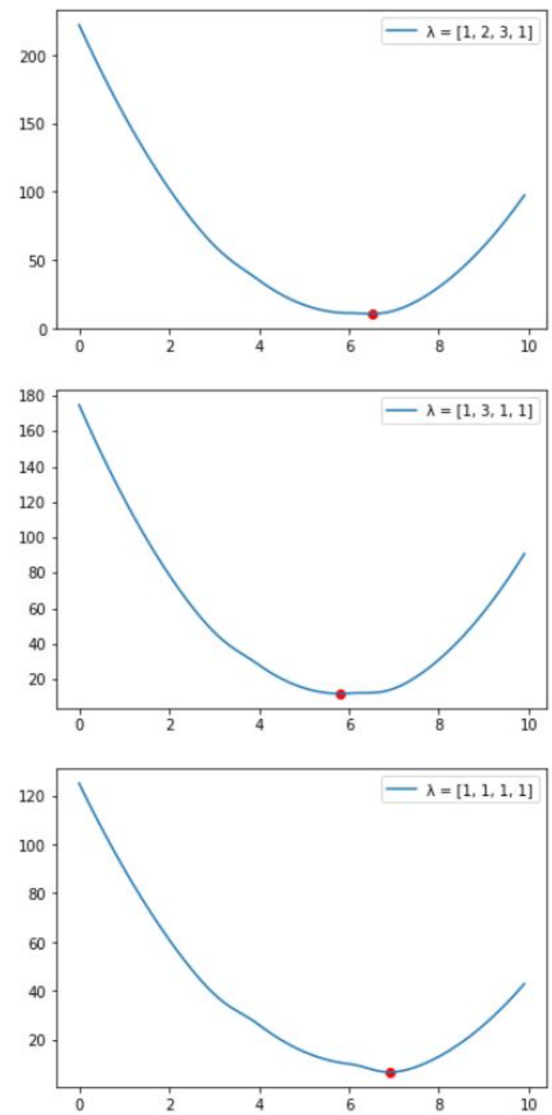

Figure 9: Efficient solutions for a location problem, using various weights

The first term is a punishment term for building close to the unsteady grounds at θ, being $= 5$ at θ and decreasing in a gaussian way the further we move from it. The other terms are the costs of building far from the sites, assumed to be quadratical. We implement the robust weighted sum method in Python 3 (Algorithm 2), using three different values of λ, and get the solutions in Figure 9 shown as the red dots on the plots of $\sup_{\theta\in U}\langle\lambda, f(x,\theta)\rangle$*. So by Theorem 4.1, the solutions* $x = 6.5, 5.8, 6.9$ *are efficient and equally optimal unless we specify additional priorities.*

In the previous example, an efficient $\hat{x}$ was relatively easy to find since X and U were one-dimensional closed intervals, and f was continuous. There are no general algorithms for solving $\mathcal{W}_{P(U)}(\lambda)$ without knowing specific properties of f, X, and U.

4.2 ϵ - constraint method

We give a robust analogue for the ϵ-constraint method. Recall from Section 2.2 that the deterministic ϵ-constraint scalarization problem is written as

$$\epsilon C_P(\epsilon, i) : \min_{x\in X} f_i(x,\theta) \text{ s.t } f_{-i}(x,\theta) \leqq \epsilon_{-i}.$$

To make ϵC_P robust, we take the minimum of $\sup_{\theta\in U} f_i(x,\theta)$ over the set of all feasible x which satisfies $f_j(x,\theta) \leqq \epsilon_j, \forall\theta \in U, \forall j \neq i$, which we'll denote $X(\epsilon, i)$. Then we are assured that all components of f other than f_i will be as low as desired no matter the uncertain parameter outcome. We call this the robust ϵ-constraint method $\epsilon C_{P(U)}$:

$$\epsilon C_{P(U)}(\epsilon, i) : \min_{x\in X(\epsilon,i)} \sup_{\theta\in U} f_i(x,\theta),$$

where $X(\epsilon, i) := \{x \in X \mid f_{-i}(x,\theta) \leqq \epsilon_{-i},\ \forall\theta \in U\}$

Now we present the analogues to theorems 2.5, 2.6 for the uncertain case:

Theorem 4.2 ([2, Theorem 4.6])**.** *For a MOP $P(U)$, any $\epsilon \in \mathbb{R}^p$, any index $i \in \{1, \dots, p\}$, and an $\hat{x} \in X$:*

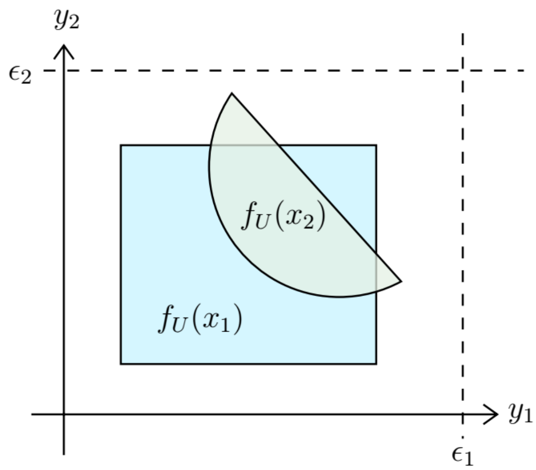

Figure 10: Situation where efficient x_2 can not be found

1. *$\hat{x}$ uniquely solves $\epsilon C_{P(U)}(\epsilon, i) \iff \hat{x}$ is rse.*
2. *$\hat{x}$ solves $\epsilon C_{P(U)}(\epsilon, i)$ and $\max_{\theta \in U} f_i(x, \theta)$ exists for all $x \in X \iff \hat{x}$ is rwe.*

There are no similar analogues for Theorem 2.7 to the uncertain case. So if we want to obtain re solutions using the ϵ-constraint method, we search for rse solutions.

Example 8. *Figure 10 shows a situation where we will never obtain the re solution x_2 with the ϵ-constraint method no matter what ϵ we choose. This is because x_1 will always be in $X(\epsilon, 1)$ or $X(\epsilon, 2)$ whenever x_2 is, and the supremum of $f_i(x_1, \theta)$ over $\theta \in U$ will always be smaller than that of $f_i(x_2, \theta)$, for $i = 1, 2$. This contrasts with Theorem (2.7) for the deterministic case, where the ϵ-constraint method will always find all efficient solutions under suitable conditions.*

4.3 Objective-wise worst case method

We now provide a method which does not have a deterministic analogue for finding efficient solutions for uncertain MOP's. For WSM and ϵ-constraint scalarization, we turn an uncertain MOP into an uncertain single-objective problem which we then solve using a minimax approach. For the objective-wise worst case method (OWC) we instead turn the uncertain MOP into a deterministic MOP which we find efficeint solutions of, typically via a deterministic scalarization method, such as those presented in this section (2). This is done by defining the function$f_U^{OWC} : X \to \mathbb{R}^p$ such that

$$f_U^{OWC}(x) = [\sup_{\theta \in U} f_1(x,\theta) \quad \dots \quad \sup_{\theta \in U} f_p(x,\theta)],$$

and then solving the problem

$$OWC_{P(U)} : \min_{x \in X} f_U^{OWC}(x).$$

So for each individual $x \in X$ we take the θ-parameter which gives the supremum of f-component j for $j = 1, 2, \dots, p$ and minimize this over all $x \in X$ (we minimize it in the pareto-optimality sense here, but the method of course also allows for other types of minimization).

Remark. *The method makes sense intuitively since when $|U| = 1$ it is reduced to a deterministic MOP and for $p = 1$ it is reduced to a single dimensional minimax problem.*

We show that solving the OWC problem gives efficient solutions.

Theorem 4.3 ([2, Theorem 4.11])**.** *Let $\hat{x} \in X$ be a strictly/weakly efficient solution to $OWC_{P(U)}$. Then $\hat{x}$ is rse/rwe for $P(U)$.*

Proof. 1. (strict) Assume $\hat{x}$ is not rse. Then there is some $\bar{x} \neq \hat{x}$ such that $f_U(\bar{x}) \subseteq f_U(\hat{x}) - \mathbb{R}^p_{\geqq}$. Then by 3.2 we'd get

$$\sup_{\theta \in U} f_i(\bar{x},\theta) \leq \sup_{\theta \in U} f_i(\hat{x},\theta) \, \forall i \in, \dots, p.$$

This means $f_U^{OWC}(\bar{x}) \leqq f_U^{OWC}(\hat{x})$, so $\hat{x}$ is not a strictly efficient solution to $OWC_{P(U)}$

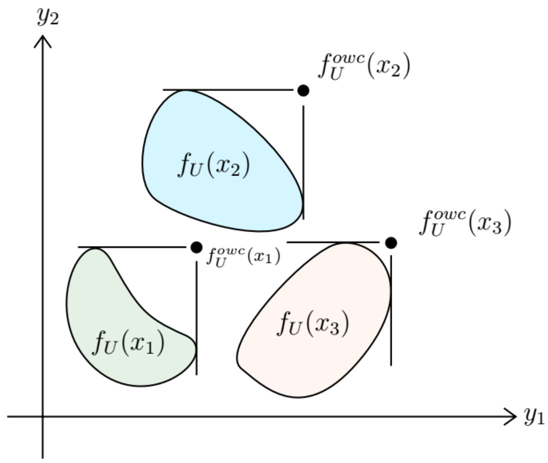

Figure 11: Candidates for efficient solutions found with OWC

2. (weak) Assume $\hat{x}$ is not rwe. Then there is some $\bar{x}$ such that $f_U(\bar{x}) \in f_U(\hat{x}) - \mathbb{R}^p_>$. Then by 3.2 we'd get

$$\max_{\theta \in U} f_i(\bar{x}, \theta) < \max_{\theta \in U} f_i(\hat{x}, \theta) \, \forall i \in 1, \dots, p.$$

(we assume the maximas exist) This means $f_U^{OWC}(\bar{x}) < f_U^{OWC}(\hat{x})$, so $\hat{x}$ is not a weakly efficient solution to $OWC_{P(U)}$.

□

Example 9. *Figure 11 illustrates Theorem 4.3. Here* $f_U^{OWC}(x)$ *is shown for* $x = x_1, x_2, x_3$. *We see that* x_1 *is strictly efficient for the problem* $OWC_{P(U)}$ *and* x_3 *is weakly efficient. Then we know* x_1 *is rse and* x_3 *is rwe for* $P(U)$.

4.4 Objective-wise uncertain problems

One problem with the OWC-method is that there might not exist any parameter realization $\theta^{\max}$ so that $f_U^{OWC}(x) = f(x, \theta^{\max})$ for some or any x. This is because $\theta_i = \arg\max_{\theta \in U} f_i(x, \theta)$ might not be the same θ as $\theta_j = \arg\max_{\theta \in U} f_j(x, \theta)$, $j \neq i$ (assuming here that the maximum does in fact exist). If we could guarantee that a single parameter $\theta^{\max}$ would provide the maximum for each component of the objective function, we could prove additional theorems on the OWC method. This would also be useful for solving practical problems. We thus introduce objective-wise uncertain problems, which concerns special classes of MOP's where the sought parameter realization θ_R is guaranteed to exist.

Definition 4.1. *A problem is said to be objective-wise uncertain (owu) if the uncertainties of each objective function component f_j are independent of each other, meaning that each function component $f_j(x, \theta)$ is either deterministic or is a function of an independent subset of the uncertainty set U only, so that*

$$U = U_1 \times U_2 \times \ldots \times U_k,$$

where each U_i is independent of the others. If a problem is owu, we can write

$$f_j(x, \theta) = f_j(x, \theta_{i(j)}),\ \theta_{i(j)} \in U_{i(j)},$$

where the uncertainty index $i = i(j) \in \{1, \ldots, k\}$ uniquely corresponds to the function index j, but is not neccesarily equal.

It is clear to see that for an owu set, we can find the sought $\theta^{\max}$ for a given x by simply solving $\arg\max_{\theta_{i(j)} \in U_{i(j)}} f_j(x, \theta_{i(j)}) =: \theta^{\max}_{i(j)}(x)$ And then set

$$\theta^{\max}(x) = [\theta_1^{\max}(x), \ldots, \theta_k^{\max}(x)] \in U_1 \times \ldots \times U_k = U. \tag{1}$$

The following theorem summarizes our discussion:

Theorem 4.4 ([2, Corrolary 5.3])**.** *For an owu problem where $\max_{\theta \in U} f_j(x)$ exists for all $j =\in 1, \ldots, p$ and all $x \in X$, we have*

$$f_U^{OWC}(x) = f(x, \theta^{\max}(x)),$$

where $\theta^{\max}(x)$ is as defined in equation (1).

For owu problems $P(U)$, we have two very useful theorems which shows the strength of the OWC-method. The first says that the solution set of the problem is identical to the solution set of the corresponding deterministic OWC-problem $OWC_{P(U)}$. So an owu problem is really a deterministic problem. The second theorem, which is a corollary to the first one, tells us that all its solutions can be found with the robust ϵ-constraint method. Recall for the next theorem that we use the [re/rwe/rse] – terminology when talking about uncertain MOP's and the [strictly/$\cdot$/weakly] efficient- terminology when talking about deterministic MOP's.

Theorem 4.5. *For an owu problem $P(U)$, where we assume that $max_{\theta \in U} f_j(x, \theta)$ exists, we have: $\hat{x}$ is rse/re/rwe for $P(U)$ if and only if it is strictly/$\cdot$/weakly efficient for $OWC_{P(U)}$.*

Proof. Since $f_U^{0WC}(x) = f(x, \theta^{\max}(x))$ where $\theta^{\max} \in U$ by 4.4 is in $f_U(x)$, we have $f_U^{OWC}(x) \in f_U(x)$, so that

$$f_U^{OWC}(x) - \mathbb{R}^p_{[\geqq/\geq/>]} \subseteq f_U(x) - \mathbb{R}^p_{[\geqq/\geq/>]}.$$

But since $f_U^{OWC}(x)$ exceeds all elements in $f_U(x)$ in every element, we also have the reverse set inclusion, so that

$$f_U^{OWC}(x) - \mathbb{R}^p_{[\geqq/\geq/>]} = f_U(x) - \mathbb{R}^p_{[\geqq/\geq/>]}.$$

Using Theorem 3.1, this means that $\hat{x}$ is rse/re/rwe for $P(U)$ if and only if it is rse/re/rwe for $OWC_{P(U)}$. Since the latter problem is deterministic, this is the same as saying if and only if $\hat{x}$ is strictly/$\cdot$/weakly efficient for that problem. $\square$

To prove the next theorem, which shows all the solutions of an owu problem may be found with the robust ϵ-constraint method, we need a lemma:

Lemma 4.6 ([2, Theorem 5.5])**.** *For an uncertain MOP $P(U)$, the set of solutions to the deterministic single objective problems $\epsilon C_{P(U)}$ and $\epsilon C_{OWC_{P(U)}}$ are identical.*

Proof. To clarify the statement: The first problem means that we use the minimax robust ϵ-constraint method on $P(U)$, and the second problem means that we use the deterministic ϵ-constraint on $OWC_{P(U)}$. The first problem can be written as $\min \sup_{\theta \in U} f_i(x, \theta)$ over all $x \in X$ such that $f_k(x, \theta) \leq \epsilon_k \, \forall k \neq i$, $\forall \theta \in U$. The second problem can be written as $\min g_i(x)$ over all $x \in X$ such that $g_k(x) \leq \epsilon_k \forall k \neq i$, where $g_j(x) := \sup_{\theta \in U} f_j(x,)$. Since these two inequalities/sets are equivalent, the lemma follows. □

Theorem 4.7 ([2, Theorem 5.6])**.** *For an owu problem $P(U)$, where we assume that $\max_{\theta \in U} f_j(x, \theta)$ exists, we have that $\hat{x} \in X$ is re $\iff \exists \hat{\epsilon} \in \mathbb{R}^p$ such that $\hat{x}$ is an optimal solution to $\epsilon C_{P(U)}(\hat{\epsilon}, i) \forall i \in 1, \ldots, p$.*

Proof. By Theorem 4.5 we have that $\hat{x}$ is efficient for $OWC_{P(U)}$. We know by Theorem 2.7 that there exists an $\hat{\epsilon} \in \mathbb{R}^p$ s.t. $\hat{x}$ solve $\epsilon C_{OWC_{P(U)}}(\hat{\epsilon}, i) \, \forall i \in 1, \ldots, p$. But then by Lemma 4.6, we have that $\hat{x}$ also solves $\epsilon C_{P(U)}(\hat{\epsilon}, i)$. □

5 A practical problem presented as a MOP

For the rest of this thesis, we will consider an application of multi objective programming to solving a practical problem. We will consider so-called *location problems*. These are problems where we have a set of several points $a_i, \ldots, a_m \in \mathbb{R}^n$ scattered in the $\mathbb{R}^n$-plane, denoted A, and want to find a point $\hat{x}$ in $\mathbb{R}^n$which is closest to those points in some sense. This is relevant in problems from industry and logistics, among others.

Example 10. *Suppose we have several docks on a coast where fish are delivered by the trawlers, and we want to find the optimal location for a processing factory in terms of proximity. Then the set L is the $\mathbb{R}^2$-coordinates of the docks, and the point $\hat{x}$ is the location of the factory.*

A natural notion of a closest point $\hat{x}$ to A is one which is an efficient solution of the problem

$$P : \min_{x \in X} D(x) = [d(a_1, x), d(a_2, x), \ldots, d(a_m, x)],$$

where $d(x) = d(a_i, x)$ is a distance metric from a_i to x. For d, we will use the euclidean distance $d_2(x, y) = \sqrt{x^2 + y^2}$, although there exists many possible choices for this metric.

5.1 Solving the location problem using WSM

We use the WSM-method to solve the location problem. So we want to solve

$$\mathcal{W}_P(\lambda) : \min_x \sum_{i=1}^{m} \lambda_i d_2(a_i, x).$$

The WSM method is particularly well-suited to location problems since there is a clear relationship between the weights λ_i and the importance of minimizing the distance from x to a_i. In our example, the size of dock i and/or relatively poor quality of the roads from dock i to x will lead to an increased priority of a close distance to that dock, so that λ_i will be increased.

WSM-problems with Euclidean distance may be approximately solved using *Weiszfeld's algorithm* ([5]). This algorithm is centered around the iteration $x_{k+1} = T(x_k)$, where

$$T(x) := \frac{\sum_{i=1}^{m} v^i \frac{a^i}{d_2(a^i,x)}}{\sum_{i=1}^{m} v^i \frac{1}{d_2(a^i,x)}}.$$

x^p is the previous estimator of the minimizer and x_{k+1} is the new one. We know from single-objective optimization that for a continuous convex function $f : \mathbb{R}^n$, a point $x^* \in \mathbb{R}^n$ is a minimizer of f if and only $\nabla f(x^*) = 0$. Denote by x^j the j-th component of x. It can be shown that

$$\frac{\partial D(x)}{\partial x^j} = x^j \sum_{i=1}^{m} v_i \frac{1}{d_2(a_i, x)} - \sum_{i=1}^{m} v_i \frac{a_i^j}{d_2(a_i, x)},$$

so we see that $\nabla f(x^*) = 0$ only when $x = T(x)$. For most practical problem the iteration $x_{k+1} = T(x_k)$ will converge to a fixed point, so the procedure is reasonable.

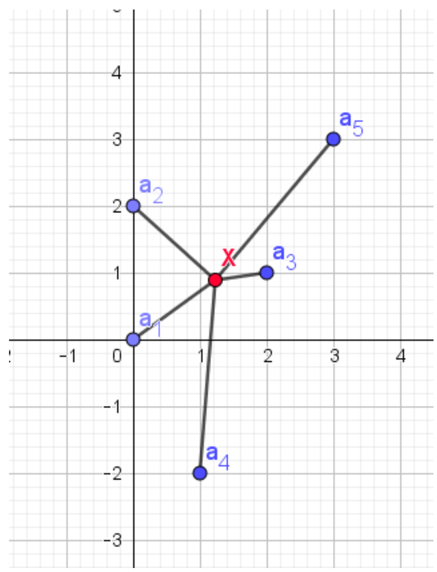

Figure 12: Optimal location of the refinery

Example 11. *Continuing our example with the fishing dock stations, if there are docks on locations* $a_1 = [0,0]$, $a_2 = [0,2]$, $a_3 = [2,1]$, $a_4 = [1,-2]$, $a_5 = [3,3]$, *and all places are considered equally important* ($\lambda = [1,1,1,1,1]$)*, then* $D(x) = [d_2(a_1,x),\dots,d_2(a_5,x)]$, *and implementing the WSM-method in Python 3 (Algorithm 3) using Weiszfeld's algorithm, we get the optimal location of the refinery as* $x = [1.230, 0.887]$, *shown as the red dot in Figure 12.*

5.2 Solving an uncertain location problem using the OWC-method

Assume now that in Example 11 there is some uncertainty as to the number of fish that will be brought in at docks $i = 1,2$. This can be represented as an additional term $\theta_i \in U_i = [0.5, 1.5]$, $i \in 1,2$, that is multiplied with the first two components of $D(x)$ to represent lowered/increased importance of getting that component as small as possible (a high θ_i representing more fish at dock i and thus an increased importance of proximity to that dock). Assume also that there is the potential of some kind of catastrophe at dock 3, determined by $\theta_3 = 0$ if nothing happens and $\theta_3 = 1$ if there is a catastrophe. So $U_3 = \{0, 1\}$. Let then the third term of D be replaced by $f(x) = (1 - \theta_3)d_2(a_3, x) + \theta_3 \frac{10}{d_2(a_3,x)}$, so that in the case of a catastrophe, the closer the factory is to dock 3, the bigger are the negative effects on the cost. We see that the uncertainty set U can be written as $U_1 \times U_2 \times U_3$, so that the problem is owu (Section 4.4).

We solve this problem with the OWC-method. We have:

$$\theta_i^{\max} = 1.5,\ i = 1, 2,$$

$$\theta_3^{\max}(x) = \left\{ \begin{array}{ll} 1, & \text{if } d_2(a_3, x) \leq \frac{10}{d_2(a_3,x)}, \\ 0, & \text{else} \end{array} \right\}.$$

That $\theta_3^{\max} = 0$ is possible might seem strange, since that means a catastrophe could increase profitability. But this is imaginable if, for example, we are required by the local government to use all the docks even when it is unprofitable to use some of them. Then if something happens at for example dock 3, so that we are permitted to not use it, then this could actually be good for minimizing $D(x)$. So we get

$$D_U^{OWC}(x) = [1.5d_2(a_1, x),\ 1.5d_2(a_2, x), f(x), d_2(a_4, x), d_2(a_5, x)].$$

This is deterministic and easily solved by the ϵ-constraint method. By Theorem 4.7, solving for all possible ϵ-constraints will give all possible optimal solutions. We could then for example use a discretization of ϵ-vectors over a suitably large ball in $\mathbb{R}^5$ to obtain a good idea of what all efficient solutions will look like.

References

[1] M. Ehrgott, *Multicriteria Optimization*, Springer, Berlin, 2005

[2] M. Ehrgott, J. Ide, A. Schöbel *Minmax robustness for multi-objective optimization problems*, European Journal of Operational Research, Elsevier, 2014

[3] J. Borwein, *On the existence of Pareto efficient points*, Mathematics of Operations research, INFORMS, 1983

[4] H. Corley, *An existence result for maximization with respect to cones*, Journal of Optimization Theory and Applications, Springer, 1980

[5] A. Beck, S.Sabach, *Weiszfeld's Method: Old and New Results.* Journal of Optimization Theory and Applications, Springer, 2015

[6] Algorithm 1, https://pastebin.com/rrEeCudU

[7] Algorithm 2, https://pastebin.com/ExTq7xej

[8] Algorithm 3, https://pastebin.com/EpxS4QUf

Notes

1. All the figures appearing in the examples in this thesis have been created by myself, but I have had help in the programming aspect of producing the figures, based on handdrawn illustrations.

NTNU

Kunnskap for en bedre verden