3.10. Use Karush-Kuhn-Tucker conditions for single objective optimization (see Theorem 3.20) and Exercise 3.9 to derive optimality conditions for efficient solutions.

3.11. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called quasi-convex if $f(\alpha x^1 + (1-\alpha)x^2) \leq \max\{f(x^1), f(x^2)\}$ for all $\alpha \in (0,1)$. It is well known that f is quasi-convex if and only if $L_{\leq}(f(x))$ is convex for all x (this is a nice exercise on level sets).

Give an Example of a multicriteria optimization problem with $\mathcal{X} \subset \mathbb{R}$ convex, $f_k : \mathbb{R} \to \mathbb{R}$ quasi-convex such that $\mathcal{X}_E$ is not connected. Hint: Monotone increasing or decreasing functions are quasi-convex, in particular those with horizontal parts in the graph.

4

Scalarization Techniques

The traditional approach to solving multicriteria optimization problems of the Pareto class is by scalarization, which involves formulating a single objective optimization problem that is related to the MOP

$$\min_{x \in \mathcal{X}} (f_1(x), \dots, f_p(x)) \tag{4.1}$$

by means of a real-valued scalarizing function typically being a function of the objective functions of the MOP (4.1), auxiliary scalar or vector variables, and/or scalar or vector parameters. Sometimes the feasible set of the MOP is additionally restricted by new constraint functions related to the objective functions of the MOP and/or the new variables introduced.

In Chapter 3 we introduced the "simplest" method to solve multicriteria problems, the weighted sum method, where we solve

$$\min_{x \in \mathcal{X}} \sum_{k=1}^{p} \lambda_k f_k(x). \tag{4.2}$$

The weighted sum problem (4.2) uses the vector of weights $\lambda \in \mathbb{R}^p_{\geq}$ as a parameter. We have seen that the method enables computation of the properly efficient and weakly efficient solutions for convex problems by varying λ. The following Theorem summarizes the results.

Theorem 4.1. *1. Let $\hat{x} \in \mathcal{X}$ be an optimal solution of (4.2). The following statements hold.*
- *If $\lambda > 0$ then $\hat{x} \in \mathcal{X}_{pE}$.*
- *If $\lambda \geq 0$ then $\hat{x} \in \mathcal{X}_{wE}$.*
- *If $\lambda \geq 0$ and $\hat{x}$ is a unique optimal solution of (4.2) then $\hat{x} \in \mathcal{X}_{sE}$.*

2. Let $\mathcal{X}$ be a convex set and $f_k, k = 1, \dots, p$ be convex functions. Then the following statements hold.

- *If* $\hat{x} \in \mathcal{X}_{pE}$ *then there is some* $\lambda > 0$ *such that* $\hat{x}$ *is an optimal solution of* (4.2).
- *If* $\hat{x} \in \mathcal{X}_{wE}$ *then there is some* $\lambda \geq 0$ *such that* $\hat{x}$ *is an optimal solution of* (4.2).

For nonconvex problems, however, it may work poorly. Consider the following example.

Example 4.2. Let $\mathcal{X} = \{x \in \mathbb{R}^2_{\geqq} : x_1^2 + x_2^2 \geq 1\}$ and $f(x) = x$. In this case $\mathcal{X}_E = \{x \in \mathcal{X} : x_1^2 + x_2^2 = 1\}$, yet $\hat{x}^1 = (1,0)$ and $\hat{x}^2 = (1,0)$ are the only feasible solutions that are optimal solutions of (4.2) for any $\lambda \geq 0$.

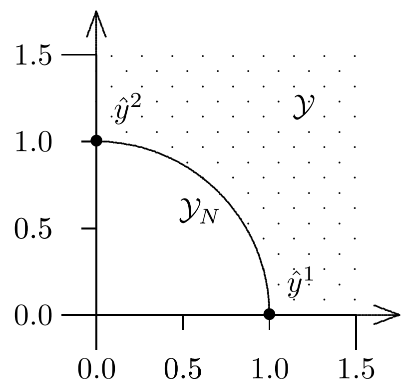

Fig. 4.1. The weighted sum method fails for nonconvex problems.

□

In this chapter we introduce some other scalarization methods, which are also applicable when $\mathcal{Y}$ is not $\mathbb{R}^p_{\geqq}$-convex.

4.1 The ε-Constraint Method

Besides the weighted sum approach, the ε-constraint method is probably the best known technique to solve multicriteria optimization problems. There is no aggregation of criteria, instead only one of the original objectives is minimized, while the others are transformed to constraints. It was introduced byHaimes *et al.* (1971), and an extensive discussion can be found in Chankong and Haimes (1983).

We substitute the multicriteria optimization problem (4.1) by the ε-constraint problem

$$\begin{aligned} &\min_{x \in \mathcal{X}} f_j(x) \\ \text{subject to } &f_k(x) \leq \varepsilon_k \quad k = 1, \ldots, p \quad k \neq j, \end{aligned} \tag{4.3}$$

where $\varepsilon \in \mathbb{R}^p$. The component ε_j is irrelevant for (4.3), but the convention to include it will be convenient later.

Figure 4.2 illustrates a bicriterion problem, where an upper bound constraint is put on $f_1(x)$. The optimal values of (4.3) problem with $j = 2$ for two values of ε_1 are indicated. These show that the constraints $f_k(x) \leq \varepsilon_k$ might or might not be active at an optimal solution of (4.3).

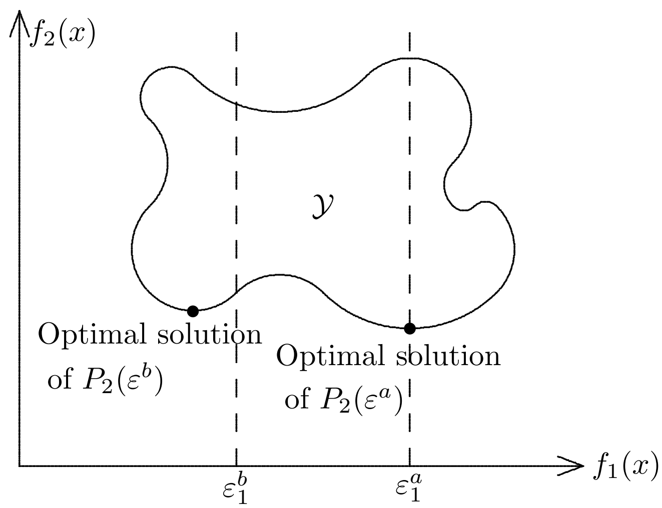

Fig. 4.2. Optimal solutions of ε-constraint problems (4.3).

To justify the approach we show that optimal solutions of (4.3) problems are at least weakly efficient. A necessary and sufficient condition for efficiency shows that this method works for general problems, no convexity assumption is needed. We will also prove a result relating (4.3) to the weighted sum problem (4.2).

Proposition 4.3. *Let $\hat{x}$ be an optimal solution of* (4.3) *for some j. Then $\hat{x}$ is weakly efficient.*

Proof. Assume $\hat{x} \notin \mathcal{X}_{wE}$. Then there is an $x \in \mathcal{X}$ such that $f_k(x) < f_k(\hat{x})$ for all $k = 1, \ldots, p$. In particular, $f_j(x) < f_j(\hat{x})$. Since $f_k(x) < f_k(\hat{x}) \leq \varepsilon_k$ for $k \neq j$, the solution x is feasible for (4.3). This is a contradiction to $\hat{x}$ being an optimal solution of (4.3). $\square$

In order to strengthen Proposition 4.3 to obtain efficiency we require the optimal solution of (4.3) to be unique. Note the similarity to Theorem 3.4 and Proposition 3.8 for the weighted sum scalarization.

Proposition 4.4. *Let $\hat{x}$ be a unique optimal solution of* (4.3) *for some j. Then $\hat{x} \in \mathcal{X}_{sE}$ (and therefore $\hat{x} \in \mathcal{X}_E$).*

Proof. Assume there is some $x \in \mathcal{X}$ with $f_k(x) \leq f_k(\hat{x}) \leq \varepsilon_k$ for all $k \neq j$. If in addition $f_j(x) \leq f_j(\hat{x})$ we must have $f_j(x) = f_j(\hat{x})$ because $\hat{x}$ is an optimal solution of (4.3). So x is an optimal solution of (4.3). Thus, uniqueness of the optimal solution implies $x = \hat{x}$ and $\hat{x} \in \mathcal{X}_{sE}$. □

In general, efficiency of $\hat{x}$ is related to $\hat{x}$ being an optimal solution of (4.3) for all $j = 1, \ldots, p$ with the same ε used in all of these problems.

Theorem 4.5. *The feasible solution $\hat{x} \in \mathcal{X}$ is efficient if and only if there exists an $\hat{\varepsilon} \in \mathbb{R}^p$ such that $\hat{x}$ is an optimal solution of* (4.3) *for all $j = 1, \ldots, p$.*

Proof. "$\Longrightarrow$" Let $\hat{\varepsilon} = f(\hat{x})$. Assume $\hat{x}$ is not an optimal solution of (4.3) for some j. Then there must be some $x \in \mathcal{X}$ with $f_j(x) < f_j(\hat{x})$ and $f_k(x) \leq \hat{\varepsilon}_k = f_k(\hat{x})$ for all $k \neq j$, that is, $\hat{x} \notin \mathcal{X}_E$.

"$\Longleftarrow$" Suppose $\hat{x} \notin \mathcal{X}_E$. Then there is an index $j \in \{1, \ldots, p\}$ and a feasible solution $x \in \mathcal{X}$ such that $f_j(x) < f_j(\hat{x})$ and $f_k(x) \leq f_k(\hat{x})$ for $k \neq j$. Therefore $\hat{x}$ cannot be an optimal solution of (4.3) for any ε for which it is feasible. Note that any such ε must have $f_k(\hat{x}) \leq \varepsilon_k$ for $k \neq j$. □

Theorem 4.5 shows that with appropriate choices of ε all efficient solutions can be found. However, as the proof shows, these ε_j values are equal to the actual objective values of the efficient solution one would like to find. A confirmation or check of efficiency is obtained rather than the discovery of efficient solutions.

We denote by

$$\mathcal{E}_j := \{\varepsilon \in \mathbb{R}^p : \{x \in \mathcal{X} : f_k(x) \leq \varepsilon_k,\ k \neq j\} \neq \emptyset\}$$

the set of right hand sides for which (4.3) is feasible and by

$$\mathcal{X}_j(\varepsilon) := \{x \in \mathcal{X} : x \text{ is an optimal solution of (4.3)}\}$$

for $\varepsilon \in \mathcal{E}_j$ the set of optimal solutions of (4.3). From Theorem 4.5 and Proposition 4.3 we have that for each $\varepsilon \in \cap_{j=1}^p \mathcal{E}_j$

$$\bigcap_{j=1}^{p} \mathcal{X}_j(\varepsilon) \subset \mathcal{X}_E \subset \mathcal{X}_j(\varepsilon) \subset \mathcal{X}_{wE} \tag{4.4}$$

for all $j = 1, \ldots, p$ (cf. (3.37) for weighted sum scalarization).

Our last result in this section provides a link between the weighted sum method and the ε-constraint method.

Theorem 4.6 (Chankong and Haimes (1983)).

1. *Suppose $\hat{x}$ is an optimal solution of* $\min_{x \in \mathcal{X}} \sum_{k=1}^{p} \lambda_k f_k(x)$. *If $\lambda_j > 0$ there exists $\hat{\varepsilon}$ such that $\hat{x}$ is an optimal solution of* (4.3), *too.*
2. *Suppose $\mathcal{X}$ is a convex set and $f_k : \mathbb{R}^n \to \mathbb{R}$ are convex functions. If $\hat{x}$ is an optimal solution of* (4.3) *for some j, there exists $\hat{\lambda} \in \mathbb{R}^p_{\geq}$ such that $\hat{x}$ is optimal for* $\min_{x \in \mathcal{X}} \sum_{k=1}^{p} \hat{\lambda}_k f_k(x)$.

Proof. 1. As in the previous proof we show that we can set $\hat{\varepsilon} = f(\hat{x})$. From optimality of $\hat{x}$ for a weighted sum problem we have

$$\sum_{k=1}^{p} \lambda_k (f_k(x) - f_k(\hat{x})) \geq 0$$

for all $x \in \mathcal{X}$. Suppose $\hat{x}$ is not optimal for (4.3) with right hand sides $\hat{\varepsilon}$. The contradiction follows from the fact that for any $x' \in \mathcal{X}$ with $f_j(x') < f_j(\hat{x})$ and $f_k(x') \leq f_k(\hat{x})$ for $k \neq j$

$$\lambda_j (f_j(x') - f_j(\hat{x})) + \sum_{k \neq j} \lambda_k (f_k(x^0) - f_k(\hat{x})) < 0 \tag{4.5}$$

because $\lambda_j > 0$.

2. Suppose $\hat{x}$ solves (4.3) optimally. Then there is no $x \in \mathcal{X}$ satisfying $f_j(x) < f_j(\hat{x})$ and $f_k(x) \leq f_k(\hat{x}) \leq \varepsilon_k$ for $k \neq j$. Using convexity of f_k we apply Theorem 3.16 to conclude that there must be some $\hat{\lambda} \in \mathbb{R}^p_{\geq}$ such that $\sum_{k=1}^{p} \hat{\lambda}_k (f_k(x) - f_k(\hat{x})) \geq 0$ for all $x \in \mathcal{X}$. Since $\hat{\lambda} \in \mathbb{R}^p_{\geq}$ we get

$$\sum_{k=1}^{p} \hat{\lambda}_k f_k(x) \geq \sum_{k=1}^{p} \hat{\lambda}_k f_k(\hat{x}) \tag{4.6}$$

for all $x \in \mathcal{X}$. Therefore $\hat{\lambda}$ is the desired weighting vector. □

A further result in this regard, showing when an optimal solution of the weighted sum problem is also an optimal solution of the (4.3) problem for all $j = 1, \ldots, p$ is given as Exercise 4.1.

4.2 The Hybrid Method

It is possible to combine the weighted sum method with the ε-constraint method. In that case, the scalarized problem to be solved has a weighted sum objective and constraints on *all* objectives. Let x^0 be an arbitrary feasible point for an MOP. Consider the following problem:

$$\begin{aligned} \min \sum_{k=1}^{p} \lambda_k f_k(x) \\ \text{subject to } f_k(x) \leq f_k(x^0) \quad k = 1, \ldots, p \\ x \in \mathcal{X} \end{aligned} \tag{4.7}$$

where $\lambda \in \mathbb{R}^p_{\geq}$.

Theorem 4.7. *Guddat* et al. *(1985) Let $\lambda \in \mathbb{R}^p_{>}$. A feasible solution $x^0 \in \mathcal{X}$ is an optimal solution of problem (4.7) if and only if $x^0 \in \mathcal{X}_E$.*

Proof. Let $x^0 \in \mathcal{X}$ be efficient. Then there is no $x \in \mathcal{X}$ such that $f(x) \leq f(x^0)$. Thus any feasible solution of (4.7) satisfies $f(x) = f(x^0)$ and is an optimal solution.

Let x^0 be an optimal solution of (4.7). If there were an $x \in \mathcal{X}$ such that $f(x) \leq f(x^0)$ the positive weights would imply

$$\sum_{k=1}^{p} \lambda_k f_k(x) < \sum_{k=1}^{p} \lambda_k f_k(x^0).$$

Thus x^0 is efficient. □

4.3 The Elastic Constraint Method

For the ε-constraint method we have no results on properly efficient solutions. In addition, the scalarized problem (4.3) may be hard to solve in practice due to the added constraints $f_k(x) \leq \varepsilon_k$. In order to address this problem we can "relax" these constraints by allowing them to be violated and penalizing any violation in the objective function. Ehrgott and Ryan (2002) used this idea to develop the e;elastic constraint scalarization

$$\begin{aligned} \min f_j(x) + \sum_{k \neq j} \mu_k s_k \\ \text{subject to } f_k(x) - s_k \leq \varepsilon_k \quad k \neq j \\ s_k \geq 0 \quad k \neq j \\ x \in \mathcal{X}, \end{aligned} \tag{4.8}$$

where $\mu_k \geq 0, k \neq j$. The feasible set of (4.8) in x variables is $\mathcal{X}$, i.e. the feasible set of the original multicriteria optimization problem (4.1).

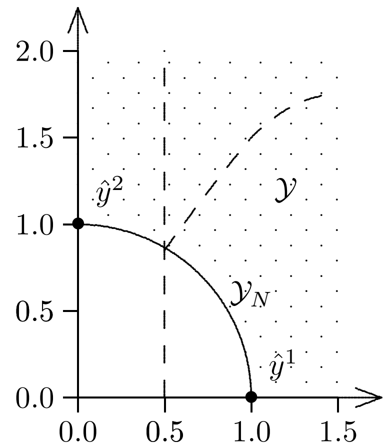

Fig. 4.3. Feasible set and objective function of Problem (4.8).

Note that if $(\hat{x}, \hat{s})$ is an optimal solution of (4.8), then we may without loss of generality assume that $\hat{s}_k = \max\{0, \varepsilon_k - f_k(\hat{x})\}$.

In Figure 4.3 (4.8) for $j = 2$ is illustrated for the bicriterion problem of Example 4.2. The vertical dotted line marks the value $\varepsilon_1 = 0.5$. The dotted curve shows the objective function of (4.8) as a function of component y_1 of nondominated points $\mathcal{Y}_N$. The idea of the method is that, by penalizing violations of the constraint $f_1(x) \leq \varepsilon_1$, a minimum is attained with the constraint active. As can be seen here, the minimum of (4.8) will be attained at $x = (0.5, 0.5)$.

We obtain the following results:

Proposition 4.8. *Let* $(\hat{x}, \hat{s})$ *be an optimal solution of* (4.8) *with* $\mu \geqq 0$. *Then* $\hat{x} \in \mathcal{X}_{wE}$.

Proof. Suppose $\hat{x}$ is not weakly efficient. Then there is some $x \in \mathcal{X}$ such that $f_k(x) < f_k(\hat{x}), k = 1, \ldots, p$. Then $(x, \hat{s})$ is feasible for (4.8) with an objective value that is smaller than that of $(\hat{x}, \hat{s})$. □

Under additional assumptions we get stronger results.

Proposition 4.9. *If* $\hat{x}$ *is unique in an optimal solution of* (4.8), *then* $\hat{x} \in \mathcal{X}_{sE}$ *is a strictly efficient solution of the MOP.*

Proof. Assume that $x \in \mathcal{X}$ is such that $f_k(x) \leq f_k(\hat{x}), k = 1, \ldots, p$. Then $(x, \hat{s})$ is a feasible solution of (4.8). Since the objective function value of $(x, \hat{s})$ is not worse than that of $(\hat{x}, \hat{s})$, uniqueness of $\hat{x}$ implies that $x = \hat{x}$. □

The following example shows that even if $\mu > 0$ an optimal solution of (4.8) may be just weakly efficient.

Example 4.10. Consider

$$\mathcal{X} = \left\{ (x_1, x_2) \in \mathbb{R}^2_{\geqq} : (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1 \right\} + \mathbb{R}^2_{\geqq}$$

and $f(x) = x$. Let $\varepsilon_1 > 1$. Then $(\hat{x}_1, \hat{x}_2, \hat{s}_1) = (\hat{x}_1, 0, 0)$ is an optimal solution of (4.8) with $j = 2$ for all $1 \leq \hat{x}_1 \leq \varepsilon_1$. If $\hat{x}_1 > 1$ this solution is weakly efficient, but not efficient. This result is independent of the choice of μ. □

The problem here is the possible existence of weakly efficient solutions that satisfy the constraints $f_k(x) \leq \varepsilon_k$ for all $k \neq j$. If, however, all ε_k are chosen in such a way that no merely weakly efficient solution satisfies the ε-constraints, an optimal solution of (4.8) with $\mu > 0$ will yield an efficient solution.

We now turn to the problem of showing that (properly) efficient solutions are optimal solutions of (4.8) for appropriate choices of k, ε, and μ. The following corollary follows immediately from Theorem 4.5 by choosing $\varepsilon = f(\hat{x})$, $\hat{s} = 0$ and $\mu_k = \infty$ for all $k = 1, \ldots, p$.

Corollary 4.11. *Let $\hat{x} \in \mathcal{X}_E$. Then there exist $\varepsilon, \mu \geqq 0$ and $\hat{s}$ such that $(\hat{x}, \hat{s})$ is an optimal solution of (4.8) for all $j \in \{1, \ldots, p\}$.*

A more careful analysis shows that for properly efficient solutions, we can do without the infinite penalties.

Theorem 4.12. *Let $\mathcal{Y}_N$ be externally stable. Let $\hat{x} \in \mathcal{X}_{pE}$ be properly efficient. Then, for every $j \in \{1, \ldots, p\}$ there are $\varepsilon, \hat{s}, \mu^j$ with $\mu^j_k < \infty$ for all $k \neq j$ such that $(\hat{x}, \hat{s})$ is an optimal solution of (4.3) for all $\mu \in \mathbb{R}^{p-1}$, $\mu \geqq \mu^j$.*

Proof. We choose $\varepsilon_k := f_k(\hat{x}), k = 1, \ldots, p$. Thus, we can choose $\hat{s} = 0$. Let $j \in \{1, \ldots, p\}$. Because $\hat{x}$ is properly efficient there is $M > 0$ such that for all $x \in \mathcal{X}$ with $f_j(x) < f_j(\hat{x})$ there is $k \neq j$ such that $f_k(\hat{x}) < f_k(x)$ and $\frac{f_j(\hat{x}) - f_j(x)}{f_k(x) - f_k(\hat{x})} < M$.

We define μ^j by $\mu^j_k := \max(M, 0)$ for all $k \neq j$.

Let $x \in \mathcal{X}$ and $s \in \mathbb{R}$ be such that $s_k = \max\{0, f_k(x) - \varepsilon_k\} = \max\{0, f_k(x) - f_k(\hat{x})\}$ for all $k \neq j$, i.e. the smallest possible value it can take. We need to show that

$$f_j(x) + \sum_{k \neq j} \mu_k s_k \geq f_j(\hat{x}) + \sum_{k \neq j} \mu_k \hat{s}_k = f_k(\hat{x}). \tag{4.9}$$

First, we prove that we can assume $x \in \mathcal{X}_E$ in (4.9). Otherwise there is $x' \in \mathcal{X}_E$ with $f(x') \leq f(x)$ (because $\mathcal{Y}_N$ is externally stable, see Definition

2.20) and s' with $s'_k = \max\{0, f_k(x') - \varepsilon_k\}$. Since $s' \leqq s$ we get that $f_j(x') + \sum_{k \neq j} \mu_k s'_k \leq f_k(x) + \sum_{k \neq j} \mu_k s_k$ for any $\mu \geqq 0$.

Now let $x \in \mathcal{X}_E$. We consider the case $f_j(x) \geq f_j(\hat{x})$. Then

$$f_j(x) + \sum_{k \neq j} \mu_k s_k > f_k(\hat{x}) + 0 = f_j(\hat{x}) + \sum_{k \neq j} \mu_k^j \hat{s}_k$$

for any $\mu \geq 0$.

Now consider the case $f_j(x) < f_j(\hat{x})$ and let $\mathcal{I}(x) := \{k \neq j : f_k(x) > f_k(\hat{x})\}$. As both $\hat{x}$ and x are efficient, $\mathcal{I}(x) \neq \emptyset$. Furthermore, we can assume $s_k = 0$ for all $k \notin \mathcal{I}(x), k \neq j$. Let $k' \in \mathcal{I}(x)$. Then

$$\begin{aligned}
f_j(x) + \sum_{k \neq j} \mu_k s_k &\geq f_j(x) + \sum_{k \neq j} \mu_k^j s_k \\
&\geq f_j(x) + \sum_{k \in \mathcal{I}(x)} \frac{f_j(\hat{x}) - f_j(x)}{f_k(x) - f_k(\hat{x})} s_k \\
&\geq f_j(x) + \frac{f_j(\hat{x}) - f_j(x)}{f_{k'}(x) - f_{k'}(\hat{x})} s_{k'} \\
&= f_j(x) + \frac{f_j(\hat{x}) - f_j(x)}{f_{k'}(x) - f_{k'}(\hat{x})} (f_{k'}(x) - f_{k'}(\hat{x})) \\
&= f_j(\hat{x}) = f_j(\hat{x}) + \sum_{k \neq j} \mu_k \hat{s}_k.
\end{aligned}$$

This follows from $\mu_k \geq \mu_k^j$, the definition of μ_k^j, nonnegativity of all terms, $s_k = f_k(x) - f_k(\hat{x})$ for $k \in \mathcal{I}(x)$ and $\hat{s} = 0$. $\square$

We can also see, that for $x \in \mathcal{X}_E \setminus \mathcal{X}_{pE}$ finite values of μ are not sufficient.

Example 4.13. Let $p = 2$ and $\mathcal{X} = \{x \in \mathbb{R}^2 : (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1\}$ with $f(x) = x$. Then $(1, 0)$ and $(0, 1)$ are efficient, but not properly efficient. The scalarization

$$\begin{aligned}
\min x_2 + \mu s& \\
\text{subject to } x_1 - s &\leq 0 \\
x &\in \mathcal{X}
\end{aligned}$$

is equivalent to (has the same optimal solution x as)

$$\min\{x_2 + \mu x_1 : (x_1 - 1)^2 + (x_2 - 1)^2 = 1\}.$$

It is easy to see that the unique optimal solution is given by $x_1 = 1 - \sqrt{1 - \frac{1}{\mu + 1}}$ and it is necessary that $\mu \to \infty$ to get $x_1 \to 0$.

Note, however, that in order to obtain $(0, 1)$, we can also consider

$$\begin{aligned} \min\, & x_1 + \mu s \\ \text{subject to } & x_2 - s \leq 1 \\ & x \in \mathcal{X}. \end{aligned}$$

It is clear that $x_1 = 0, x_2 = 1, s = 0$ is an optimal solution of this problem for any $\mu \geq 0$. □

It is worth noting that in the elastic constraint method ε-constraints of (4.3) are relaxed in a manner similar to penalty function methods in nonlinear programming. This may help solving the scalarized problem in practice.

4.4 Benson's Method

The method and results described in this section are from Benson (1978). The idea is to choose some initial feasible solution $x^0 \in \mathcal{X}$ and, if it is not itself efficient, produce a dominating solution that is. To do so nonnegative deviation variables $l_k = f_k(x^0) - f_k(x)$ are introduced, and their sum maximized. This results in an x dominating x^0, if one exists, and the objective ensures that it is efficient, pushing x as far from x^0 as possible.

The substitute problem (4.10) for given x^0 is

$$\begin{aligned} \max & \sum_{k=1}^{p} l_k \\ \text{subject to } & f_k(x^0) - l_k - f_k(x) = 0 \quad k = 1, \ldots, p \\ & l \geqq 0 \\ & x \in \mathcal{X}. \end{aligned} \tag{4.10}$$

An illustration in objective space (Figure 4.4) demonstrates the idea. The initial feasible, but dominated, point $f(x^0)$ has values greater than the efficient point $f(\hat{x})$. Maximizing the total deviation $\hat{l}_1 + \hat{l}_2$, the intention is to find a dominating solution, which is efficient.

First of all, solving (4.10) is a check for efficiency of the initial solution x^0 itself. We will see this result again later, when we deal with linear problems in Chapter 6.

Theorem 4.14. *The feasible solution $x^0 \in \mathcal{X}$ is efficient if and only if the optimal objective value of* (4.10) *is* 0.

Proof. Let (x, l) be a feasible solution of (4.10). Because of the nonnegativity constraint $l_k \geq 0$ for $k = 1, \ldots, p$ and the definition of l_k as $f_k(x^0) - f_k(x)$ we have

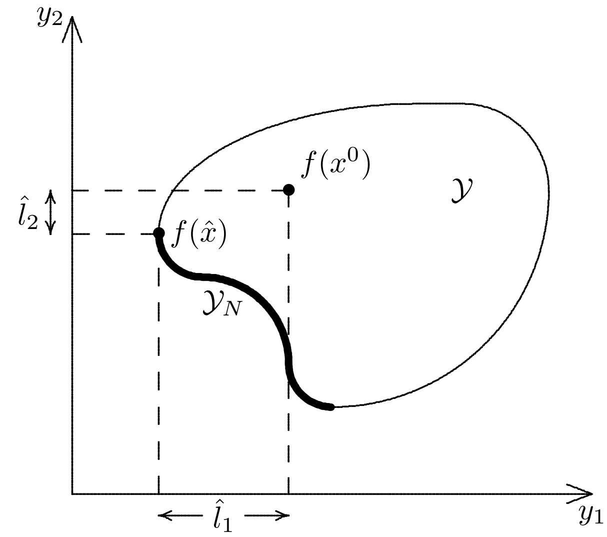

Fig. 4.4. Illustration of Benson's problem (4.10).

$$\sum_{k=1}^{p} l_k = 0 \Longleftrightarrow l_k = 0 \quad k = 1, \ldots, p$$
$$\Longleftrightarrow f_k(x^0) = f_k(x) \quad k = 1, \ldots, p$$

Thus, if the optimal value is 0, and $x \in \mathcal{X}$ is such that $f(x) \leqq f(x^0)$ it must hold that $f(x) = f(x^0)$, i.e. x^0 is efficient. If, on the other hand, x^0 is efficient, the feasible set of (4.10) consists of those (x, l) for which $x \in \mathcal{X}$ and $f(x) = f(x^0)$ and thus $l = 0$. $\square$

That the initial solution x^0 is efficient cannot be expected in general. The strength of the method lies in the fact that whenever problem (4.10) has a finite optimal solution value, the optimal solution is efficient. Under convexity assumptions, we can even show that when the objective function of (4.8) is unbounded, no properly efficient solutions exist. From an application point of view, this constitutes a pathological situation: all efficient solutions will have unbounded trade-offs. However, this can only happen in situations where existence of efficient solutions is not guaranteed in general.

Proposition 4.15. *If problem* (4.10) *has an optimal solution* $(\hat{x}, \hat{l})$ *(and the optimal objective value is finite) then* $\hat{x} \in \mathcal{X}_E$.

Proof. Suppose $\hat{x} \notin \mathcal{X}_E$. Then there is some $x' \in \mathcal{X}$ such that $f_k(x') \leq f_k(\hat{x})$ for all $k = 1, \ldots, p$ and $f_j(x') < f_j(\hat{x})$ for at least one j. We define $l' := f(x^0) - f(x')$. Then (x', l') is feasible for (4.10) because

$$l'_k = f_k(x^0) - f_k(x') \geq f_k(x^0) - f_k(\hat{x}) = \hat{l}_k \geq 0. \tag{4.11}$$

Furthermore, $\sum_{k=1}^p l'_k > \sum_{k=1}^p \hat{l}_k$ as $l'_j > \hat{l}_j$. This is impossible because $(\hat{x}, \hat{l})$ is an optimal solution of (4.10). □

The question of what happens if (4.10) is unbounded can be answered under convexity assumptions.

Theorem 4.16 (Benson (1978)). *Assume that the functions $f_k, k = 1, \dots, p$ are convex and that $\mathcal{X} \subset \mathbb{R}^n$ is a convex set. If (4.10) has no finite optimal objective value then $\mathcal{X}_{pE} = \emptyset$.*

Proof. Since (4.10) is unbounded, for every real number $M \geq 0$ we can find $x^M \in \mathcal{X}$ such that $l = f(x^0) - f(x^M) \geq 0$ and

$$\sum_{k=1}^p l_k = \sum_{k=1}^p (f_k(x^0) - f_k(x^M)) > M. \tag{4.12}$$

Assume that $\hat{x}$ is properly efficient in Geoffrion's sense. From Theorem 3.15 we know that there are weights $\lambda_k > 0$ for $k = 1, \dots, p$ such that $\hat{x}$ is an optimal solution of $\min_{x \in \mathcal{X}} \sum_{k=1}^p \lambda_k f_k(x)$. Therefore $\sum_{k=1}^p \lambda_k (f_k(x) - f_k(\hat{x})) \geq 0$ for all $x \in \mathcal{X}$, and in particular

$$\sum_{k=1}^p \lambda_k (f_k(x^0) - f_k(\hat{x})) \geq 0. \tag{4.13}$$

We define $\hat{\lambda} := \min\{\lambda_1, \dots, \lambda_p\} > 0$ and for some arbitrary, but fixed $M' \geq 0$ let $M := M'/\hat{\lambda}$. From (4.12) we know that for this M there is some $x^M \in \mathcal{X}$ satisfying $f_k(x^0) - f_k(x^M) \geq 0$ for all $k = 1, \dots, p$ and

$$\hat{\lambda} \sum_{k=1}^p (f_k(x^0) - f_k(x^M)) > \hat{\lambda} M = \frac{M'}{\hat{\lambda}} \cdot \hat{\lambda} = M'. \tag{4.14}$$

This implies that

$$M' < \sum_{k=1}^p \hat{\lambda} (f_k(x^0) - f_k(x^M)) \leq \sum_{k=1}^p \lambda_k (f_k(x^0) - f_k(x^M)) \tag{4.15}$$

is true for all $M' \geq 0$ because of the definition of $\hat{\lambda}$ and because M' was chosen arbitrarily. We can therefore use $M' = \sum_{k=1}^p \lambda_k (f_k(x^0) - f_k(\hat{x}))$ to get

$$\sum_{k=1}^p \lambda_k (f_k(x^0) - f_k(\hat{x})) < \sum_{k=1}^p \lambda_k (f_k(x^0) - f_k(x^M)), \tag{4.16}$$

i.e. $\sum_{k=1}^p \lambda_k f_k(x^M) < \sum_{k=1}^p \lambda_k f_k(\hat{x})$, contradicting optimality of $\hat{x}$ for the weighted sum problem. □

Recalling that $\mathcal{Y}_N \subset \operatorname{cl} \mathcal{Y}_{pN}$ if in addition to convexity $\mathcal{Y} = f(\mathcal{X})$ is $\mathbb{R}^p_{\geqq}$-closed (Theorem 3.17) we can strengthen Theorem 4.16 to emptiness of $\mathcal{X}_E$.

Corollary 4.17. *Assume $\mathcal{X} \subset \mathbb{R}^n$ is convex, $f_k : \mathbb{R}^n \to \mathbb{R}$ are convex for $k = 1, \ldots, p$ and $f(\mathcal{X})$ is $\mathbb{R}^p_{\geqq}$-closed. If (4.10) is unbounded then $\mathcal{X}_E = \emptyset$.*

Proof. From Theorem 3.17 we know that $\mathcal{Y}_N \subset \operatorname{cl} S(\mathcal{Y}) = \operatorname{cl} \mathcal{Y}_{pN}$. From Theorem 4.16 $\mathcal{Y}_{pN} = \emptyset$ whence $\operatorname{cl} \mathcal{Y}_{pE} = \emptyset$ and $\mathcal{Y}_N = \emptyset$. Thus $\mathcal{X}_E = \emptyset$. □

Example 4.18 (Wiecek (1995)). Consider the multicriteria optimization problem with a single variable

$$\begin{aligned} \min \left(x^2 - 4, (x-1)^4\right) \\ \text{subject to } -x - 100 \leq 0. \end{aligned}$$

Benson's problem (4.10) in this case is

$$\begin{aligned} \max\ l_1 + l_2 \\ \text{subject to } -x - 100 &\leq 0 \\ (x^0)^2 - 4 - l_1 - x^2 + 4 &= 0 \\ (x^0 - 1)^4 - l_2 - (x-1)^4 &= 0 \\ l &\geqq 0. \end{aligned}$$

We solve the problem for two choices of x^0. First, consider $x^0 = 0$. We obtain

$$\max l_1 + l_2 \tag{4.17}$$
$$\text{subject to } -x - 100 \leq 0 \tag{4.18}$$
$$x^2 + l_1 = 0 \tag{4.19}$$
$$1 - l_2 - (x-1)^4 = 0 \tag{4.20}$$
$$l_1, l_2 \geqq 0 \tag{4.21}$$

From (4.19) and (4.21) $l_1 = 0$ and $x = 0$. Then (4.20) and (4.21) imply $l_2 = 0$. Therefore $x = 0, l = (0, 0)$ is the only feasible solution of (4.10) with $x^0 = 0$ and Theorem 4.14 implies that $x^0 = 0 \in \mathcal{X}_E$.

The (strictly, weakly) efficient sets for this problem here are all equal to $[0, 1]$ (use the result in Exercise 2.8 to verify this). Therefore let us try (4.10) with an initial solution $x^0 = 2$, to see if $x^0 \notin \mathcal{X}_E$ can be confirmed, and to find a dominating efficient solution.

The problem becomes

$$\begin{aligned}
\max l_1 + l_2 & \\
\text{subject to } -x - 100 &\leq 0 \\
-x^2 + 4 - l_1 &= 0 \\
1 - (x-1)^4 - l_2 &= 0 \\
l_1, l_2 &\geqq 0.
\end{aligned}$$

From the constraints we deduce $0 \leq l_1 \leq 4$ and $0 \leq l_2 \leq 1$. Therefore the optimal objective value is bounded, and according to Proposition 4.15 an optimal solution of (4.10) with $x^0 = 2$ is efficient. Because $x = 0, l_1 = 4, l_2 = 0$ is feasible for (4.10), the optimal objective value is nonzero. Theorem 4.14 implies that $x^0 = 2$ is not efficient. The (unique) optimal solution of the problem is $\hat{x} \approx 0.410$.

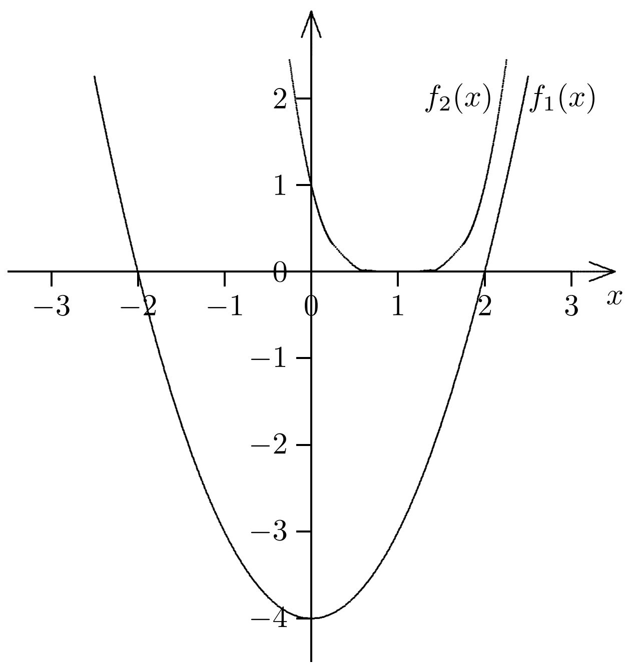

Fig. 4.5. Objective functions in Example 4.18.

$\square$

4.5 Compromise Solutions – Approximation of the Ideal Point

The best possible outcome of a multicriteria problem would be the ideal point y^I (see Definition 2.22). Yet when the objectives are conflicting the ideal values are impossible to obtain. However, the ideal point can serve as a reference point, with the goal to seek for solutions as close as possible to the ideal point. This is the basic idea of compromise programming.

Given a distance measure

$$d : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}_{\geqq}, \tag{4.22}$$

the compromise programming problem is given by

$$\min_{x \in \mathcal{X}} d(f(x), y^I). \tag{4.23}$$

In this text, we will only consider metrics derived from norms as distance measures, i.e. $d(y^1, y^2) = \|y^1 - y^2\|$. In particular for $y^1, y^2, y^3 \in \mathcal{Y}$: d is symmetric $d(y^1, y^2) = d(y^2, y^1)$, satisfies the triangle inequality $d(y^1, y^2) \leq d(y^1, y^3) + d(y^3, y^2)$, and $d(y^1, y^2) = 0$ if and only if $y^1 = y^2$.

The compromise programming problem (4.23) has a nice interpretation in terms of the level sets $\{y \in \mathbb{R}^p : \|y - y^I\| \leq c\}$. These sets contain all points of distance c or less to the ideal point y^I. Therefore the goal of the compromise programming problem is to find the smallest value c such that the intersection of the corresponding level set with $\mathcal{Y} = f(\mathcal{X})$ is nonempty. Figure 4.6 illustrates this perspective for the l_1 distance $\|y^1 - y^2\|_1 := \sum_{k=1}^p |y^1_k - y^2_k|$, the l_∞ distance $\|y^1 - y^2\|_\infty := \max_{k=1}^p |y^1_k - y^2_k|$, and a distance measure d derived from a norm γ with asymmetric level sets.

Whether an optimal solution of problem (4.23) is efficient depends on properties of the distance measure d, and therefore on properties of norm $\|\cdot\|$, from which d is derived.

Definition 4.19. *1. A norm $\|\cdot\| : \mathbb{R}^p \to \mathbb{R}_{\geqq}$ is called* monotone*, if $\|y^1\| \leq \|y^2\|$ holds for all $y^1, y^2 \in \mathbb{R}^p$ with $|y^1_k| \leq |y^2_k|$, $k = 1, \ldots, p$ and moreover $\|y^1\| < \|y^2\|$ if $|y^1_k| < |y^2_k|$, $k = 1, \ldots, p$.*
2. A norm $\|\cdot\|$ is called strictly monotone*, if $\|y^1\| < \|y^2\|$ holds whenever $|y^1_k| \leq |y^2_k|$, $k = 1, \ldots, p$ and $|y^1_j| \neq |y^2_j|$ for some j.*

With definition 4.19 we can prove the following basic results.

Theorem 4.20. *1. If $\|\cdot\|$ is monotone and $\hat{x}$ is an optimal solution of* (4.23) *then $\hat{x}$ is weakly efficient. If $\hat{x}$ is a unique optimal solution of* (4.23) *then $\hat{x}$ is efficient.*

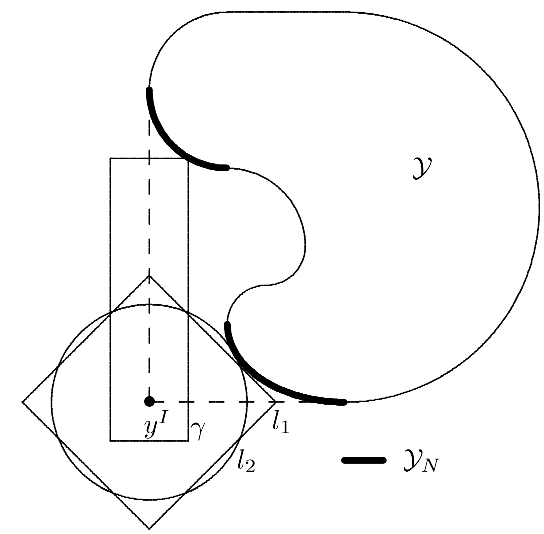

Fig. 4.6. Level Sets $\{y : \|y - y^I\| \leq c\}$ for different norms.

2. If $\|\cdot\|$ is strictly monotone and $\hat{x}$ is an optimal solution of (4.23) *then $\hat{x}$ is efficient.*

Proof. 1. Suppose $\hat{x}$ is an optimal solution of (4.23) and $\hat{x} \notin \mathcal{X}_{wE}$. Then there is some $x' \in \mathcal{X}$ such that $f(x') < f(\hat{x})$. Therefore $0 \leq f_k(x') - y_k^I < f_k(\hat{x}) - y_k^0$ for $k = 1, \ldots, p$ and

$$\|f(x') - y^I\| < \|f(\hat{x}) - y^0\|, \tag{4.24}$$

a contradiction.
Now assume that $\hat{x}$ is a unique optimal solution of (4.23) and that $\hat{x} \notin \mathcal{X}_E$. Then there is some $x' \in \mathcal{X}$ such that $f(x') \leq f(\hat{x})$. Therefore $0 \leq f_k(x) - y_k^I \leq f_k(\hat{x}) - y_k^I$ for $k = 1, \ldots, p$ with one strict inequality, and

$$\|f(x) - y^I\| \leq \|f(\hat{x}) - y^I\|. \tag{4.25}$$

From optimality of $\hat{x}$ equality must hold, which contradicts the uniqueness of $\hat{x}$.

2. Suppose $\hat{x}$ is an optimal solution of (4.23) and $\hat{x} \notin \mathcal{X}_E$. Then there are $x' \in \mathcal{X}$ and $j \in \{1, \ldots, p\}$ such that $f_k(x') \leq f_k(\hat{x})$ for $k = 1, \ldots, p$ and $f_j(x') < f_j(\hat{x})$. Therefore $0 \leq f_k(x) - y_k^I \leq f_k(\hat{x}) - y_k^I$ for all $k = 1, \ldots, p$ and $0 \leq f_j(x) - y_j^I < f_j(\hat{x}) - y_j^I$. Again the contradiction

$$\|f(x) - y^0\| < \|f(\hat{x}) - y^0\| \tag{4.26}$$

follows. □

The most important class of norms is the class of l_{p}-norms $\|\cdot\| = \|\cdot\|_{\mathsf{p}}$, i.e.

$$\|y\|_{\mathsf{p}} = \left(\sum_{k=1}^{p} |y_k|^{\mathsf{p}}\right)^{\frac{1}{\mathsf{p}}} \tag{4.27}$$

for $1 \leq \mathsf{p} \leq \infty$. The l_{p} norm $\|\ \|_{\mathsf{p}}$ is strictly monotone for $1 \leq \mathsf{p} < \infty$ and monotone for $\mathsf{p} = \infty$. The special cases $\mathsf{p} = 1$ with $\|y\| = \sum_{k=1}^{p} |y_k|$ and $\mathsf{p} = \infty$ with $\|y\| = \max_{k=1}^{p} |y_k|$ are of major importance.

As long as we just minimize the distance between a feasible point in objective space and the ideal point, we will find one (weakly) efficient solution for each choice of a norm. The results can be strengthened if we allow weights in the norms. From now on we only consider l_{p}-norms. The weighted compromise programming problems are

$$\min_{x \in \mathcal{X}} \left(\sum_{k=1}^{p} \lambda_k (f_k(x) - y_k^I)^{\mathsf{p}}\right)^{\frac{1}{\mathsf{p}}} \tag{4.28}$$

for general p, and

$$\min_{x \in \mathcal{X}} \max_{k=1,\ldots,p} \lambda_k (f_k(x) - y_k^I), \tag{4.29}$$

for $\mathsf{p} = \infty$.

Here we assume, as usual, that the vector of weights $\lambda \in \mathbb{R}^p_{\geq}$ is nonnegative and nonzero. Note that the functions $\|\cdot\|_{\mathsf{p}}^{\lambda} : \mathbb{R}^p \to \mathbb{R}_{\geqq}$ are not necessarily norms if some of the weights λ_k are zero. It is also of interest to observe that for $\mathsf{p} = 1$ (4.28) can be written as

$$\min_{x \in \mathcal{X}} \sum_{k=1}^{p} (\lambda_k f_k(x) - y_k^I) = \min_{x \in \mathcal{X}} \left(\sum_{k=1}^{p} \lambda_k f_k(x)\right) - \sum_{k=1}^{p} \lambda_k y_k^I.$$

Hence weighted sum scalarization can be seen as a special case of weighted compromise programming. We can therefore exclude this case from now on. The emphasis on the distinction between $1 < \mathsf{p} < \infty$ and $\mathsf{p} = \infty$ is justified for two reasons: The latter is the most interesting case, and the most widely used, and the results are often different from those for $\mathsf{p} < \infty$.

For (4.28) and (4.29) we can prove some basic statements analogous to Theorem 4.20.

Theorem 4.21. *An optimal solution $\hat{x}$ of* (4.28) *with $\mathsf{p} < \infty$ is efficient if one of the following conditions holds.*

1. $\hat{x}$ is a unique optimal solution of (4.28).

2. $\lambda_k > 0$ *for all* $k = 1, \dots, p$.

Proof. Assume $\hat{x}$ is a minimizer of (4.28) but $\hat{x} \notin \mathcal{X}_E$. Then there is some $x' \in \mathcal{X}$ dominating $\hat{x}$.

1. In this case, x' must also be an optimal solution of (4.28), which due to $x \neq \hat{x}$ is impossible.
2. From $\lambda > 0$ we have $0 < \lambda_k(f_k(x) - y^I_k) \leq \lambda_k(f_k(\hat{x}) - y^I_k)$ for all $k = 1, \dots, p$ with strict inequality for some k. Taking power p and summing up preserves strict inequality, which contradicts $\hat{x}$ being an optimal solution of (4.28). □

Proposition 4.22. *Let* $\lambda > 0$ *be a strictly positive weight vector. Then the following statements hold.*

1. *If* $\hat{x}$ *is an optimal solution of* (4.29) *then* $\hat{x} \in \mathcal{X}_{wE}$.
2. *If* $\mathcal{Y}_N$ *is externally stable (see Definition 2.20) and* (4.29) *has an optimal solution then at least one of its optimal solutions is efficient.*
3. *If* (4.29) *has a unique optimal solution* $\hat{x}$, *then* $\hat{x} \in \mathcal{X}_E$.

Proof. 1. The proof is standard and left out. See the proofs of Theorems 4.20 and 4.21.
2. Assume that (4.29) has optimal solutions, but none of them is is efficient. Let $\hat{x}$ be an optimal solution of (4.29). Because $\mathcal{Y}_N$ is externally stable there must be an $x \in \mathcal{X}_E$ with $f(x) \leq f(\hat{x})$. Then $\lambda_k(f_k(x) - y^I_k) \leq \lambda_k(f_k(\hat{x}) - y^I_k)$ for $k = 1, \dots, p$, which means x is optimal for (4.29), too.
3. This part can be shown as usual. If $\mathcal{Y}_N$ is externally stable it follows directly from the second statement. □

Actually, all the results we proved so far remain valid, if the ideal point y^I is replaced by any other reference point y^R, as long as this reference point is chosen to satisfy $y^R \leqq y^I$.

Definition 4.23. *A point* $y^U := y^I_i - \varepsilon$, *where* $\varepsilon \in \mathbb{R}^p_>$ *has small positive components is called a* utopia point.

Note that not even minimizing the single objectives independently of one another will yield the utopia values: $f_k(x) > y^U_k$ for all feasible solutions $x \in \mathcal{X}$ and all $k = 1, \dots, p$. The advantage of using utopia points instead of ideal points will become clear from the following theorems. The first complements Proposition 4.22 by a necessary and sufficient condition for weak efficiency.

Theorem 4.24 (Choo and Atkins (1983)). *A feasible solution* $\hat{x} \in \mathcal{X}$ *is weakly efficient if and only if there is a weight vector* $\lambda > 0$ *such that* $\hat{x}$ *is an optimal solution of the problem*

$$\min_{x\in\mathcal{X}} \max_{k=1,\dots,p} \lambda_k(f_k(x) - y_k^U). \tag{4.30}$$

Proof. "$\Longleftarrow$" The proof of sufficiency is the same standard proof as that of the first part of Proposition 4.22.

"$\Longrightarrow$" We define appropriate weights and show that they do the job. Let $\lambda_k := 1/(f_k(\hat{x}) - y_k^U)$. These weights are positive and finite. Suppose $\hat{x}$ is not optimal for (4.30). Then there is a feasible $x \in \mathcal{X}$ such that

$$\max_{k=1,\dots,p} \lambda_k(f_k(x) - y_k^U) < \max_{k=1,\dots,p} \frac{1}{f_k(\hat{x}) - y_k^U}(f_k(\hat{x}) - y_k^U) = 1$$

and therefore

$$\lambda_k(f_k(x) - y_k^U) < 1 \text{ for all } k = 1,\dots,p.$$

Dividing by λ_k we get $f_k(x) - y_k^U < f_k(\hat{x}) - y_k^U$ for all $k = 1,\dots,p$ and thus $f(x) < f(\hat{x})$, contradicting $\hat{x} \in \mathcal{X}_{wE}$. □

With Theorem 4.24 we have a complete characterization of weakly efficient solutions for general, nonconvex problems. However, as for the ε-constraint method, we have to accept the drawback that in practice the result will only be useful as a check for weak efficiency, because $f(\hat{x})$ is needed to define the weights to prove optimality of $\hat{x}$. It should also be noted that if y^U is replaced by y^I in Theorem 4.24 then not even

$$\mathcal{Y}_{pN} \subset \bigcup_{\lambda\in\mathbb{R}^p_>} \left\{\hat{y} : \max_{k=1,\dots,p} \lambda_k|\hat{y}_k - y_k^I| \leq \max_{k=1,\dots,p} \lambda_k|y_k - y_k^I| \text{ for all } y \in \mathcal{Y}\right\}$$

is true, see Exercise 4.8.

We are now able to prove the main result of this section. It is the formal extension of the main result on the weighted sum scalarization in Chapter 3. We have noted earlier that (4.28) contains the weighted sum problem as a special case (setting $\mathsf{p} = 1$). For this special case we have seen in Theorem 3.17 that for $\mathbb{R}^p_{\geqq}$-convex and $\mathbb{R}^p_{\geqq}$-bounded $\mathcal{Y}$

$$\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_{pN} \subset \mathcal{Y}_N \subset \mathrm{cl}(S(\mathcal{Y})).$$

For the general problem (4.28) we can therefore expect more general results, when convexity is relaxed. Theorem 4.25 is this generalization. Before we can prove the theorem, we have to introduce some notation to enhance readability of the proof.

Let

$$\Lambda := \left\{ \lambda \in \mathbb{R}^p_{\geqq} : \sum_{k=1}^{p} \lambda_k = 1 \right\}$$
$$\Lambda^0 := \operatorname{ri} \Lambda = \left\{ \lambda \in \mathbb{R}^p_{\geq} : \sum_{k=1}^{p} \lambda_k = 1 \right\}.$$

For $\lambda \in \Lambda$ and $y \in \mathcal{Y}$ we shall write

$$\lambda \odot y = (\lambda_1 y_1, \ldots, \lambda_p y_p).$$

Furthermore, in analogy to $\mathrm{Opt}(\lambda, \mathcal{Y})$ and $S(\mathcal{Y})$, the set of best approximations of y^I for a certain weight λ and norm $\| \, \|_{\mathsf{p}}$ is denoted by

$$\mathcal{A}(\lambda, \mathsf{p}, \mathcal{Y}) := \left\{ \hat{y} \in \mathcal{Y} : \left\| \lambda \odot (\hat{y} - y^U) \right\|_{\mathsf{p}} = \min_{y \in \mathcal{Y}} \left\| w \odot (y - y^U) \right\|_{\mathsf{p}} \right\} \tag{4.31}$$
$$\mathcal{A}(\mathcal{Y}) := \bigcup_{\lambda \in \Lambda^0} \bigcup_{1 \leq \mathsf{p} < \infty} \mathcal{A}(\lambda, \mathsf{p}, \mathcal{Y}). \tag{4.32}$$

From Theorem 4.21 and Theorem 4.24 we already know that

$$\mathcal{A}(\mathcal{Y}) \subset \mathcal{Y}_N \subset \mathcal{Y}_{wN} = \bigcup_{\lambda \in \Lambda^0} \mathcal{A}(\lambda, \infty, \mathcal{Y}). \tag{4.33}$$

The main result will show that this can be strengthened to

$$\mathcal{A}(\mathcal{Y}) \subset \mathcal{Y}_{pE} \subset \mathcal{Y}_E \subset \operatorname{cl}(\mathcal{A}(\mathcal{Y})) \tag{4.34}$$

for $\mathbb{R}^p_{\geqq}$-closed sets $\mathcal{Y}$, a complete analogy to Theorem 3.17 for nonconvex sets.

In the proof of Theorem 4.25 some of the essential arguments are based on the following properties of l_{p}-norms:

(P1) $\|y\|_\infty \leq \|y\|_{\mathsf{p}}$ for all $1 \leq \mathsf{p} < \infty$ and all $y \in \mathbb{R}^p$,
(P2) $\|y\|_{\mathsf{p}} \to \|y\|_\infty$ as $\mathsf{p} \to \infty$ holds for any $y \in \mathbb{R}^p$,
(P3) $\| \cdot \|_{\mathsf{p}}$ is strictly monotone for all $1 \leq \mathsf{p} < \infty$.

Theorem 4.25 (Sawaragi *et al.* (1985)). *If $\mathcal{Y}$ is $\mathbb{R}^p_{\geqq}$-closed then*

$$\mathcal{A}(\mathcal{Y}) \subset \mathcal{Y}_{pN} \subset \mathcal{Y}_N \subset \operatorname{cl}(\mathcal{A}(\mathcal{Y})).$$

Proof. The proof is divided into two main parts, corresponding to the two inclusions $\mathcal{A}(\mathcal{Y}) \subset \mathcal{Y}_{pN}$ and $\mathcal{Y}_N \subset \operatorname{cl}(\mathcal{A}(\mathcal{Y}))$.

Part 1: $\mathcal{A}(\mathcal{Y}) \subset \mathcal{Y}_{pN}$. Let $\hat{y} \in \mathcal{A}(\mathcal{Y})$. By definition of $\mathcal{A}(\mathcal{Y})$ there is a positive weight vector $\lambda \in \Lambda^0$ and some $\mathsf{p} \in [1, \infty)$ such that

$$\left\| \lambda \odot (\hat{y} - y^U) \right\|_{\mathsf{p}} \leq \left\| \lambda \odot (y - y^U) \right\|_{\mathsf{p}} \tag{4.35}$$

for all $y \in \mathcal{Y}$. Let us assume that $\hat{y} \notin \mathcal{Y}_{pN}$, which according to Benson's definition 2.44 means that there are sequences $\{\beta_k\} \subset \mathbb{R}$, $\{y^k\} \subset \mathcal{Y}$, and $\{d^k\} \subset \mathbb{R}^p_{\geqq}$ with $\beta_k > 0$ and

$$\beta_k(y^k + d^k - \hat{y}) \to -d \text{ for some } d \in \mathbb{R}^p_{\geq}. \tag{4.36}$$

We distinguish the two cases $\{\beta_k\}$ bounded and $\{\beta_k\}$ unbounded and use (4.36) to construct a point $\tilde{y}$, respectively a sequence y^k, which do not satisfy (4.35).

$\{\beta_k\}$ bounded: In this case we can assume, without less of generality, that β_k converges to some number $\beta_0 \geq 0$ (taking a subsequence, if necessary). If $\beta_0 = 0$ the fact $y^k + d^k - \hat{y} \geq y^U - \hat{y}$ implies

$$\beta_k(y^k + d^k - \hat{y}) \geq \beta_k(y^U - \hat{y}). \tag{4.37}$$

Because the left hand side term in (4.37) converges to $-d$, and the right hand side term to 0, we get $-d \geq 0$, a contradiction.
If, on the other hand, $\beta_0 > 0$ we have that $y^k + d^k - \hat{y} \to (-d)/\beta_0$, which is nonzero, and $y^k + d^k \to \hat{y} - d/\beta_0$. Since $y^k + d^k \in \mathcal{Y} + \mathbb{R}^p_{\geqq}$ and this set is closed, it must be that the limit $\hat{y} - d/\beta_0 \in \mathcal{Y} + \mathbb{R}^p_{\geqq}$. From this observation we conclude that there is some $\tilde{y} \in \mathcal{Y}$ such that $\hat{y} \geq \tilde{y}$. Positive weights and strict monotonicity of the l_{p}-norm finally yield $\|\lambda \odot (\hat{y} - y^U)\|_{\mathsf{p}} > \|\lambda \odot (\tilde{y} - y^U)\|_{\mathsf{p}}$.

$\{\beta_k\}$ unbounded: Taking subsequences if necessary, we can here assume $\beta_k \to \infty$, which by the convergence in (4.36) gives $y^k + d^k - \hat{y} \to 0$. Because $\hat{y}_k > y^U_k$ for all $k = 1, \ldots, p$ we can find a sufficiently large $\beta' > 0$ so that

$$0 \leq \hat{y} - \frac{d}{\beta} - y^U < \hat{y} - y^U \tag{4.38}$$

for all $\beta > \beta'$. We use strict monotonicity of the norm and $\lambda > 0$ to obtain

$$\left\| \lambda \odot \left(\hat{y} - \frac{d}{\beta} - y^U \right) \right\|_{\mathsf{p}} < \left\| \lambda \odot (\hat{y} - y^U) \right\|_{\mathsf{p}} \tag{4.39}$$

for all $\beta > \beta'$. Since $\beta_k \to \infty$ we will have $\beta_k > \beta'$ for all $k \geq k_0$ with a sufficiently large k_0. Therefore

$$\begin{aligned} \left\| \lambda \odot (y^k + d^k - y^U) \right\|_{\mathsf{p}} &= \left\| \lambda \odot (y^k + d^k - \hat{y} + \tfrac{d}{\beta_k} + \hat{y} - \tfrac{d}{\beta_k} - y^U) \right\|_{\mathsf{p}} \\ &\leq \left\| \lambda \odot (y^k + d^k - \hat{y}) \right\|_{\mathsf{p}} + \tfrac{\|\lambda \odot d\|_{\mathsf{p}}}{\beta_k} + \\ &\quad \left\| \lambda \odot \left(\hat{y} - \tfrac{d}{\beta_k} - y^U \right) \right\|_{\mathsf{p}}. \end{aligned} \tag{4.40}$$

We know that the first term on the right hand side of the inequality of (4.40) converges to 0. The sequence β_k being unbounded implies the second term converges to 0, too. Thus from (4.40) and (4.39)

$$\begin{aligned}\lim_{k\to\infty}\left\|\lambda\odot(y^k+d^k-y^U)\right\|_{\mathsf{p}} &\leq \lim_{k\to\infty}\left\|\lambda\odot\left(\hat{y}-\tfrac{d}{\beta_k}-y^U\right)\right\|_{\mathsf{p}} \\ &< \left\|\lambda\odot(\hat{y}-y^U)\right\|_{\mathsf{p}}.\end{aligned} \tag{4.41}$$

But since $y^k+d^k-y^U \geqq y^k-y^U \geq 0$, applying monotonicity of the norm once more, (4.41) implies $\lim\limits_{k\to\infty}\|\lambda\odot(y^k-y^U)\|_{\mathsf{p}} < \|\lambda\odot(\hat{y}-y^U)\|_{\mathsf{p}}$.

Part 2: $\mathcal{Y}_N \subset \mathrm{cl}(\mathcal{A}(\mathcal{Y}))$. We prove this part by showing that for all $\hat{y} \in \mathcal{Y}_N$ and for all $\varepsilon > 0$ there is some $y^\varepsilon \in \mathcal{A}(\mathcal{Y})$ in an ε-neighbourhood of $\hat{y}$. Then, taking the closure of $\mathcal{A}(\mathcal{Y})$, the result follows. The ε-neighbourhood is defined according to the l_∞-norm.

I.e. let $\hat{y} \in \mathcal{Y}_N$ and let $\varepsilon > 0$. We show that there is some $y^\varepsilon \in \mathcal{A}(\mathcal{Y})$ with $\|y^\varepsilon - \hat{y}\|_\infty = \max_{k=1,\dots,p}|y^\varepsilon_k - \hat{y}_k| < \varepsilon$.

First we proof an auxiliary claim: For each $\varepsilon > 0$ there is $y' > \hat{y}$ such that $\|y-\hat{y}\|_\infty < \varepsilon$ for all y in the section $(y' - \mathbb{R}^p_{\geqq}) \cap \mathcal{Y}$, see Figure 4.7. To see this, assume that for some $\varepsilon > 0$ there is no such y'. Then there must be a sequence $\{\hat{y}^k\} \subset \mathbb{R}^p$ with $\hat{y}^k \geq \hat{y}$, $\hat{y}^k \to \hat{y}$ such that for all k there is $y^k \in (\hat{y}^k - \mathbb{R}^p_{\geqq}) \cap \mathcal{Y}$ with $\|y^k - \hat{y}\| \geq \varepsilon$.

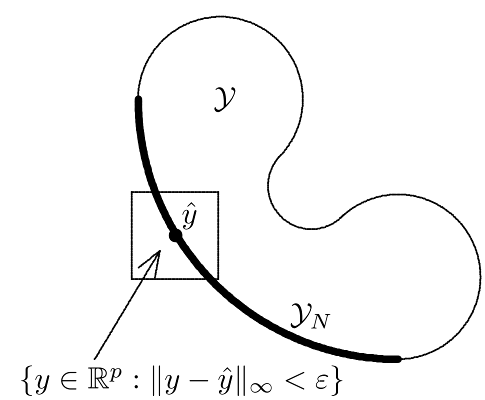

Fig. 4.7. ε-neighbourhoods of nondominated points in the l_∞-norm.

Because $\mathcal{Y}+\mathbb{R}^p_{\geqq}$ is closed and $\mathcal{Y} \subset y^U + \mathbb{R}_{\geqq}$, i.e. $\mathcal{Y}$ is bounded below we can assume without loss of generality that $y^k \to y'' + d''$, where $y'' \in \mathcal{Y}, d'' \geqq 0$ and $\|y''+d''-\hat{y}\|_\infty \geq \varepsilon$. On the other hand $y''+d'' \in (\hat{y}-\mathbb{R}^p_{\geqq}) \cap (\mathcal{Y}+\mathbb{R}^p_{\geqq}) = \{\hat{y}\}$ (since $\hat{y} \in \mathcal{Y}_N$), a contradiction.

For y' from the claim we know $y^U \leq \hat{y} < y'$ and thus there is some $\lambda \in \Lambda^0$, and $\beta > 0$ such that $y' - y^U = \beta(1/\lambda_1, \ldots, 1/\lambda_p)$. Hence

$$\lambda_k(\hat{y}_k - y_k^U) < \lambda_k(y_k' - y_k^U) = \beta \tag{4.42}$$

for all $k = 1, \ldots, p$ and

$$\|\lambda \odot (\hat{y} - y^U)\|_\infty < \beta. \tag{4.43}$$

Choose $y(\mathsf{p}) \in \mathcal{A}(\lambda, \mathsf{p}, \mathcal{Y})$. Note that $\mathcal{A}(\lambda, \mathsf{p}, \mathcal{Y})$ is nonempty because $\mathcal{Y} + \mathbb{R}^p_{\geqq}$ is closed. We obtain

$$\begin{aligned} \|\lambda \odot (y(\mathsf{p}) - y^U)\|_\infty &\leq \|\lambda \odot (y(\mathsf{p}) - y^U)\|_{\mathsf{p}} \\ &\leq \|\lambda \odot (\hat{y} - y^U)\|_{\mathsf{p}} \\ &\to \|\lambda \odot (\hat{y} - y^U)\|_\infty \ < \ \beta, \end{aligned} \tag{4.44}$$

where we have used (P1), the definition of $\mathcal{A}(\lambda, \mathsf{p}, \mathcal{Y})$, and (P2), respectively.

This means we have $\|\lambda \odot (y(\mathsf{p}) - y^U)\|_\infty \leq \beta$, if p is sufficiently large. By the definition of the l_∞-norm

$$y_k(\mathsf{p}) - y_k^U \leq \frac{\beta}{\lambda_k} = y_k' - y_k^U \text{ for all } k = 1, \ldots, p, \tag{4.45}$$

i.e. $y(\mathsf{p}) \leq y'$ or $y(\mathsf{p}) \in (y' - \mathbb{R}^p_{\geqq}) \cap \mathcal{Y}$ and therefore, using the auxiliary claim, we can choose $y^\varepsilon := y(\mathsf{p})$ for sufficiently large p. □

We know that if $\mathcal{Y} + \mathbb{R}^p_{\geqq}$ is convex, $\mathsf{p} = 1$ will always work for $y(\mathsf{p}) \in \mathcal{A}(\lambda, \mathsf{p}, \mathcal{Y})$ and that $\mathsf{p} = \infty$ can be chosen for arbitrary sets. The proof of the second part of the theorem suggests that, if $\mathcal{Y}$ is not $\mathbb{R}_{\geqq}$-convex, p has to be bigger than one. The value of p seems to be related to the degree of nonconvexity of $\mathcal{Y}$. An Example, where $1 < \mathsf{p} < \infty$ can be chosen to generate $\mathcal{Y}_N$ by solving (4.28) is given in Exercise 4.7.

At the end of this section we have two examples. The first one shows that the inclusion $\operatorname{cl} \mathcal{A}(\mathcal{Y}) \subset \mathcal{Y}_N$ may not be true. In the second we solve the problem from Example 4.18 by the compromise programming method.

Example 4.26. Let $\mathcal{Y} := \{y \in \mathbb{R}^2 : y_1^2 + (y_2 - 1)^2 \leq 1\} \cup \{y \in \mathbb{R}^2 : y_1 \geq 0,\ y_2 \geq -1\}$. Here the efficient set is $\mathcal{Y}_N = \{y \in \mathcal{Y} : y_1^2 + (y_2 - 1)^2 = 1, y_2 \leq 1; y_1 > -1\} \cup \{(0, -1)\}$, see Figure 4.8.

Therefore $0 \notin \mathcal{Y}_N$ but $0 \in \operatorname{cl} \mathcal{A}(\mathcal{Y})$. Note that the efficient points with $y_2 < 1$ and $y_1 < 0$ are all generated as optimal solutions of (4.28) with any choice of $y^U < (-1, -1)$ for appropriate λ and p. □

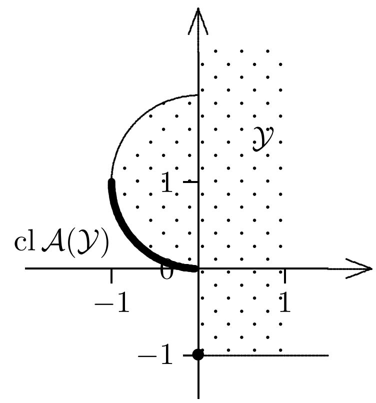

Fig. 4.8. cl $\mathcal{A}(\mathcal{Y})$ may contain dominated points.

Example 4.27. We apply the compromise programming method to the problem of Example 4.18:

$$\begin{aligned} &\min(x^2 - 4, (x-1)^4) \\ &\text{subject to } -x - 100 \leq 0. \end{aligned}$$

Let $\lambda = (0.5, 0.5)$ and $\mathsf{p} = 2$. The ideal point is $y^I = (-4, 0)$ and we choose $y^U = (-5, -1)$. So (4.28) with $\mathsf{p} = 2$ and y^U as reference point is

$$\begin{aligned} \min \sqrt{\frac{1}{2}(x^2 - 4 + 5)^2 + \frac{1}{2}((x-1)^4 + 1)^2} \\ \text{subject to } -x - 100 \leq 0. \end{aligned} \tag{4.46}$$

Observing that the compromise programming objective is convex, that the problem is in fact unconstrained, and that the derivative of the objective function in (4.46) is zero if and only if the derivative of the term under the root is zero we set

$$\phi(x) = \frac{1}{2}(x^2 + 1)^2 + \frac{1}{2}((x-1)^4 + 1)^2$$

and compute

$$\begin{aligned} \phi'(x) &= (x^2 + 1)2x + ((x-1)^4 + 1) \cdot 4(x-1)^3 \\ &= 2x^3 + 2x + 4(x-1)^7 + 4(x-1)^3 \end{aligned}$$

From $\phi'(x) = 0$ we obtain $\hat{x} \approx 0.40563$ as unique minimizer. Theorem 4.21 confirms that $\hat{x} \in \mathcal{X}_E$. □

4.6 The Achievement Function Method

A certain class of real-valued functions $s_r : \mathbb{R}^p \to \mathbb{R}$, referred to as achievement functions, can be used to scalarize the MOP (4.1). The scalarized problem is given by

$$\begin{aligned} &\min s_R(f(x)) \\ &\text{subject to } x \in \mathcal{X}. \end{aligned} \tag{4.47}$$

Similar to distance functions discussed in Section 4.5 above, certain properties of achievement functions guarantee that problem (4.47) yields (weakly) efficient solutions.

Definition 4.28. *An achievement function $s_R : \mathbb{R}^p \to \mathbb{R}$ is said to be*

1. *increasing if for $y^1, y^2 \in \mathbb{R}^p$, $y^1 \leqq y^2$ then $s_R(y^1) \leq s_R(y^2)$,*
2. *strictly increasing if for $y^1, y^2 \in \mathbb{R}^p$, $y^1 < y^2$ then $s_R(y^1) < s_R(y^2)$,*
3. *strongly increasing if for $y^1, y^2 \in \mathbb{R}^p$, $y^1 \leq y^2$ then $s_R(y^1) < s_R(y^2)$.*

Theorem 4.29 (Wierzbicki (1986a,b)).

1. *Let an achievement function s_R be increasing. If $\hat{x} \in \mathcal{X}$ is a unique optimal solution of problem (4.47) then $\hat{x} \in \mathcal{X}_{sE}$.*
2. *Let an achievement function s_R be strictly increasing. If $\hat{x} \in \mathcal{X}$ is an optimal solution of problem (4.47) then $\hat{x} \in \mathcal{X}_{wE}$.*
3. *Let an achievement function s_R be strongly increasing. If $\hat{x} \in \mathcal{X}$ is an optimal solution of problem (4.47) then $\hat{x} \in \mathcal{X}_E$.*

We omit the proof, as it is very similar to the proofs of Theorems 4.20, 4.21 and Proposition 4.22, see Exercise 4.11.

Among many achievement functions satisfying the above properties we mention the strictly increasing function

$$s_R(y) = \max_{k=1,\ldots,p} \{\lambda_k (y_k - y_k^R)\}$$

and the strongly increasing functions

$$\begin{aligned} s_R(y) &= \max_{k=1,\ldots,p} \{\lambda_k (y_k - y_k^R)\} + \rho_1 \sum_{k=1}^{p} \lambda_k (y_k - y_k^R) \\ s_R(y) &= -\|y - y^R\|^2 + \rho_2 \|(y - y^R)_+\|^2, \end{aligned}$$

where $y^R \in \mathbb{R}^p$ is a reference point, $\lambda \in \mathbb{R}^p_>$ is a vector of positive weights, $\rho_1 > 0$ and sufficiently small, $\rho_2 > 1$ is a penalty parameter, and $(y - y^R)_+$ is a vector with components $\max\{0, y_k - r_k\}$ (Wierzbicki, 1986a,b).

4.7 Notes

In Guddat *et al.* (1985), Theorem 4.7 is also generalized for scalarizations in the form of problem (4.7) with an objective function being strictly increasing on $\mathbb{R}^p$ (cf. Definition 4.28). See also Exercise 4.10.

The formulation of the scalarized problem of Benson's method (4.10) has been used by Ecker and Hegner (1978) and Ecker and Kouada (1975) in multiobjective linear programming earlier. In fact, already Charnes and Cooper (1961) have formulated the problem and proved Theorem 4.14.

Some discussion of compromise programming that covers several aspects we neglected here can be found in Yu (1985). Two further remarks on the proof of Theorem 4.25 are in order. First, the statement remains true, if y^I is chosen as reference point. However, the proof needs modification (we have used $y > y^U$ in both parts). We refer to Sawaragi *et al.* (1985) for this extension. Second, we remark that the definition of the l_{p}-norms has never been used. Therefore the theorem is valid for any family of norms with properties (P1) – (P3). This fact has been used by several researchers to justify methods for generation of efficient solutions, e.g. Choo and Atkins (1983). Other norms used for compromise programming are include the augmented l_∞-norm in Steuer and Choo (1983); Steuer (1985) and the modified l_∞-norm by Kaliszewski (1987).

There are many more scalarization methods available in the literature than we can present here. They can roughly be classified as follows.

Weighting methods These include weighted sum method (Chapter 3), the weighted t-th power method White (1988), and the weighted quadratic method Tind and Wiecek (1999)

Constraint methods We have discussed the ε-constraint method (Section 4.1), the hybrid method (Section 4.2), the elastic constraint method (Section 4.3) and Benson's method (Section 4.4). See also Exercises 4.1, 4.2 and 4.10 for more.

Reference point methods The most important in this category are the compromise programming method of Section 4.5 and the (more general) achievement function method (Section 4.6). But goal programming (see e.g. Tamiz and Jones (1996)) and the weighted geometric mean approach of Lootsma *et al.* (1995) also fit in this category.

Direction based methods There is a wide variety of direction based methods, including the reference direction approach Korhonen and Wallenius (1988), the Pascoletti-Serafini method Pascoletti and Serafini (1984) and the gauge-based technique of Klamroth *et al.* (2002).

Of course, some methods can be associated with several of these categories. A survey with the most important results can be found in Ehrgott and Wiecek (2005).

Exercises

4.1. Suppose $\hat{x}$ is the unique optimal solution of

$$\min_{x \in \mathcal{X}} \sum_{k=1}^{p} \lambda_i f_i(x)$$

with $\lambda \in \mathbb{R}^p_{\geq}$. Then there exists some $\hat{\varepsilon} \in \mathbb{R}^p$ such that $\hat{x}$ is an optimal solution of (4.3) for all $j = 1, \ldots, p$.

4.2 (Corley (1980)). Show that $\hat{x} \in \mathcal{X}_E$ if and only if there are $\lambda \in \mathbb{R}^p_{>}$ and $\varepsilon \in \mathbb{R}^p$ such that $\hat{x}$ is an optimal solution of

$$\begin{aligned} &\min_{x \in \mathcal{X}} \sum_{k=1}^{p} \lambda_k f_k(x) \\ &\text{subject to } f(x) \leq \varepsilon. \end{aligned} \tag{4.48}$$

4.3. Show, by choosing the parameters μ and ε in (4.8) appropriately, that both the weighted sum problem (4.2) and the ε-constraint problem (4.3) are special cases of (4.8).

4.4. Consider the following bicriterion optimization problem.

$$\begin{array}{ll} \min & -6x_1 - 4x_2 \\ \min & \qquad\quad -x_1 \\ \text{s.t.} & \quad x_1 + x_2 \leq 100 \\ & \quad 2x_1 + x_2 \leq 150 \\ & \qquad x_1, x_2 \geq 0. \end{array}$$

Use $\varepsilon = 0$ and solve the ε-constraint problem (4.3) with $j = 1$. Check if the optimal solution $\hat{x}$ of $P_1(0)$ is efficient using Benson's test (4.10).

4.5. Consider $\min\limits_{x \in \mathcal{X}} (f_1(x), \ldots, f_p(x))$ and assume $0 < \min_{x \in \mathcal{X}} f_k(x)$ for all $k = 1, \ldots, p$. Prove that $x \in \mathcal{X}_{wE}$ if and only if x is an optimal solution of

$$\min_{x \in \mathcal{X}} \max_{k=1,\ldots,p} \lambda_k f_k(x)$$

for some $\lambda \in \mathbb{R}^p_{>}$.

4.6. Find an efficient solution of the problem of Exercise 4.4 using the compromise programming method. Use $\lambda = (1/2, 1/2)$ and find an optimal solution of (4.28) for $\mathsf{p} = 1, 2, \infty$.

4.7. Consider finding a compromise solution by maximizing the distance to the nadir point.

1. Let $\|\cdot\|$ be a norm. Show that an optimal solution of the problem

$$\begin{aligned} &\max_{x\in\mathcal{X}} \|f(x)-y^N\| \\ &\text{subject to } f_k(x) \leq y_k^n \quad k=1,\ldots,p \end{aligned} \tag{4.49}$$

is weakly efficient. Give a condition under which an optimal solution of (4.49) is efficient.
2. Another possibility is to solve

$$\begin{aligned} &\max_{x\in\mathcal{X}} \min_{k=1,\ldots,p} |f_k(x)-y_k^N| \\ &\text{subject to} f_k(x) \leq y_k^n, \quad k=1,\ldots,p. \end{aligned} \tag{4.50}$$

Prove that an optimal solution of (4.50) is weakly efficient.

4.8. Let $\mathcal{Y} = \{y \in \mathbb{R}^2 : y_1 + y_2 \geq 1,\ 0 \leq y_1 \leq 1\}$. Show that $\hat{y} = (0,1) \in \mathcal{Y}_{pN}$ according to Benson's definition, but that there is no $\lambda \in \Lambda^0$ such that $\hat{y} \in \mathcal{A}(\lambda, \infty, \mathcal{Y})$, if y^I is used as reference point in (4.28).

4.9. Let $\mathcal{Y} = \left\{(y_1, y_2) \in \mathbb{R}^2_{\geqq} : y_1^2 + y_2^2 \geq 1\right\}$. Verify that there is $1 < p < \infty$ such that

$$\mathcal{Y}_N = \bigcup_{\lambda\in\Lambda^0} \mathcal{A}(\lambda, p, \mathcal{Y}).$$

Choose either y^I or y^U in the definition of $\mathcal{A}(\lambda, p, \mathcal{Y})$ and N_p^λ.

4.10 (Soland (1979)). A function $s : \mathbb{R}^p \to \mathbb{R}$ is called strongly increasing, if for $y^1, y^2 \in \mathbb{R}^p$ with $y^1 \leq y^2$ the inequality $s(y^1) < s(y^2)$ holds (see Definition 4.28).

Consider the following single objective optimization problem, where $\varepsilon \in \mathbb{R}^p$ and $f : \mathbb{R}^n \to \mathbb{R}^p$.

$$\begin{aligned} &\min s(f(x)) \\ &\text{subject to } x \in \mathcal{X} \\ &\qquad f(x) \leq \varepsilon. \end{aligned} \tag{4.51}$$

Let s be strongly increasing. Prove that $x \in \mathcal{X}_E$ if and only if there is $\varepsilon \in \mathbb{R}^p$ such that x is an optimal solution of (4.51) with finite objective value.

4.11. Prove Theorem 4.29.

4.12. An achievement function $s_R : \mathbb{R} \to \mathbb{R}$ is called order representing if s_R is strictly increasing (Definition 4.28) for any $y^R \in \mathbb{R}^p$ and in addition

$$\{y \in \mathbb{R}^p : s_R(y) < 0\} = y^R - \mathbb{R}^p_>$$

holds for all $y^R \in \mathbb{R}^p$. Which of the functions

$$\begin{aligned} s_R(y) &= d(y, y^R) = \|y - y^R\|, \\ s_R(y) &= \max_{k=1,\ldots,p} \{\lambda_k(y_k - y_k^R)\}, \\ s_R(y) &= \max_{k=1,\ldots,p} \{\lambda_k(y_k - y_k^R)\} + \rho_1 \sum_{k=1}^{p} \lambda_k(y_k - y_k^R) \\ s_R(y) &= -\|y - y^R\|^2 + \rho_2\|(y - y^R)_+\|^2 \end{aligned}$$

is order representing?

4.13. Show that Benson's problem (4.10), the weighted sum scalarization (4.2) with $\lambda \in \mathbb{R}^p_>$, the compromise programming problem (4.28) with $1 \leq \mathsf{p} < \infty$ and $\lambda \in \mathbb{R}^p_>$, and Corley's problem (4.48) (see Exercise 4.2) can all be seen as special cases of (4.51).