3

The Weighted Sum Method and Related Topics

In this chapter we will investigate to what extent an MOP

$$\min_{x \in \mathcal{X}} (f_1(x), \ldots, f_p(x)) \tag{3.1}$$

of the Pareto class

$$(\mathcal{X}, f, \mathbb{R}^p)/\mathrm{id}/(\mathbb{R}^p, \leq)$$

can be solved (i.e. its efficient solutions be found) by solving single objective problem problems of the type

$$\min_{x \in \mathcal{X}} \sum_{k=1}^{p} \lambda_k f_k(x), \tag{3.2}$$

which in terms of the classification of Section 1.5 is written as

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

where $\langle \lambda, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^p$. We call the single objective (or scalar) optimization problem (3.2) a *weighted sum scalarization* of the MOP (3.1).

As in the previous chapter, we will usually look at the objective space $\mathcal{Y}$ first and prove results on the relationships between (weakly, properly) nondominated points and values $\sum_{k=1}^{p} \lambda_k y_k$. From those, we can derive results on the relationships between $\mathcal{X}_{(w,p)E}$ and optimal solutions of (3.2).

We use these results to prove Fritz-John and Kuhn-Tucker type optimality conditions for (weakly, properly) efficient solutions (Section 3.3). Finally, we investigate conditions that guarantee that nondominated and efficient sets are connected (Section 3.4).

Let $\mathcal{Y} \subset \mathbb{R}^p$. For a fixed $\lambda \in \mathbb{R}^p_{\geqq}$ we denote by

$$\mathcal{S}(\lambda, \mathcal{Y}) := \left\{ \hat{y} \in \mathcal{Y} : \ \langle \lambda, \hat{y} \rangle = \min_{y \in \mathcal{Y}} \langle \lambda, y \rangle \right\} \tag{3.4}$$

the set of optimal points of $\mathcal{Y}$ with respect to λ.

Figure 3.1 gives an example of a set $\mathcal{S}(\lambda, \mathcal{Y})$ consisting of two points y^1 and y^2. These points are the intersection points of a line $\{y \in \mathbb{R}^p : \langle \lambda, y \rangle = \hat{c}$. Obviously, y^1 and y^2 are nondominated. Considering c as a parameter, and the family of lines $\{y \in \mathbb{R}^p : \langle \lambda, y \rangle = c\}$, we see that in Figure 3.1 $\hat{c}$ is chosen as the smallest value of c such that the intersection of the line with $\mathcal{Y}$ is nonempty.

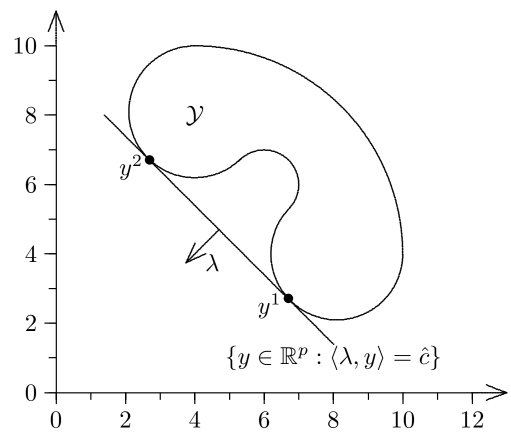

Fig. 3.1. A set $\mathcal{S}(\lambda, \mathcal{Y})$.

Graphically, to find $\hat{c}$ we can start with a large value of the parameter c and translate the line in parallel towards the origin as much as possible while keeping a nonempty intersection with $\mathcal{Y}$. Analytically, this means finding elements of $\mathcal{S}(\lambda, \mathcal{Y})$. The obvious questions are:

1. Does this process always yield nondominated points? (Is $\mathcal{S}(\lambda, \mathcal{Y}) \subset \mathcal{Y}_N$?) and
2. if so, can all nondominated points be detected this way? (Is $\mathcal{Y}_N \subset \cup_{\lambda \in \mathbb{R}^p_{\geq}} \mathcal{S}(\lambda, \mathcal{Y})$?)

Note that due to the definition of nondominated points, we have to consider nonnegative weighting vectors $\lambda \in \mathbb{R}^p_{\geq}$ only. However, the distinction between nonnegative and positive weights turns out to be essential. Therefore we distinguish optimal points of $\mathcal{Y}$ with respect to nonnegative and strictly positive weights, and define

$$\mathcal{S}(\mathcal{Y}) := \bigcup_{\lambda \in \mathbb{R}^p_>} \mathcal{S}(\lambda, \mathcal{Y}) = \bigcup_{\{\lambda > 0 : \sum_{k=1}^p \lambda_k = 1\}} \mathcal{S}(\lambda, \mathcal{Y}) \tag{3.5}$$

$$\text{and } \mathcal{S}_0(\mathcal{Y}) := \bigcup_{\lambda \in \mathbb{R}^p_\geq} \mathcal{S}(\lambda, \mathcal{Y}) = \bigcup_{\{\lambda \geq 0 : \sum_{k=1}^p \lambda_k = 1\}} \mathcal{S}(\lambda, \mathcal{Y}). \tag{3.6}$$

Clearly, the assumption $\sum_{k=1}^p \lambda_k = 1$ can always be made. It just normalizes the weights, but does not change $\mathcal{S}(\lambda, \mathcal{Y})$. It will thus be convenient to have the notation

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

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

It is also evident that using $\lambda = 0$ does not make sense, as $\mathcal{S}(0, \mathcal{Y}) = \mathcal{Y}$. We exclude this case henceforth. Finally,

$$\mathcal{S}(\mathcal{Y}) \subset \mathcal{S}_0(\mathcal{Y}) \tag{3.7}$$

follows directly from the definition. The results in the following two sections extend (3.7) by including links with efficient sets.

In many of the results of this chapter we will need some convexity assumptions. However, requiring $\mathcal{Y}$ to be convex is usually too restrictive a requirement. After all, we are looking for nondominated points, which, bearing Proposition 2.3 in mind are located in the "south-west" of $\mathcal{Y}$. Hence, we define $\mathbb{R}^p_{\geqq}$-convexity.

Definition 3.1. *A set $\mathcal{Y} \in \mathbb{R}^p$ is called $\mathbb{R}^p_{\geqq}$-convex, if $\mathcal{Y} + \mathbb{R}^p_{\geqq}$ is convex.*

Every convex set $\mathcal{Y}$ is clearly $\mathbb{R}^p_{\geqq}$-convex. The set $\mathcal{Y}$ of Figure 3.1 is neither convex nor $\mathbb{R}^p_{\geqq}$-convex. Figure 2.4 shows a nonconvex set $\mathcal{Y}$ which is $\mathbb{R}^p_{\geqq}$-convex.

A fundamental result about convex sets is that nonintersecting convex sets can be separated by a hyperplane.

Theorem 3.2. *Let $\mathcal{Y}_1, \mathcal{Y}_2 \subset \mathbb{R}^p$ be nonempty convex sets. There exists some $y^* \in \mathbb{R}^p$ such that*

$$\inf_{y \in \mathcal{Y}_1} \langle y, y^* \rangle \geq \sup_{y \in \mathcal{Y}_2} \langle y, y^* \rangle \tag{3.8}$$

$$\textit{and } \sup_{y \in \mathcal{Y}_1} \langle y, y^* \rangle > \inf_{y \in \mathcal{Y}_2} \langle y, y^* \rangle \tag{3.9}$$

if and only if $\mathrm{ri}(\mathcal{Y}_1) \cap \mathrm{ri}(\mathcal{Y}_2) = \emptyset$. *In this case* $\mathcal{Y}_1$ *and* $\mathcal{Y}_2$ *are said to be properly separated by a hyperplane with normal* y^*.

Recall that $\mathrm{ri}(\mathcal{Y}_i)$ is the relative interior of $\mathcal{Y}_i$, i.e. the interior in the space of appropriate dimension $\dim(\mathcal{Y}_i) \leq p$. A proof of Theorem 3.2 can be found in Rockafellar (1970, p. 97).

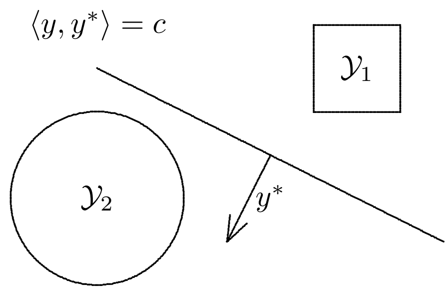

Fig. 3.2. Properly separated sets $\mathcal{Y}_1$ and $\mathcal{Y}_2$.

We will also use the following separation theorem.

Theorem 3.3. *Let* $\mathcal{Y} \subset \mathbb{R}^p$ *be a nonempty, closed, convex set and let* $y^0 \in \mathbb{R}^p \setminus \mathcal{Y}$. *Then there exists a* $y^* \in \mathbb{R}^p \setminus \{0\}$ *and* $\alpha \in \mathbb{R}$ *such that*

$$\langle y^*, y^0 \rangle < \alpha < \langle y^*, y \rangle$$

for all $y \in \mathcal{Y}$.

We will derive results on efficient solutions $x \in \mathcal{X}$ of a multicriteria optimization problem from results on nondominated points $y \in \mathcal{Y}$. This is done as follows. For results that are valid for any set $\mathcal{Y}$ we obtain analogous results simply by invoking the fact that efficient solutions are preimages of nondominated points. For results that are only valid under some conditions on $\mathcal{Y}$ (usually $\mathbb{R}^p_{\geqq}$-convexity) , appropriate assumptions on $\mathcal{X}$ and f are required. To ensure $\mathbb{R}^p_{\geqq}$-convexity of $\mathcal{Y}$ the assumption of convexity of $\mathcal{X}$ and all objective functions f_k.

3.1 Weighted Sum Scalarization and (Weak) Efficiency

In this section, we show that optimal solutions of the weighted sum problem (3.2) with positive (nonnegative) weights are always (weakly) efficient and that under convexity assumptions all (weakly) efficient solutions are optimal solutions of scalarized problems with positive (nonnegative) weights.

Theorem 3.4. *For any set $\mathcal{Y} \subset \mathbb{R}^p$ we have $\mathcal{S}_0(\mathcal{Y}) \subset \mathcal{Y}_{wN}$.*

Proof. Let $\lambda \in \mathbb{R}^p_{\geq}$ and $\hat{y} \in \mathcal{S}(\lambda, \mathcal{Y})$. Then

$$\sum_{k=1}^{p} \lambda_k \hat{y}_k \leq \sum_{k=1}^{p} \lambda_k y_k \text{ for all } y \in \mathcal{Y}.$$

Suppose that $\hat{y} \notin \mathcal{Y}_{wN}$. Then there is some $y' \in \mathcal{Y}$ with $y'_k < \hat{y}_k$, $k = 1, \ldots, p$. Thus,

$$\sum_{k=1}^{p} \lambda_k y'_k < \sum_{k=1}^{p} \lambda_k \hat{y}_k,$$

because at least one of the weights λ_k must be positive. This contradiction implies the result. $\square$

For $\mathbb{R}^p_{\geqq}$-convex sets we can prove the converse inclusion.

Theorem 3.5. *If $\mathcal{Y}$ is $\mathbb{R}^p_{\geqq}$-convex, then $\mathcal{Y}_{wN} = \mathcal{S}_(\mathcal{Y})$.*

Proof. Due to Theorem 3.4 we only have to show $\mathcal{Y}_{wN} \subset \mathcal{S}_(\mathcal{Y})$. We first observe that $\mathcal{Y}_{wN} \subset (\mathcal{Y} + \mathbb{R}^p_{>})_{wN}$ (The proof of this fact is the same as that of Proposition 2.3, replacing $\mathbb{R}^p_{\geqq}$ by $\mathbb{R}^p_{>}$).

Therefore, if $\hat{y} \in \mathcal{Y}_{wN}$, we have

$$(\mathcal{Y}_{wN} + \mathbb{R}^p_{>} - \hat{y}) \cap (-\mathbb{R}^p_{>}) = \emptyset.$$

This means that the intersection of the relative interiors of the two convex sets $\mathcal{Y} + \mathbb{R}^p_{>} - \hat{y}$ and $-\mathbb{R}^p_{>}$ is empty. By Theorem 3.2 there is some $\lambda \in \mathbb{R}^p \setminus \{0\}$ such that

$$\langle \lambda, y + d - \hat{y} \rangle \geq 0 \geq \langle \lambda, -d' \rangle \tag{3.10}$$

for all $y \in \mathcal{Y}, d \in \mathbb{R}^p_{>}, d' \in \mathbb{R}^p_{>}$.

Since $\langle \lambda, -d' \rangle \leq 0$ for all $d' \in \mathbb{R}^p_{>}$ we can choose $d' = e_k + \varepsilon e$ – where e_k is the k-th unit vector, $e = (1, \ldots, 1) \in \mathbb{R}^p$ is a vector of all ones, and $\varepsilon > 0$ is arbitrarily small – to see that $\lambda_k \geq 0, k = 1, \ldots, p$. On the other hand, choosing $d = \varepsilon e$ in $\langle \lambda, y + d - \hat{y} \rangle \geq 0$ implies

$$\langle \lambda, y \rangle + \varepsilon \langle \lambda, e \rangle \geq \langle \lambda, \hat{y} \rangle \tag{3.11}$$

for all $y \in \mathcal{Y}$ and thus

$$\langle \lambda, y \rangle > \langle \lambda, \hat{y} \rangle. \tag{3.12}$$

Therefore $\lambda \in \mathbb{R}^p_{\geq}$ and $\hat{y} \in \mathcal{S}(\lambda, \mathcal{Y}) \subset \mathcal{S}_(\mathcal{Y})$. $\square$

With Theorems 3.4 and 3.5 we have the first extension of inclusion (3.7), namely

$$\mathcal{S}(\mathcal{Y}) \subset \mathcal{S}_{(}(\mathcal{Y}) \subset \mathcal{Y}_{wN} \tag{3.13}$$

in general and

$$\mathcal{S}(\mathcal{Y}) \subset \mathcal{S}_{(}(\mathcal{Y}) = \mathcal{Y}_{wN} \tag{3.14}$$

for $\mathbb{R}_{\geqq}$-convex sets.

Next we relate $\mathcal{S}(\mathcal{Y})$ and $\mathcal{S}_{(}(\mathcal{Y})$ to $\mathcal{Y}_N$.

Theorem 3.6. *Let $\mathcal{Y} \subset \mathbb{R}^p$. Then $\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_N$.*

Proof. Let $\hat{y} \in \mathcal{S}(\mathcal{Y})$. Then there is some $\lambda \in \mathbb{R}^p_{>}$ satisfying $\sum_{k=1}^p \lambda_k \hat{y}_k \leq \sum_{k=1}^p \lambda_k y_k$ for all $y \in \mathcal{Y}$.

Suppose $\hat{y} \notin \mathcal{Y}_N$. Hence there must be $y' \in \mathcal{Y}$ with $y' \leq y$. Multiplying componentwise by the weights gives $\lambda_k y'_k \leq \lambda_k \hat{y}_k$ for all $k = 1, \ldots, p$ and strict inequality for one k. This strict inequality together with the fact that all λ_k are positive implies $\sum_{k=1}^p \lambda_k y'_k < \sum_{k=1}^p \lambda_k \hat{y}_k$, contradicting $\hat{y} \in \mathcal{S}(\mathcal{Y})$. □

Corollary 3.7. *$\mathcal{Y}_N \subset \mathcal{S}_{(}(\mathcal{Y})$ if $\mathcal{Y}$ is an $\mathbb{R}^p_{\geqq}$-convex set.*

Proof. This result is an immediate consequence of Theorem 3.5 since $\mathcal{Y}_N \subset \mathcal{Y}_{wN} = \mathcal{S}_{(}(\mathcal{Y})$. □

Theorem 3.6 and Corollary 3.7 yield

$$\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_N; \quad \mathcal{S}_{(}(\mathcal{Y}) \subset \mathcal{Y}_{wN} \tag{3.15}$$

in general and

$$\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_N \subset \mathcal{S}_{(}(\mathcal{Y}) = \mathcal{Y}_{wN} \tag{3.16}$$

for $\mathbb{R}^p_{\geqq}$-convex sets.

Theorem 3.6 can be extended by the following proposition.

Proposition 3.8. *If $\hat{y}$ is the unique element of $\mathcal{S}(\lambda, \mathcal{Y})$ for some $\lambda \in \mathbb{R}^p_{\geq}$ then $\hat{y} \in \mathcal{Y}_N$.*

Proof. The easy proof is left to the reader, see Exercise 3.2.

In Exercise 3.3 the reader is asked for examples where the inclusions in (3.15) and (3.16) are strict, demonstrating that these are the strongest relationship between weighted sum optimal points and (weakly) nondominated points that can be proved for general and $\mathbb{R}^p_{\geqq}$-convex sets, without additional assumptions, like the uniqueness of Proposition 3.8

Let us now summarize the analogies of the results of this section in terms of the decision space, i.e. (weakly) efficient solutions of multicriteria optimization problems.

Proposition 3.9. *Suppose that $\hat{x}$ is an optimal solution of the weighted sum optimization problem*

$$\min_{x \in \mathcal{X}} \sum_{k=1}^{p} \lambda_k f_k(x). \tag{3.17}$$

with $\lambda \in \mathbb{R}^p_{\geq}$. Then the following statements hold.

1. *If $\lambda \in \mathbb{R}^p_{\geq}$ then $\hat{x} \in \mathcal{X}_{wE}$.*
2. *If $\lambda \in \mathbb{R}^p_{>}$ then $\hat{x} \in \mathcal{X}_{E}$.*
3. *If $\lambda \in \mathbb{R}^p_{\geq}$ and $\hat{x}$ is a unique optimal solution of (3.17) then $\hat{x} \in \mathcal{X}_{sE}$.*

Proof. The assertions follow directly from Theorem 3.4, Theorem 3.6, and Proposition 3.8 with the uniqueness of $\hat{x}$, respectively. □

Proposition 3.10. *Let $\mathcal{X}$ be a convex set, and let f_k be convex functions, $k = 1, \ldots, p$. If $\hat{x} \in \mathcal{X}_{wE}$ there is some $\lambda \in \mathbb{R}^p_{\geq}$ such that $\hat{x}$ is an optimal solution of* (3.17)*.*

The proof follows from Theorem 3.5. Note that there is no distinction between $\mathcal{X}_{wE}$ and $\mathcal{X}_E$ here, an observation that we shall regrettably have to make for almost all methods to find efficient solutions. This problem is caused by the (possibly) strict inclusions in (3.16). Therefore the examples of Exercise 3.3 and the usual trick of identifying decision and objective space should convince the reader that this problem cannot be avoided.

At the end of this section, we point out that Exercises 3.4 and 3.7 show how to generalize the weighted sum scalarization for nondominated points with respect to a convex and pointed cone $\mathcal{C}$.

Remembering that $\mathcal{Y}_{pN} \subset \mathcal{Y}_N$, we continue our investigations by looking at relationships between $\mathcal{Y}_{pN}$ and $\mathcal{S}(\mathcal{Y})$.

3.2 Weighted Sum Scalarization and Proper Efficiency

Here we will establish the relationships between properly nondominated points (in Benson's or Geoffrion's sense) and optimal points of weighted sum scalarizations with positive weights. The main result shows that these points coincide for convex sets. A deeper result shows that in this situation the difference between nondominated and properly nondominated points is small: The set of properly nondominated points is dense in the nondominated set.

From now on we denote the set of properly efficient points in Geoffrion's sense by $\mathcal{Y}_{pE}$.. Note that due to Theorem 2.48 Geoffrion's and Benson's definitions are equivalent for efficiency defined by $\mathbb{R}^p_{\geqq}$.

Unless otherwise stated, $\mathcal{X}_{pN}$ will denote the set of properly efficient solutions of a multicriteria optimization problem in Geoffrion's sense. Our first result shows that an optimal solution of (3.2) is a properly efficient solution of (3.1) if $\lambda > 0$.

Theorem 3.11 (Geoffrion (1968)). *Let* $\lambda_k > 0$, $k = 1, \ldots, p$ *with* $\sum_{k=1}^{p} \lambda_k = 1$ *be positive weights. If* $\hat{x}$ *is an optimal solution of* (3.2) *then* $\hat{x}$ *is a properly efficient solution of* (3.1)

Proof. Let $\hat{x}$ be an optimal solution of (3.2). To show that $\hat{x}$ is efficient suppose there exists some $x' \in \mathcal{X}$ with $f(x') \leq f(\hat{x})$. Positivity of the weights λ_k and $f_i(x') < f_i(\hat{x})$ for some $i \in \{1, \ldots, p\}$ imply the contradiction

$$\sum_{k=1}^{p} \lambda_k f_k(x') < \sum_{k=1}^{p} \lambda_k f_k(\hat{x}). \tag{3.18}$$

To show that $\hat{x}$ is properly efficient, we choose an appropriately large number M such that assuming there is a trade-off bigger than M yields a contradiction to optimality of $\hat{x}$ for the weighted sum problem. Let

$$M := (p-1) \max_{i,j} \frac{\lambda_j}{\lambda_i}. \tag{3.19}$$

Suppose that $\hat{x}$ is not properly efficient. Then there exist $i \in \{1, \ldots, p\}$ and $x \in \mathcal{X}$ such that $f_i(x) < f_i(\hat{x})$ and $f_i(\hat{x}) - f_i(x) > M(f_j(x) - f_j(\hat{x}))$ for all $j \in \{1, \ldots, p\}$ such that $f_j(\hat{x}) < f_j(x)$. Therefore

$$f_i(\hat{x}) - f_i(x) > \frac{p-1}{\lambda_i} \lambda_j (f_j(x) - f_j(\hat{x})) \tag{3.20}$$

for all $j \neq i$ by the choice of M (note that the inequality is trivially true if $f_j(\hat{x}) > f_j(x)$). Multiplying each of these inequalities by $\lambda_i/(p-1)$ and summing them over $j \neq i$ yields

$$\begin{aligned}
\lambda_i(f_i(\hat{x}) - f_i(x)) &> \sum_{j \neq i} \lambda_j (f_j(x) - f_j(\hat{x})) \\
\Rightarrow \quad \lambda_i f_i(\hat{x}) - \lambda_i f_i(x) &> \sum_{j \neq i} \lambda_j f_j(x) - \sum_{j \neq i} \lambda_j f_j(\hat{x}) \\
\Rightarrow \quad \lambda_i f_i(\hat{x}) + \sum_{j \neq i} \lambda_j f_j(\hat{x}) &> \lambda_i f_i(x) + \sum_{j \neq i} \lambda_j f_j(x) \\
\Rightarrow \quad \sum_{i=1}^{p} \lambda_i f_i(\hat{x}) &> \sum_{i=1}^{p} \lambda_i f_i(x),
\end{aligned}$$

contradicting optimality of $\hat{x}$ for (3.2). Thus $\hat{x}$ is properly efficient. $\square$

Theorem 3.11 immediately yields $\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_{pE}$, strengthening the left part of (3.15).

Corollary 3.12. *Let* $\mathcal{Y} \subset \mathbb{R}^p$. *Then* $\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_{pN}$.

Now that we have a sufficient condition for proper nondominance and proper efficiency, the natural question is, whether this condition is also necessary. In general it is not. We shall illustrate this graphically.

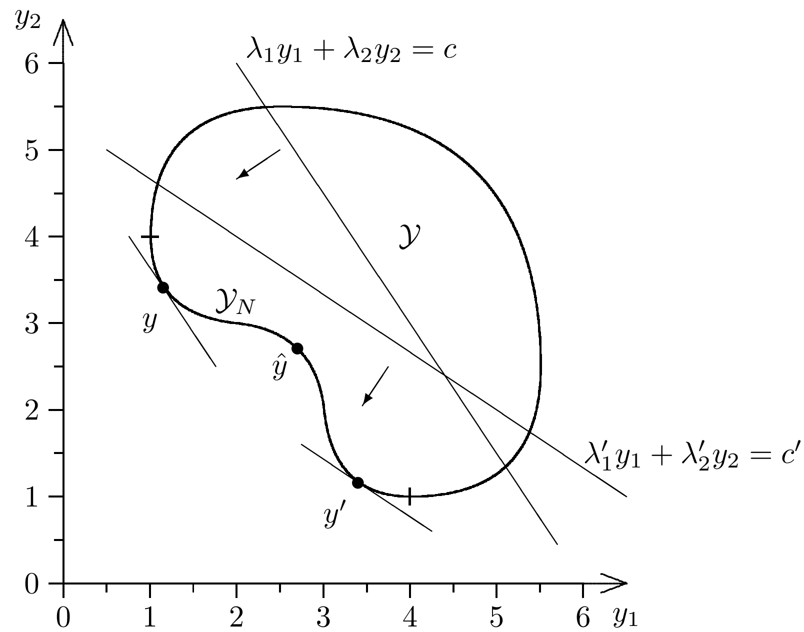

Fig. 3.3. Properly nondominated $\hat{y} \in \mathcal{Y}_N$.

In Figure 3.3, the feasible set in objective space for a nonconvex problem is shown ($\mathcal{Y}$ is not $\mathbb{R}_{\geqq}$-convex). Since all objective vectors $y = (f_1(x), \ldots, f_p(x))$, which attain the same value $c = \sum_{k=1}^{p} \lambda_k f_k(x)$ of the weighted sum objective, are located on a straight line, the minimization problem (3.4) amounts to pushing this line towards the origin, until the intersects $\mathcal{Y}$ only on the boundary of $\mathcal{Y}$. In Figure 3.3 this is illustrated for two weighting vectors (λ_1, λ_2) and (λ'_1, λ'_2), that lead to the nondominated points y and y'. It is obvious that the third point $\hat{y} \in \mathcal{Y}_N$ is properly nondominated, but none of its preimages x under f can be an optimal solution of (3.2) for any choice of $(\lambda_1, \ldots, \lambda_p) \in \mathbb{R}^p_>$.

The converse of Theorem 3.11 and Corollary 3.12 can be shown for $\mathbb{R}^p_{\geqq}$-convex sets. We shall give a prove in objective space using Benson's definition and a proof in decision space that uses Geoffrion's definition.

Theorem 3.13. *If $\mathcal{Y}$ is $\mathbb{R}^p_{\geqq}$-convex then $\mathcal{Y}_{pE} \subset \mathcal{S}(\mathcal{Y})$.*

Proof. Let $\hat{y} \in \mathcal{Y}_{pE}$, i.e.

$$\mathrm{cl}(\mathrm{cone}(\mathcal{Y} + \mathbb{R}^p_{\geqq} - \hat{y})) \cap (-\mathbb{R}^p_{\geqq}) = \{0\}. \tag{3.21}$$

By definition, $\mathrm{cl}(\mathrm{cone}(\mathcal{Y} + \mathbb{R}^p_{\geqq} - \hat{y}))$ is a closed convex cone.

The idea of the proof is that if there exists a $\lambda \in \mathbb{R}^p_{>}$ such that

$$\langle \lambda, d \rangle \geq 0 \text{ for all } d \in \mathrm{cl}(\mathrm{cone}(\mathcal{Y} + \mathbb{R}_{\geqq} - y^*)) =: \mathcal{K} \tag{3.22}$$

we get, in particular,

$$\langle \lambda, y - y^* \rangle \geq 0 \text{ for all } y \in \mathcal{Y}, \tag{3.23}$$

i.e. $\langle \lambda, y \rangle \geq \langle \lambda, \hat{y} \rangle$ for all $y \in \mathcal{Y}$ and thus $\hat{y} \in \mathcal{S}(\mathcal{Y})$. This is true, because $\mathcal{Y} - \hat{y} \subset \mathrm{cl}(\mathrm{cone}(\mathcal{Y} + \mathbb{R}^p_{\geqq} - \hat{y}))$. We now prove the existence of $\lambda \in \mathbb{R}^p_{\geqq}$ with property (3.22).

Assume no such λ exists. Both $\mathbb{R}^p_{>}$ and

$$\mathcal{K}^\circ := \{\mu \in \mathbb{R}^p : \ \langle \mu, d \rangle \geq 0 \text{ for all } d \in \mathcal{K}\} \tag{3.24}$$

are convex sets and because of our assumption have nonintersecting relative interiors. Therefore we can apply Theorem 3.2 to get some nonzero $y^* \in \mathbb{R}^p_{\geq}$ and $\beta \in \mathbb{R}$ such that

$$\langle y^*, \mu \rangle \leq \beta \text{ for all } \mu \in \mathbb{R}^p_{>} \tag{3.25}$$

$$\langle y^*, \mu \rangle \geq \beta \text{ for all } \mu \in K^\circ. \tag{3.26}$$

Using $\mu' = \alpha\mu$ for some arbitrary but fixed $\mu \in K^\circ$ and letting $\alpha \to \infty$ in (3.26) we get $\beta = 0$. Therefore

$$\langle y^*, \mu \rangle \leq 0 \text{ for all } \mu \in \mathbb{R}^p_{>}. \tag{3.27}$$

Selecting $\mu = \varepsilon e + e_k = (\varepsilon, \ldots, \varepsilon, 1, \varepsilon, \ldots, \varepsilon)$ and letting $\varepsilon \to 0$ in (3.27) we obtain $y^*_k \leq 0$ for all $k = 1, \ldots, p$, i.e.

$$y^* \in -\mathbb{R}^p_{\geq}. \tag{3.28}$$

Let

$$\mathcal{K}^{\circ\circ} := \{y \in \mathbb{R}^p : \ \langle y, \mu \rangle \geq 0 \text{ for all } \mu \in \mathcal{K}^\circ\}. \tag{3.29}$$

According to (3.27), $y^* \in K^{\circ\circ}$. Once we have shown that $\mathcal{K}^{\circ\circ} \subset \mathrm{cl}\,\mathcal{K} = \mathcal{K}$ we know that

$$y^* \in \mathcal{K}. \tag{3.30}$$

Finally (3.28) and (3.30) imply that $y^* \in \mathcal{K} \cap (-\mathbb{R}^p_{\geqq})$ with $y^* \neq 0$ contradicting the proper nondominance conditions (3.21). Therefore the desired λ satisfying (3.22) exists.

To complete the proof, we have to show that $\mathcal{K}^{\circ\circ} \subset \operatorname{cl}\mathcal{K} = \mathcal{K}$. Let $y \in \mathbb{R}^p,\ y \notin \mathcal{K}$. Using Theorem 3.3 to separate $\{y\}$ and $\mathcal{K}$ we get $y^* \in \mathbb{R}^p,\ y^* \neq 0$ and $\alpha \in \mathbb{R}$ with $\langle d, y^* \rangle > \alpha$ for all $d \in \mathcal{K}$ and $\langle y, y^* \rangle < \alpha$. Then $0 \in \mathcal{K}$ implies $\alpha < 0$ and therefore $\langle y, y^* \rangle < 0$. Taking $d = \alpha d'$ for some arbitrary but fixed d' and letting $\alpha \to \infty$ we get $\langle d, y^* \rangle \geq 0$ for all $d \in \mathcal{K}$, i.e $y^* \in \mathcal{K}^\circ$. So $\langle y, y^* \rangle < 0$ implies $y \notin \mathcal{K}^{\circ\circ}$. Hence $\mathcal{K}^{\circ\circ} \subset \mathcal{K}$. □

The properties of and relationships among $\mathcal{K}, \mathcal{K}^\circ$, and $\mathcal{K}^{\circ\circ}$ we have used here are true for cones $\mathcal{K}$ in general, not just for the one used above. See Exercise 3.6 for more details. Let us now illustrate Theorem 3.13.

Example 3.14. Consider the set $\mathcal{Y} = \{(y_1, y_2) : y_1^2 + y_2^2 \leq 1\}$. Here

$$\mathcal{Y}_N = \{(y_1, y_2) : y_1^2 + y_2^2 = 1,\ y_1 \leq 0,\ y_2 \leq 0\}, \tag{3.31}$$

$$\mathcal{Y}_{pN} = \mathcal{Y}_N \setminus \{(-1, 0), (0, -1)\}. \tag{3.32}$$

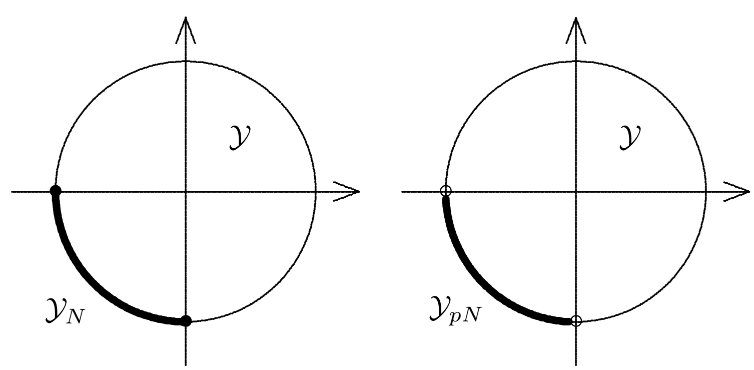

Fig. 3.4. Sets $\mathcal{Y}_N$ and $\mathcal{Y}_{pN}$.

All properly nondominated points $\hat{y}$ are optimal points of weighted sum scalarizations (indeed, the weights correspond to the normals of the tangents to the circle at $\hat{y}$). The two boundary points of $\mathcal{Y}_N$ are not properly nondominated. But $(-1, 0)$ and $(0, -1)$ are unique optimal solutions of

$$\min_{y \in \mathcal{Y}} \lambda_1 y_1 + \lambda_2 y_2 \tag{3.33}$$

for $\lambda = (1, 0)$ and $\lambda = (0, 1)$, respectively, and therefore belong to the nondominated set $\mathcal{Y}_N$, see Proposition 3.8. □

Theorem 3.15 (Geoffrion (1968)). *Let $\mathcal{X} \subset \mathbb{R}^n$ be convex and assume $f_k : \mathcal{X} \to \mathbb{R}$ are convex for $k = 1, \ldots, p$. Then $\hat{x} \in \mathcal{X}$ is properly efficient if and only if $\hat{x}$ is an optimal solution of* (3.2)*, with strictly positive weights $\lambda_k, k = 1, \ldots, p$.*

Proof. Due to Theorem 3.11 we only have to prove necessity of the condition. Let $\hat{x} \in \mathcal{X}$ be properly efficient. Then, by definition, there exists a number $M > 0$ such that for all $i = 1, \ldots, p$ the system

$$\begin{aligned} f_i(x) &< f_i(\hat{x}) \\ f_i(x) + M f_j(x) &< f_i(\hat{x}) + M f_j(\hat{x}) \text{ for all } j \neq i \end{aligned} \tag{3.34}$$

has no solution. To see that, simply rearrange the inequalities in (2.41).

A property of convex functions, which we state as Theorem 3.16 below implies that for the ith such system there exist $\lambda_k^i \geq 0$, $k = 1, \ldots, p$ with $\sum_{k=1}^p \lambda_k^i = 1$ such that for all $x \in \mathcal{X}$ the following inequalities holds.

$$\begin{aligned} & \lambda_i^i f_i(x) + \sum_{k \neq i} \lambda_k^i (f_i(x) + M f_k(x)) \geq \lambda_i^i f_i(\hat{x}) + \sum_{k \neq i} \lambda_k^i (f_i(\hat{x}) + M f_k(\hat{x})) \\ \Leftrightarrow\ & \lambda_i^i f_i(x) + \sum_{k \neq i} \lambda_k^i f_i(x) + M \sum_{j \neq i} \lambda_j^i f_k(x) \geq \lambda_i^i f_i(\hat{x}) + \sum_{k \neq i} \lambda_k^i f_i(\hat{x}) + M \sum_{k \neq i} \lambda_j^i f_k(\hat{x}) \\ \Rightarrow\ & \sum_{k=1}^p \lambda_k^i f_i(x) + M \sum_{k \neq i} \lambda_k^i f_k(x) \geq \sum_{k=1}^p \lambda_k^i f_i(\hat{x}) + M \sum_{k \neq i} \lambda_k^i f_j(\hat{x}) \\ \Leftrightarrow\ & f_i(x) + M \sum_{k \neq i} \lambda_k^i f_k(x) \geq f_i(\hat{x}) + M \sum_{k \neq i} \lambda_k^i f_k(\hat{x}) \end{aligned}$$

We have such an inequality for each $i = 1, \ldots, p$ and now simply sum over i to obtain

$$\begin{aligned} & \sum_{i=1}^p f_i(x) + M \sum_{i=1}^p \sum_{k \neq i} \lambda_k^i f_k(x) \geq \sum_{i=1}^p f_i(\hat{x}) + M \sum_{k=1}^p \sum_{k \neq i} \lambda_j^i f_k(\hat{x}) \\ \Rightarrow\ & \sum_{k=1}^p \left(1 + M \sum_{i \neq k} \lambda_k^i \right) f_k(x) \geq \sum_{k=1}^p \left(1 + M \sum_{i \neq k} \lambda_k^i \right) f_k(x^*) \end{aligned}$$

for all $x \in \mathcal{X}$.

We can now normalize the values $(1 + M \sum_{k \neq i} \lambda_k^i)$, so that they sum up to one to obtain positive $\lambda_i, i = 1, \ldots, p$ for which $\hat{x}$ is optimal in (3.2). $\square$

The theorem, which we have used, is the following. For a proof we refer to Mangasarian (1969, p. 65).

Theorem 3.16. *Let $\mathcal{X} \subset \mathbb{R}^n$ be a convex set, let $h_k : \mathbb{R}^n \to \mathbb{R}$ be convex functions, $k = 1, \ldots, p$. Then, if the system $h_k(x) < 0$, $k = 1, \ldots, p$ has no solution $x \in \mathcal{X}$, there exist $\lambda_k \geq 0$, $\sum_{k=1}^p \lambda_k = 1$ such that all $x \in \mathcal{X}$ satisfy*

$$\sum_{k=1}^{p} \lambda_k h_k(x) \geq 0. \tag{3.35}$$

With these results on proper nondominance and proper efficiency we can extend (3.15) and (3.16) as follows:

$$\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_{pE} \subset \mathcal{Y}_E \text{ and } \mathcal{S}_(\mathcal{Y}) \subset \mathcal{Y}_{wE} \tag{3.36}$$

holds for general sets, whereas for $\mathbb{R}^p_{\geqq}$-convex sets

$$\mathcal{S}(\mathcal{Y}) = \mathcal{Y}_{pE} \subset \mathcal{Y}_E \subset \mathcal{Y}_{wE} = \mathcal{S}_(\mathcal{Y}). \tag{3.37}$$

A closer inspection of the inclusions reveals that the gap between $\mathcal{Y}_{wE}$ and $\mathcal{Y}_E$ might be quite large, even in the convex cases (see Example 2.27 for an illustration). This is not possible for the gap between $\mathcal{Y}_{pE}$ and $\mathcal{Y}_E$.

Theorem 3.17 (Hartley (1978)). *If $\mathcal{Y} \neq \emptyset$ is $\mathbb{R}^p_{\geqq}$-closed and $\mathbb{R}^p_{\geqq}$-convex, the following inclusions hold:*

$$\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_N \subset \operatorname{cl} \mathcal{S}(\mathcal{Y}) = \operatorname{cl} \mathcal{Y}_{pN}. \tag{3.38}$$

Proof. The only inclusion we have to show is $\mathcal{Y}_N \subset \operatorname{cl} \mathcal{S}(\mathcal{Y})$. Since $\mathcal{Y}_N = (\mathcal{Y} + \mathbb{R}^p_{\geqq})_N$ and $\mathcal{S}(\mathcal{Y}) = \mathcal{S}(\mathcal{Y} + \mathbb{R}^p_{\geqq})$, we only prove it for a closed convex set $\mathcal{Y}$. Without loss of generality we shall also assume that $\hat{y} = 0 \in \mathcal{Y}_N$.

The proof proceeds in two steps: First we show the result for compact $\mathcal{Y}$, applying a minimax theorem to the scalar product on two compact convex sets. We shall then prove the general case by reduction to the compact case.

Case 1: $\mathcal{Y}$ is compact and convex.

Choose $d \in \mathbb{R}^p_{>}$ and $\mathcal{C}(\varepsilon) := \varepsilon d + \mathbb{R}^p_{\geqq}$ for $0 < \varepsilon \in \mathbb{R}$. If ε is sufficiently small, $\mathcal{C}(\varepsilon) \cap \mathcal{B}(0,1)$ is nonempty. Thus, both $\mathcal{Y}$ and $\mathcal{C}(\varepsilon) \cap \mathcal{B}(0,1)$ are nonempty, convex, and compact.

Applying the Sion-Kakutani minimax theorem (Theorem 3.18 below) to $\Phi = \langle \cdot, \cdot \rangle$ with $\mathcal{C} = \mathcal{C}(\varepsilon) \cap \mathcal{B}(0,1)$ and $\mathcal{D} = \mathcal{Y}$ we get the existence of $y(\varepsilon) \in \mathcal{Y}$ and $\lambda(\varepsilon) \in \mathcal{C}(\varepsilon) \cap \mathcal{B}(0,1)$ such that

$$\langle \lambda, y(\varepsilon) \rangle \leq \langle \lambda(\varepsilon), y(\varepsilon) \rangle \leq \langle \lambda(\varepsilon), y \rangle \text{ for all } y \in \mathcal{Y}, \text{ for all } \lambda \in \mathcal{C}(\varepsilon) \cap \mathcal{B}(0,1) \tag{3.39}$$

From (3.39) using $0 \in \mathcal{Y}$ we obtain $\langle \lambda, y(\varepsilon) \rangle \leq 0$ for all $\lambda \in \mathcal{C}(\varepsilon) \cap \mathcal{B}(0,1)$. Because $\mathcal{Y}$ is compact there exists a sequence $\varepsilon^k \to 0$ such that $y^k := y(\varepsilon^k) \to y' \in \mathcal{Y}$ for $k \to \infty$.

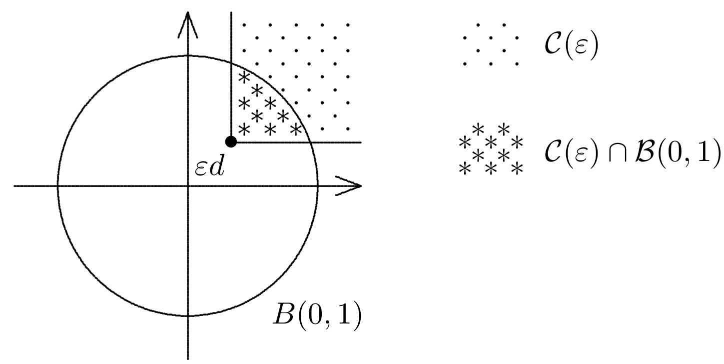

Fig. 3.5. Illustration of the first case in the proof of Theorem 3.17.

Furthermore, for each $\lambda \in \mathbb{R}^p_> \cap \mathcal{B}(0,1)$ there is some $\varepsilon' > 0$ such that $\lambda \in \mathcal{C}(\varepsilon) \cap \mathcal{B}(0,1)$ for all $\varepsilon \leq \varepsilon'$ and therefore $\langle \lambda, y^k \rangle \leq 0$ when k is large enough. The convergence $y^k \to y'$ then implies $\langle \lambda, \hat{y} \rangle \leq 0$ for all $\lambda \in \mathbb{R}^p_>$. This implies $y' \in -\mathbb{R}^p_{\geqq}$. Thus, $y' \leqq 0$ but since $\hat{y} = 0 \in \mathcal{Y}_N$ we must have $y' = 0$.

Next we show that $y' = \hat{y} = 0 \in \operatorname{cl} \mathcal{S}(\mathcal{Y})$. To this end let $\lambda^k := \lambda(\varepsilon^k)/\|\lambda(\varepsilon^k)\| \in \mathbb{R}^p_> \cap \operatorname{bd} \mathcal{B}(0,1)$, where $\lambda(\varepsilon^k)$ is the λ associated with ε^k and $y(\varepsilon^k)$ to satisfy (3.39). Therefore we have

$$\langle \lambda^k, y(\varepsilon^k) \rangle \leq \langle \lambda^k, y \rangle \text{ for all } y \in \mathcal{Y}, \tag{3.40}$$

i.e. $y^k = y(\varepsilon^k) \subset \mathcal{S}(\lambda^k, \mathcal{Y}) \subset \mathcal{S}(\mathcal{Y})$. Since $y' = \lim y^k$ this implies $\hat{y} = y' \in \operatorname{cl} \mathcal{S}(\mathcal{Y})$.

Case 2: $\mathcal{Y}$ is closed and convex (but not necessarily compact).

Again let $\hat{y} = 0 \in \mathcal{Y}_N$. $\mathcal{Y} \cap \mathcal{B}(0,1)$ is nonempty, convex, and compact and $0 \in (\mathcal{Y} \cap \mathcal{B}(0,1))_N$. Case 1 yields the existence of $\lambda^k \in \mathbb{R}^p_>$, $\|\lambda^k\| = 1$, and $y^k \in \mathcal{S}(\lambda^k, \mathcal{Y} \cap B(0,1))$ with $y^k \to 0$. We show that $y^k \in \mathcal{S}(\lambda^k, \mathcal{Y})$, which completes the proof.

Note that for k large enough $y^k \in \operatorname{int} \mathcal{B}$ (since $y^k \to 0$) and suppose $y' \in \mathcal{Y}$ exists with $\langle \lambda^k, y' \rangle < \langle \lambda^k, y^k \rangle$. Then $\alpha y' + (1-\alpha) y^k \in \mathcal{Y} \cap B(0,1)$ for sufficiently small α (see Figure 3.6).

This implies

$$\langle \lambda^k, \alpha y' + (1-\alpha) y^k \rangle = \alpha \langle \lambda^k, y' \rangle + (1-\alpha) \langle \lambda^k, y^k \rangle < \langle \lambda^k, y^k \rangle, \tag{3.41}$$

contradicting $y^k \in \mathcal{S}(\lambda^k, \mathcal{Y})$. □

The Sion-Kakutani minimax theorem that we used is stated for completeness. For a proof we refer to Stoer and Witzgall (1970, p. 232).

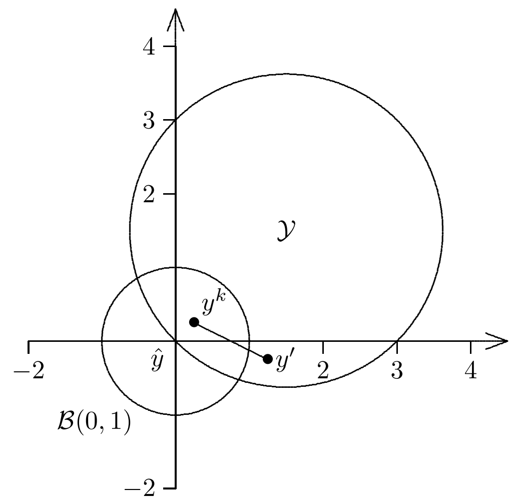

Fig. 3.6. Illustration of the second case in the proof of Theorem 3.17.

Theorem 3.18 (Sion-Kakutani minimax theorem). *Let $\mathcal{C}, \mathcal{D} \subset \mathbb{R}^p$ be nonempty, compact, convex sets and $\Phi : \mathcal{C} \times \mathcal{D} \to \mathbb{R}$ be a continuous mapping such that $\Phi(\cdot, d)$ is convex for all $d \in \mathcal{D}$ and $\Phi(c, \cdot)$ is concave for all $c \in \mathcal{C}$. Then*

$$\max_{d\in\mathcal{D}} \min_{c\in\mathcal{C}} \Phi(c,d) = \min_{c\in\mathcal{C}} \max_{d\in\mathcal{D}} \Phi(c,d). \tag{3.42}$$

Although Theorem 3.17 shows that $\mathcal{Y}_N \subset \operatorname{cl} \mathcal{Y}_{pN}$, the inclusion $\operatorname{cl} \mathcal{Y}_{pN} \subset \mathcal{Y}_N$ is not always satisfied.

Example 3.19 (Arrow et al. *(1953)).*

Consider the set $\mathcal{Y}' = \{(y_1, y_2, y_3) : (y_1 - 1)^2 + (y_2 - 1)^2 = 1,\ y_1 \leq 1,\ y_2 \leq 1,\ y_3 = 1\}$ and define

$$Y := \operatorname{conv}\left(\mathcal{Y}' \cup \{(1,0,0)\}\right), \tag{3.43}$$

shown from different angles in Figure 3.7.

$\mathcal{Y}$ is closed and convex. Note that $\hat{y} = (1,0,1) \notin \mathcal{Y}_N$ because $(1,0,0) \leq \hat{y}$. From Theorem 3.13 we know $\mathcal{Y}_{pN} = \mathcal{S}(\mathcal{Y})$. We show that all $y' \in Y'$ with $y'_1 < 1, y'_2 < 1$ are properly efficient.

Let $y' = (1 - \cos\theta, 1 - \sin\theta, 1)$ for $0 < \theta < \pi/2$ and $\lambda = (1 - \alpha)(\cos\theta, \sin\theta, 0) + \alpha(0,0,1)$ with $0 < \alpha < 1$ so that $\lambda \in \mathbb{R}^p_>$.

We compute $\langle \lambda, y - y' \rangle$ for $y = (1 - \cos\theta', 1 - \sin\theta', 1)$, $0 \leq \theta' \leq \pi/2$:

$$\begin{aligned}\langle \lambda, y - y' \rangle &= (1-\alpha)\left[\cos\theta(\cos\theta - \cos\theta') + \sin\theta(\sin\theta - \sin\theta')\right] \\ &= (1-\alpha)(1 - (\cos\theta\cos\theta' + \sin\theta\sin\theta')) \\ &= (1-\alpha)(1 - \cos(\theta - \theta')) \geq 0.\end{aligned} \tag{3.44}$$

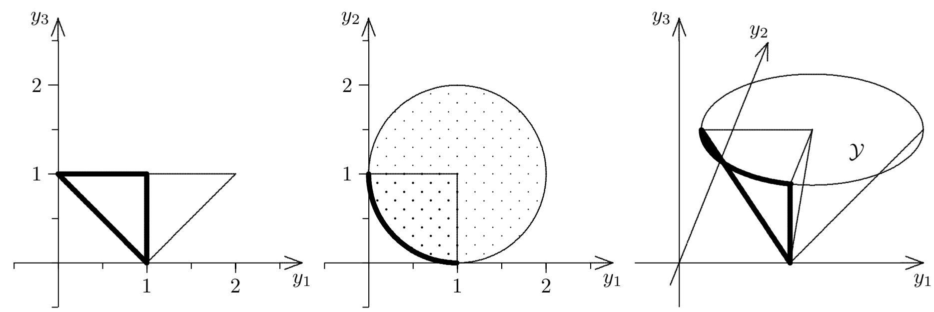

Fig. 3.7. The set $\mathcal{Y}$ in Example 3.19.

Furthermore, for $y = (1, 0, 0)$ we get

$$\begin{aligned}\langle \lambda, (1,0,0) - y' \rangle &= (1-\alpha)\left[\cos^2\theta - \sin\theta(1-\sin\theta)\right] - \alpha \\ &= (1-\alpha)(1-\sin\theta) - \alpha \quad > 0 \qquad (3.45)\end{aligned}$$

for small α. So by taking convex combinations of (3.44) and (3.45) we get $\langle \lambda, y - \overline{y} \rangle \geq 0$ for all $y \in \mathcal{Y}$ and thus $y' \in \mathcal{S}(\mathcal{Y})$. In addition, for $\theta \to 0$ we get $y' \to \hat{y}$ which is therefore in $\operatorname{cl} \mathcal{S}(\mathcal{Y})$. □

3.3 Optimality Conditions

In this section we prove necessary and sufficient conditions for weak and proper efficiency of solutions of a multicriteria optimization problem. These results follow along the lines of Karush-Kuhn-Tucker optimality conditions known from single objective nonlinear programming. We use the results to prove the yet missing link in Figure 2.17 and we give an example that shows that Kuhn and Tucker's and Geoffrion's definitions of properly efficient solutions do not always coincide.

We recall the Karush-Kuhn-Tucker necessary and sufficient optimality conditions in single objective nonlinear programming, see e.g. Bazaraa *et al.* (1993).

Theorem 3.20. *Let $f, g_j : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable functions and consider the single objective optimization problem*

$$\min\{f(x) : g_j(x) \leqq 0, j = 1, \ldots, m\}. \qquad (3.46)$$

Denote $\mathcal{X} := \{x \in \mathbb{R}^n : g_j(x) \leq 0, j = \{1, \ldots, m\}\}$.

- *If $\hat{x} \in \mathcal{X}$ is a (locally) optimal solution of* (3.46) *there is some $\hat{\mu} \in \mathbb{R}^m_{\geqq}$ such that*

$$\nabla f(\hat{x}) + \sum_{j=1}^{m} \hat{\mu}_j \nabla g_j(\hat{x}) = 0, \tag{3.47}$$

$$\sum_{j=1}^{m} \hat{\mu}_j g_j(\hat{x}) = 0. \tag{3.48}$$

- *If f, g_j are convex and there are $\hat{x} \in \mathcal{X}$ and $\hat{\mu} \in \mathbb{R}^m_{\geqq}$ such that* (3.47) *and* (3.48) *hold then $\hat{x}$ is a locally, thus globally, optimal solution of* (3.46).

We start with conditions for weak efficiency.

Theorem 3.21. *Suppose that the KT constraint qualification (see Definition 2.50) is satisfied at $\hat{x} \in \mathcal{X}$. If $\hat{x}$ is weakly efficient there exist $\hat{\lambda} \in \mathbb{R}^p_{\geq}$ and $\hat{\mu} \in \mathbb{R}^m_{\geqq}$ such that*

$$\sum_{k=1}^{p} \hat{\lambda}_k \nabla f_k(\hat{x}) + \sum_{j=1}^{m} \hat{\mu}_j \nabla g_j(\hat{x}) = 0 \tag{3.49}$$

$$\sum_{j=1}^{m} \hat{\mu}_j g_j(\hat{x}) = 0 \tag{3.50}$$

$$\hat{\lambda} \geq 0 \tag{3.51}$$

$$\hat{\lambda} \geqq 0 \tag{3.52}$$

Proof. Let $\hat{x} \in \mathcal{X}_{wE}$. We first show that there can be no $d \in \mathbb{R}^n$ such that

$$\langle \nabla f_k(\hat{x}), d \rangle < 0 \text{ for all } k = 1, \ldots, p \tag{3.53}$$

$$\langle \nabla g_j(\hat{x}), d \rangle < 0 \text{ for all } j \in \mathcal{J}(\hat{x}) := \{j : g_j(\hat{x}) = 0\}. \tag{3.54}$$

We then apply Motzkin's theorem of the alternative (Theorem 3.22) to obtain the multipliers $\hat{\lambda}_k$ and $\hat{\mu}_j$.

Suppose that such a $d \in \mathbb{R}^n$ exists. From the KT constraint qualification there is a continuously differentiable function $\theta : [0, \bar{t}] \to \mathbb{R}^n$ such that $\theta(0) = \hat{x},\ g(\theta(t)) \leq 0$ for all $t \in [0, \bar{t}]$, and $\theta'(0) = \alpha d$ with $\alpha > 0$. Thus,

$$f_k(\theta(t)) = f_k(\hat{x}) + t\langle \nabla f_k(\hat{x}), \alpha d \rangle + o(t) \tag{3.55}$$

and using $\langle \nabla f_k(\hat{x}), d \rangle < 0$ it follows that $f_k(\theta(t)) < f_k(\hat{x}),\ k = 1, \ldots, p$ for t sufficiently small, which contradicts $\hat{x} \in \mathcal{X}_{wE}$.

It remains to show that (3.53) and (3.54) imply the conditions of (3.49) – (3.52). This is achieved by using matrices $B = (\nabla f_k(\hat{x}))_{k=1,\ldots,p}$, $C =$

$(\nabla g_j(\hat{x}))_{j\in\mathcal{J}(\hat{x})}$, $D = 0$ with $l = |\mathcal{J}(\hat{x})|$ in Theorem 3.22. Then, since (3.53) and (3.54) have no solution $d \in \mathbb{R}^n$, according to Theorem 3.22 there must be $y^1 =: \hat{\lambda}, y^2 =: \hat{\mu}$, and y^3 such that $B^T y^1 + C^T y^2 = 0$ with $y^1 \geq 0$ and $y^2 \geqq 0$, i.e.

$$\sum_{k=1}^{p} \hat{\lambda}_k \nabla f_k(\hat{x}) + \sum_{j=\mathcal{J}(\hat{x})}^{m} \hat{\mu}_j \nabla g_j(\hat{x}) = 0.$$

We complete the proof by setting $\hat{\mu}_j = 0$ for $j \in \{1, \ldots, m\} \setminus \mathcal{J}(\hat{x})$. □

Theorem 3.22 (Motzkin's theorem of the alternative). *Let B, C, D be $p \times n$, $l \times n$ and $o \times n$ matrices, respectively. Then either*

$$Bx < 0,\ Cx \leqq 0,\ Dx = 0$$

has a solution $x \in \mathbb{R}^n$ or

$$B^T y^1 + C^T y^2 + D^T y^3 = 0,\ y^1 \geq 0,\ y^2 \geqq 0 \tag{3.56}$$

has a solution $y^1 \in \mathbb{R}^p,\ y^2 \in \mathbb{R}^l,\ y^3 \in \mathbb{R}^o$, but never both.

A proof of Theorem 3.22 can be found in (Mangasarian, 1969, p.28).

For convex functions, we also have a sufficient condition for weakly efficient solutions.

Corollary 3.23. *Under the assumptions of Theorem 3.21 and the additional assumption that all functions f_k and g_j are convex (3.49) – (3.52) with $\hat{\lambda} \geq 0$ and $\hat{\mu} \geqq 0$ in Theorem 3.21 are sufficient for $\hat{x}$ to be weakly efficient.*

Proof. By the second part of Theorem 3.20 and Theorem 3.21, (3.49) – (3.52) imply that $\hat{x}$ is an optimal solution of the single objective optimization problem

$$\min_{x\in\mathcal{X}} \sum_{k=1}^{p} \hat{\lambda}_k f_k(x).$$

Since $\hat{\lambda} \in \mathbb{R}_{\geq}$ this implies that $\hat{x} \in \mathcal{X}_{wE}$ according to the first statement of Proposition 3.9. □

Next, we prove similar conditions for properly efficient solutions in Kuhn and Tucker's sense and in Geoffrion's sense.

Kuhn and Tucker's definition of proper efficiency (Definition 2.49) is based on the system of inequalities (3.57) – (3.59)

$$\langle \nabla f_k(\hat{x}), d\rangle \leq 0 \quad \forall\ k = 1, \ldots, p \tag{3.57}$$

$$\langle \nabla f_i(\hat{x}), d\rangle < 0 \quad \text{for some } i \in \{1, \ldots, p\} \tag{3.58}$$

$$\langle \nabla g_j(\hat{x}), d\rangle \leq 0 \quad \forall\ j \in \mathcal{J}(\hat{x}) = \{j = 1, \ldots, m :\ g_j(\hat{x}) = 0\} \tag{3.59}$$

having no solution. We apply Tucker's theorem of the alternative, given below, to show that a dual system of inequalities then has a solution. This system yields a necessary condition for proper efficiency in Kuhn and Tucker's sense.

Theorem 3.24 (Tucker's theorem of the alternative). *Let B, C and D be $p \times n$, $l \times n$ and $o \times n$ matrices. Then either*

$$Bx \leq 0, \ Cx \leqq 0, \ Dx = 0 \tag{3.60}$$

has a solution $x \in \mathbb{R}^n$ or

$$B^T y^1 + C^T y^2 + D^T y^3 = 0, \ y^1 > 0, \ y^2 \geqq 0 \tag{3.61}$$

has a solution $y^1 \in \mathbb{R}^p, \ y^2 \in \mathbb{R}^l, \ y^3 \in \mathbb{R}^o$, but never both.

A proof of Theorem 3.24 can be found in Mangasarian (1969, p.29).

Theorem 3.25. *If $\hat{x}$ is properly efficient in Kuhn and Tucker's sense there exist $\hat{\lambda} \in \mathbb{R}^p, \ \hat{\mu} \in \mathbb{R}^m$ such that*

$$\sum_{k=1}^{p} \hat{\lambda}_k \nabla f_k(\hat{x}) + \sum_{j=1}^{m} \hat{\mu}_j \nabla g_j(\hat{x}) = 0 \tag{3.62}$$

$$\sum_{j=1}^{m} \hat{\mu}_j g_j(\hat{x}) = 0 \tag{3.63}$$

$$\hat{\lambda} > 0 \tag{3.64}$$

$$\hat{\mu} \geqq 0. \tag{3.65}$$

Proof. Because $\hat{x}$ is properly efficient in Kuhn and Tucker's sense there is no $d \in \mathbb{R}^n$ satisfying (3.57) – (3.59).

We apply Theorem 3.24 to the matrices

$$\begin{aligned} B &= (\nabla f_k(\hat{x}))_{k=1,\ldots,p} \\ C &= (\nabla g_j(\hat{x}))_{j \in \mathcal{J}(\hat{x})} \\ D &= 0 \end{aligned}$$

with $l = |J(\hat{x})|$. Since (3.57) – (3.59) do not have a solution $d \in \mathbb{R}^n$ we obtain $y^1 =: \hat{\lambda}$, $y^2 =: \hat{\mu}$ and y^3 with $\hat{\lambda}_k > 0$ for $k = 1, \ldots, p$, $\hat{\mu}_j \geq 0$ for $j \in \mathcal{J}(\hat{x})$ satisfying

$$\sum_{k=1}^{p} \hat{\lambda}_k \nabla f_k(\hat{x}) + \sum_{j \in \mathcal{J}(\hat{x})} \hat{\mu}_j \nabla g_j(\hat{x}) = 0. \tag{3.66}$$

Letting $\hat{\mu}_j = 0$ for all $j \in \{1, \ldots, m\} \setminus \mathcal{J}(\hat{x})$, the proof is complete. □

With Theorem 3.25 providing necessary conditions for Kuhn-Tucker proper efficiency and Theorem 2.51, which shows that Geoffrion's proper efficiency implies Kuhn and Tucker's under the constraint qualification we obtain Corollary 3.26 as an immediate consequence.

Corollary 3.26. *If $\hat{x}$ is properly efficient in Geoffrion's sense and the KT constraint qualification is satisfied at $\hat{x}$ then* (3.62) – (3.65) *are satisfied.*

For the missing link in the relationships of proper efficiency definitions we use the single objective Karush-Kuhn-Tucker sufficient optimality conditions of Theorem 3.20 and apply them to the weighted sum problem. We obtain the following theorem.

Theorem 3.27. *Assume that $f_k, g_j : \mathbb{R}^n \to \mathbb{R}$ are convex, continuously differentiable functions. Suppose that there are $\hat{x} \in \mathcal{X}$, $\hat{\lambda} \in \mathbb{R}^p$ and $\hat{\mu} \in \mathbb{R}^m$ satisfying* (3.62) – (3.65). *Then $\hat{x}$ is properly efficient in the sense of Geoffrion.*

Proof. Let $f(x) := \sum_{k=1}^{p} \hat{\lambda}_k \nabla f_k(x)$, which is a convex function. By the second part of Theorem 3.20 $\hat{x}$ is an optimal solution of $\min_{x \in \mathcal{X}} \sum_{k=1}^{p} \hat{\lambda}_k f_k(x)$. Since $\hat{\lambda}_k > 0$ for $k = 1, \ldots, p$ Theorem 3.15 yields that $\hat{x}$ is properly efficient in the sense of Geoffrion. □

We can derive two corollaries, the first one shows that for convex problems proper efficiency in Kuhn and Tucker's sense implies proper efficiency in Geoffrion's sense.

Corollary 3.28. *(See Theorem 2.52) Let $f_k, g_j : \mathbb{R}^n \to \mathbb{R}$ be convex, continuously differentiable functions and suppose $\hat{x}$ is properly efficient in Kuhn and Tucker's sense. Then $\hat{x}$ is properly efficient in Geoffrion's sense.*

Proof. The result follows from Theorem 3.25 and Theorem 3.27. □

The second corollary provides sufficient conditions for proper efficiency in Kuhn and Tucker's sense. It follows immediately from Theorems 3.27 and 2.51.

Corollary 3.29. *If, in addition to the assumptions of Theorem 3.27 the KT constraint qualification is satisfied at $\hat{x}$,* (3.62) – (3.65) *are sufficient for $\hat{x}$ to be properly efficient in Kuhn and Tucker's sense.*

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

Example 3.30 (Geoffrion (1968)). In the following problem, $\hat{x} = 0$ is properly efficient according to Kuhn and Tucker's definition, but not according to Geoffrion's definition. Consider

$$\min f(x) = (f_1(x), f_2(x)) = (-x^2, x^3)$$
$$\text{subject to } x \in \mathcal{X} = \{x \in \mathbb{R} : \ x \geq 0\}.$$

Figure 3.8 shows graphs of the objective functions and the feasible set in objective space, $\mathcal{Y} = f(\mathcal{X})$ as graph of $y_2(y_1) = (-y_1)^{\frac{3}{2}}$. The only constraint is given by $g(x) = -x \leq 0$.

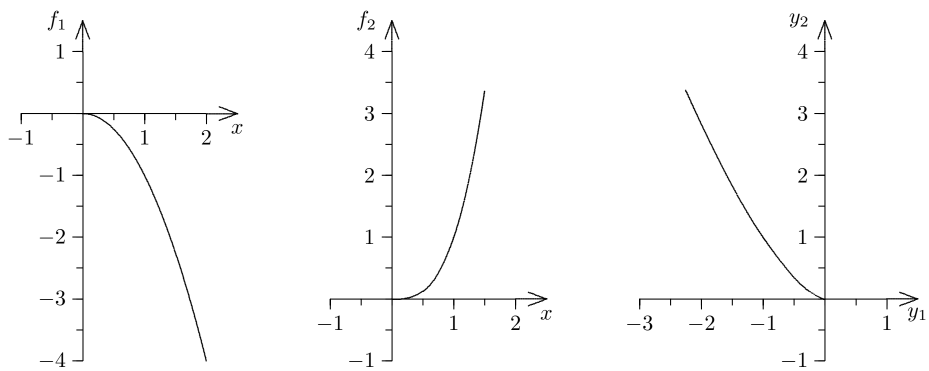

Fig. 3.8. Objective functions in Example 3.30.

To see that Definition 2.49 is satisfied compute

$$\nabla f(x) = (-2x, 3x^2) \qquad \nabla f_1(\hat{x}) = (0, 0)$$
$$\nabla g(x) = -1 \qquad \nabla g(\hat{x}) = -1$$

and choose $\hat{\lambda} = (1, 1)$, $\hat{\mu} = 0$ which satisfies (3.62) – (3.65).

To see that Definition 2.39 is not satisfied, let $\varepsilon > 0$ and compute the trade-off

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

□

The reader is asked to come up with an example, where a feasible solution $\hat{x}$ is properly efficient in Geoffrion's sense, but not in Kuhn and Tucker's sense, see Exercise 3.5.

We have shown necessary and sufficient conditions for weakly and strictly efficient solutions. Why are there none for efficient solutions? The answer is that the ones that can be proved are included in the above results. Observe that, because $\mathcal{X}_E \subset \mathcal{X}_{wE}$, the necessary condition of Theorem 3.21 holds for efficient solutions, too. On the other hand, because $\mathcal{S}(\mathcal{Y}) = \mathcal{X}_{pE} \subset \mathcal{X}_E$ for convex problems, the sufficient condition of Theorem 3.27 are sufficient for $\hat{x}$ to be efficient, too. Note that the essential difference between the conditions for weak and proper efficiency is $\hat{\lambda} \geq 0$ versus $\hat{\lambda} > 0$. We can therefore not expect any further results of this type for efficient solutions. This is pretty much the same situation we have encountered in Sections 3.1 and 3.2, where for convex problems we have been able to characterize weakly nondominated and properly nondominated points through weighted sum scalarization with $\lambda \geq 0$ and $\lambda > 0$, respectively.

3.4 Connectedness of Efficient and Nondominated Sets

We have discussed existence of nondominated points and efficient solutions and we have seen how the different concepts of efficiency relate to weighted sum scalarization. In this section, we use scalarizations to prove a topological property of the efficient and nondominated sets, connectedness. Connectedness is an important property, when it comes to determining these sets. If $\mathcal{Y}_N$ or $\mathcal{X}_E$ is connected, the whole nondominated or efficient set can possibly be explored starting from a single nondominated/efficient point using local search ideas. Connectedness will also make the task of selecting a final compromise solution from among the set of efficient solutions $\mathcal{X}_E$ easier, as there are no "gaps" in the efficient set.

In Figure 3.9 two sets $\mathcal{Y}$ are shown, one of which has a connected nondominated set and one of which has not.

Apparently, connectedness cannot be expected, when $\mathcal{Y}$ is not $\mathbb{R}^p_{\geqq}$-convex.

Definition 3.31. *A set* $\mathcal{S} \subset \mathbb{R}^p$ *is called* not connected *if it can be written as* $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$*, with* $\mathcal{S}_1, \mathcal{S}_2 \neq \emptyset$*,* $\operatorname{cl}\mathcal{S}_1 \cap \mathcal{S}_2 = \mathcal{S}_1 \cap \operatorname{cl}\mathcal{S}_2 = \emptyset$*. Equivalently,* $\mathcal{S}$ *is* not connected *if there exist open sets* $\mathcal{O}_1, \mathcal{O}_2$ *such that* $\mathcal{S} \subset \mathcal{O}_1 \cup \mathcal{O}_2$*,* $\mathcal{S} \cap \mathcal{O}_1 \neq \emptyset$*,* $\mathcal{S} \cap \mathcal{O}_2 \neq \emptyset$*,* $\mathcal{S} \cap \mathcal{O}_1 \cap \mathcal{O}_2 = \emptyset$*. Otherwise,* $\mathcal{S}$ *is called* connected.

In the proofs of the following theorems, we use some facts about connected sets which we state without proof here.

Lemma 3.32. *1. If* $\mathcal{S}$ *is connected and* $\mathcal{S} \subset \mathcal{U} \subset \operatorname{cl}\mathcal{S}$ *then* $\mathcal{U}$ *is connected.*
2. If $\{\mathcal{S}_i : i \in \mathcal{I}\}$ *is a family of connected sets with* $\cap_{i\in\mathcal{I}}\mathcal{S}_i \neq \emptyset$ *then* $\cup_{i\in\mathcal{I}}\mathcal{S}_i$ *is connected.*

We derive a preliminary result, considering $\mathcal{S}(\lambda, \mathcal{Y})$ and $\mathcal{S}(\mathcal{Y})$. From Theorem 3.17 we know $\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_N \subset \operatorname{cl}\mathcal{S}(\mathcal{Y})$ for $\mathbb{R}^p_{\geqq}$-convex sets $\mathcal{Y}$. We prove

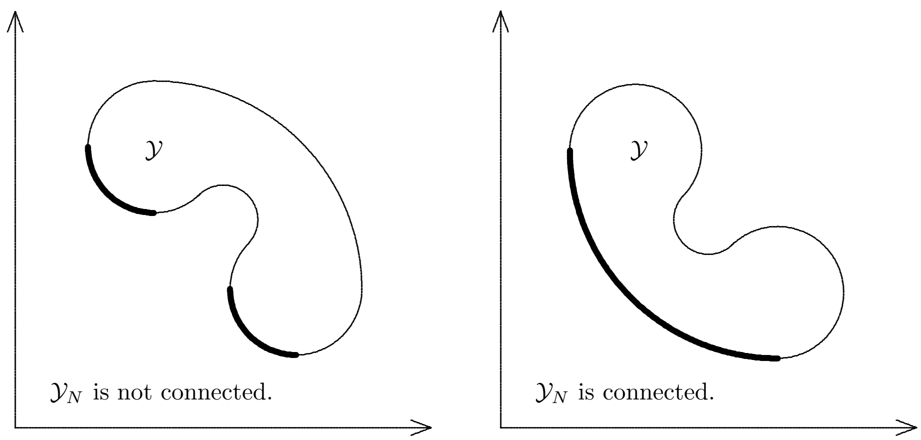

Fig. 3.9. Connectedness of $\mathcal{Y}_N$.

connectedness of $\mathcal{S}(\mathcal{Y})$ in the case that $\mathcal{Y}$ is compact, which implies the connectedness of $\mathcal{Y}_N$ with Lemma 3.32.

Proposition 3.33. *If $\mathcal{Y}$ is compact and convex then $\mathcal{S}(\mathcal{Y})$ is connected.*

Proof. Suppose $\mathcal{S}(\mathcal{Y})$ is not connected. Then we have open sets $\mathcal{Y}_1, \mathcal{Y}_2$ such that $\mathcal{Y}_i \cap \mathcal{S}(\mathcal{Y}) \neq \emptyset$ for $i = 1, 2$, $\mathcal{Y}_1 \cap \mathcal{Y}_2 \cap \mathcal{S}(\mathcal{Y}) = \emptyset$, and $\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_1 \cup \mathcal{Y}_2$. Let

$$\mathcal{L}_i := \{\lambda \in \mathbb{R}^p_> : \mathcal{S}(\lambda, \mathcal{Y}) \cap \mathcal{Y}_i \neq \emptyset\},\ i = 1, 2. \tag{3.67}$$

Because $\mathcal{S}(\lambda, \mathcal{Y})$ is convex and every convex set is connected, we know that $\mathcal{S}(\lambda, \mathcal{Y})$ is connected. Therefore

$$\mathcal{L}_i = \{\lambda \in \mathbb{R}^p_> : \mathcal{S}(\lambda, \mathcal{Y}) \subset \mathcal{Y}_i\},\ i = 1, 2 \tag{3.68}$$

and $\mathcal{L}_1 \cap \mathcal{L}_2 = \emptyset$. But since $\mathcal{Y}_i \cap \mathcal{S}(\mathcal{Y}) \neq \emptyset$ we also have $\mathcal{L}_i \cap \mathbb{R}^p_> \neq \emptyset$ for $i = 1, 2$. From $\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_1 \cup \mathcal{Y}_2$ it follows that $\mathbb{R}^p_> \subset \mathcal{L}_1 \cup \mathcal{L}_2$ (in fact, these sets are equal). By Lemma 3.34 below the sets $\mathcal{L}_i$ are open, which implies the absurd statement that $\mathbb{R}^p_\geqq$ is not connected. □

Lemma 3.34. *The sets $\mathcal{L}_i = \{\lambda \in \mathbb{R}^p_> :\ \mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_i\}$ in the proof of Proposition 3.33 are open.*

Proof. We will show the Lemma for $\mathcal{L}_1$, which by symmetry is enough. If $\mathcal{L}_1$ is not open there must be $\hat{\lambda} \in \mathcal{L}_1$ and $\{\lambda^k, k \geq 1\} \subset \mathbb{R}^p_> \setminus \mathcal{L}_1 = \mathcal{L}_2$ such that $\lambda^k \to \hat{\lambda}$.

Let $y^k \in \mathcal{S}(\lambda^k, \mathcal{Y}),\ k \geq 1$. Since $\mathcal{Y}$ is compact, we can assume (taking a subsequence if necessary) that $y^k \to \hat{y} \in \mathcal{Y}$ and $\hat{y} \in \mathcal{S}(\hat{\lambda}, \mathcal{Y})$. Note that otherwise there would be $y' \in \mathcal{Y}$ such that $\langle \hat{\lambda}, y' \rangle < \langle \hat{\lambda}, \hat{y} \rangle$ and by continuity of the scalar product, we would have $\langle \lambda^k, y' \rangle < \langle \lambda^k, y^k \rangle$ for sufficiently large k, contradicting $y^k \in \mathcal{S}(\lambda^k, \mathcal{Y})$.)

Now we have $y^k \in \mathcal{S}(\lambda^k, \mathcal{Y}) \subset (\mathcal{Y}_2 \cap \mathcal{S}(\mathcal{Y}))$ and $\mathcal{Y}_1 \cap \mathcal{Y}_2 \cap \mathcal{S}(\mathcal{Y}) = \emptyset$ so $y^k \in \mathcal{Y}_1^c$ for each $k \geq 1$. Since $\mathcal{Y}_1^c$ is closed, $\hat{y} = \lim y^k \in \mathcal{Y}_1^c$, i.e. $\hat{y} \notin \mathcal{Y}_1$ contradicting $\hat{\lambda} \in \mathcal{L}_1$. □

Theorem 3.35 (Naccache (1978)). *If $\mathcal{Y}$ is closed, convex, and $\mathbb{R}^p_{\geqq}$-compact then $\mathcal{Y}_N$ is connected.*

Proof. We will first construct compact and convex sets $\mathcal{Y}(\alpha), \alpha \in \mathbb{R}$, for which Proposition 3.33 is applicable. We apply Theorem 3.17 to get that $\mathcal{Y}(\alpha)_N \subset \operatorname{cl} \mathcal{S}(\mathcal{Y}(\alpha))$ and apply Lemma 3.32 to see that sets $\mathcal{Y}(\alpha)_N$ are connected. It is then easy to derive the claim of the theorem by showing $\mathcal{Y}_N = \cup_{\alpha \geq \hat{\alpha}} \mathcal{Y}(\alpha)_N$ for some $\hat{\alpha}$ with$\cap_{\alpha \geq \hat{\alpha}} \mathcal{Y}(\alpha)_N \neq \emptyset$ and applying Lemma 3.32 again.

To construct $\mathcal{Y}(\alpha)$ choose $d \in \mathbb{R}^p_{>}$ and define $y(\alpha) = \alpha d$, $\alpha \in \mathbb{R}$. We claim that for all $y \in \mathbb{R}^p$ there is a real number $\alpha > 0$ such that $y \in y(\alpha) - \mathbb{R}^p$ (see Figure 3.10).

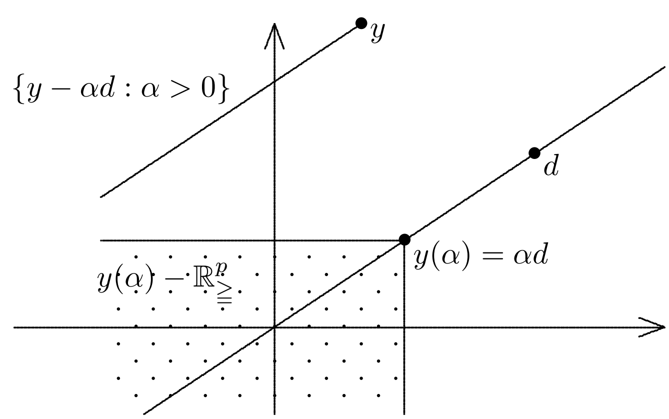

Fig. 3.10. The claim in the proof of Theorem 3.35.

To see this, observe that if it were not true there would be no $d' \in \mathbb{R}^p_{\geqq}$ such that $y = \alpha d - d'$, or $y - \alpha d = -d'$. Thus, we would have two nonempty convex sets $\{y - \alpha d : \ \alpha > 0\}$ and $-\mathbb{R}^p_{\geqq}$ which can be separated according to Theorem 3.2. Doing so provides some $y^* \in \mathbb{R}^p \setminus \{0\}$ with

$$\langle y^*, y - \alpha d \rangle \geq 0 \text{ for all } \alpha > 0, \tag{3.69}$$

$$\langle y^*, -d' \rangle \leq 0 \text{ for all } d' \in \mathbb{R}^p_{\geqq}. \tag{3.70}$$

Hence $\langle y^*, d' \rangle \geq 0$ for all $d' \in \mathbb{R}^p_{\geqq}$, in particular $\langle y^*, d \rangle > 0$ because $d \in \mathbb{R}^p_{>}$. But then $\langle \lambda, y - \alpha d \rangle < 0$ for α sufficiently large, a contradiction to (3.69).

With the claim proved, we can choose $y \in \mathcal{Y}_N$ and appropriate $\hat{\alpha} > 0$ such that $y \in y(\hat{\alpha}) - \mathbb{R}^p_{\geqq}$, which means that $(y(\hat{\alpha}) - \mathbb{R}^p_{\geqq}) \cap \mathcal{Y}_N \neq \emptyset$. We define

$$\mathcal{Y}(\alpha) := \left[\left(y(\alpha) - \mathbb{R}^p_{\geqq} \right) \cap \mathcal{Y} \right]. \tag{3.71}$$

With this notation, the claim above implies in particular that

$$\mathcal{Y}_N = \bigcup_{\alpha \geq \hat{\alpha}} Y(\alpha)_N. \tag{3.72}$$

Because $\mathcal{Y}(\alpha)$ is convex and compact ($\mathcal{Y}$ is $\mathbb{R}^p_{\geqq}$-compact) we can apply Theorem 3.17 to get

$$\mathcal{S}(\mathcal{Y}(\alpha)) \subset \mathcal{Y}(\alpha)_N \subset \mathcal{Y}\alpha_{pN}.$$

Thus, Proposition 3.33 and the first part of Lemma 3.32 imply that $\mathcal{Y}(\alpha)_N$ is connected.

Observing that $\mathcal{Y}(\alpha)_N \supset \mathcal{Y}(\hat{\alpha})_N$ for $\alpha > \hat{\alpha}$, i.e. $\cap_{\alpha \geq \hat{\alpha}} \mathcal{Y}(\alpha)_N = \mathcal{Y}(\hat{\alpha})_N \neq \emptyset$ we have expressed $\mathcal{Y}_N$ as a union of a family of connected sets with nonempty intersection (see Figure 3.11. The second part of Lemma 3.32 proves that $\mathcal{Y}_N$ is connected. □

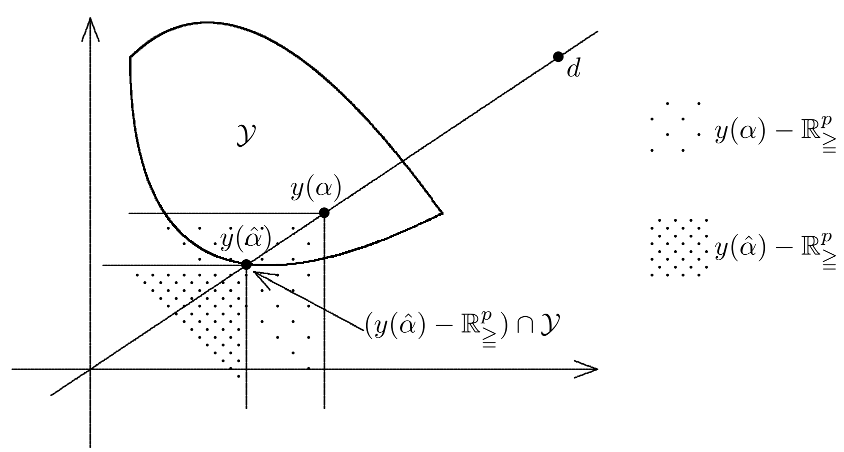

Fig. 3.11. Finding the sets $E(\alpha)$ in the proof of Theorem 3.35.

With Theorem 3.35 we have a criterion for connectedness in the objective space. What about the decision space? If we assume convexity of f, it is possible to show that $\mathcal{X}_{wE}$ is connected. Let $\mathcal{X} \subset \mathbb{R}^n$ be convex and compact and $f_k : \mathbb{R}^n \to \mathbb{R}$ be convex. We will use Theorem 3.5 ($\mathcal{Y}_{wN} = \mathcal{S}_(\mathcal{Y})$) and the following fact:

Lemma 3.36. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex on the closed convex $\mathcal{X}$. Then the set $\{\hat{x} \in \mathcal{X} : f(\hat{x}) = \inf_{x \in \mathcal{X}} f(x)\}$ is closed and convex.*

We also need a theorem providing a result on connectedness of preimages of sets, taken from Warburton (1983), where a proof can be found.

Theorem 3.37. *Let $\mathcal{V} \subset \mathbb{R}^n, \mathcal{W} \subset \mathbb{R}^p$, and assume that $\mathcal{V}$ is compact and $\mathcal{W}$ is connected. Furthermore, let $g : \mathcal{V} \times \mathcal{W} \to \mathbb{R}$ be continuous. Denote by $\mathcal{X}(w) = \operatorname{argmin}\{g(v, w) : \ v \in \mathcal{V}\}$. If $\mathcal{X}(w)$ is connected for all $w \in \mathcal{W}$ then $\cup_{w \in \mathcal{W}} \mathcal{X}(w)$ is connected.*

Theorem 3.38. *Let $\mathcal{X}$ be a compact convex set and assume that $f_k : \mathbb{R}^n \to \mathbb{R}, k = 1, \ldots, p$ are convex. Then $\mathcal{X}_{wE}$ is connected.*

Proof. Since the objective functions f_k are continuous and $\mathcal{X}$ is compact, $\mathcal{Y} = f(\mathcal{X})$ is compact. Using Theorem 3.5 we have $\mathcal{Y}_{wN} = \mathcal{S}_(\mathcal{Y})$. In terms of f and $\mathcal{X}$ this means

$$\begin{aligned} \mathcal{X}_{wP} &= \bigcup_{\lambda \in \mathbb{R}_\geq} \{\hat{x} : \textstyle\sum_{k=1}^p \lambda_k f_k(\hat{x}) \leq \sum_{k=1}^p \lambda_k f_k(x) \text{ for all } x \in \mathcal{X}\} \\ &=: \bigcup_{\lambda \in \mathbb{R}_\geq} \mathcal{X}(\lambda). \end{aligned} \tag{3.73}$$

Noting that $\langle f(\cdot), \cdot \rangle : \mathcal{X} \times \mathbb{R}_\geq \to \mathbb{R}$ is continuous, that $\mathbb{R}^p_\geq$ is connected, that $\mathcal{X}$ is compact, and that by Lemma 3.36 $\mathcal{X}(\lambda)$ is nonempty and convex (hence connected) we can apply Theorem 3.37 to get that $\mathcal{X}_{wE}$ is connected. □

We remark that the proof works in the same way to see that $\mathcal{X}_{pE}$ is connected under the same assumptions. This is true, because as in (3.73), we can write

$$\mathcal{X}_{pE} = \bigcup_{\lambda \in \mathbb{R}^p_>} \mathcal{X}(\lambda). \tag{3.74}$$

and as we observed, $\mathcal{X}(\lambda)$ is connected (convex), and of course $\mathbb{R}^p_>$ is connected.

To derive a connectedness result for $\mathcal{X}_E$ we need an additional Lemma.

Lemma 3.39. *Let $f : \mathcal{X} \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$ be a continuous function and let $\tilde{\mathcal{Y}} \subset \mathbb{R}^p$ be such that $f^{-1}(\operatorname{cl} \tilde{\mathcal{Y}}) \subset \mathcal{X}$. Then*

$$f^{-1}(\operatorname{cl} \tilde{\mathcal{Y}}) = \operatorname{cl}(f^{-1}(\tilde{\mathcal{Y}})). \tag{3.75}$$

Theorem 3.40. *Let $\mathcal{X} \subset \mathbb{R}^n$ be a convex and compact set. Assume that all objective functions f_k are convex. Then $\mathcal{X}_E$ is connected.*

Proof. Because $\mathcal{X}$ is compact and convex and f_k are convex and continuous, $\mathcal{Y} = f(\mathcal{X})$ is also compact and convex. Thus, from Theorem 3.17

$$\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_N \subset \operatorname{cl} \mathcal{S}(\mathcal{Y}). \tag{3.76}$$

Therefore, taking preimages and applying Theorem 3.13 and Corollary 3.12 ($\mathcal{Y}_{pN} = \mathcal{S}(\mathcal{Y})$) we get

$$\mathcal{X}_{pE} \subset \mathcal{X}_E \subset f^{-1}\left(\operatorname{cl} \mathcal{S}(\mathcal{Y})\right). \tag{3.77}$$

We apply Lemma 3.39 to $\tilde{\mathcal{Y}} = \mathcal{S}(\mathcal{Y})$ to get $f^{-1}(\operatorname{cl} \mathcal{S}(\mathcal{Y})) = \operatorname{cl}(f^{-1}(\mathcal{S}(\mathcal{Y}))) = \operatorname{cl} \mathcal{X}_{pE}$ and obtain

$$\mathcal{X}_{pE} \subset \mathcal{X}_E \subset \operatorname{cl} \mathcal{X}_{pE}. \tag{3.78}$$

The result now follows from Lemma 3.32. □

For once deriving results on $\mathcal{Y}$ from results on $\mathcal{X}$, we note the consequences of Theorem 3.38 and Theorem 3.40 for $\mathcal{Y}_{wN}, \mathcal{Y}_N$, and $\mathcal{Y}_{pE}$.

Corollary 3.41. *If $\mathcal{X}$ is a convex, compact set and $f_k : \mathbb{R}^n \to \mathbb{R}, k = 1, \ldots, p$ are convex functions then $\mathcal{Y}_{wN}$, $\mathcal{Y}_N$, and $\mathcal{Y}_{pN}$ are connected*

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

That a relaxation of convexity, namely quasi-convexity, is not sufficient to prove connectedness of $\mathcal{X}_E$ can be seen from Exercise 3.11

3.5 Notes

Equations (3.37) and (3.38) imply

$$\mathcal{Y}_{pE} \subset \mathcal{Y}_N \subset \operatorname{cl} \mathcal{Y}_{pN}$$

for $\mathbb{R}^p_{\geqq}$-convex and $\mathbb{R}^p_{\geqq}$-closed sets. Results of this type are called Arrow-Barankin-Blackwell theorems, after the first theorem of this type for closed convex sets, proved by Arrow *et al.* (1953). This has been generalized to orders defined by closed cones $\mathcal{C}$. Hirschberger (2002) shows that the convexity is not essential and the result remains true if *calY* is closed and $\mathcal{Y}_{pN} \neq \emptyset$.

The necessary and sufficient conditions for proper efficiency in Kuhn and Tucker's sense go back to Kuhn and Tucker (1951). Fritz-John type necessary conditions for efficiency have been proved in Da Cuhna and Polak (1967). All of the conditions we have mentioned here are first order conditions. There is of course also literature on second order necessary and sufficient conditions for efficiency. For this type of conditions it is usually assumed that the objective functions $f_k, k = 1, \ldots, p$ and the constraint functions $g_j, j = 1, \ldots, m$ of the MOP are twice continuously differentiable.

Several necessary and sufficient second-order conditions for the MOP are developed by Wang (1991). Cambini *et al.* (1997) establish second order conditions for MOPs with general convex cones while Cambini (1998) develops second order conditions for MOPs with the componentwise order. Aghezzaf (1999) and Aghezzaf and Hachimi (1999) develop second-order necessary conditions. Recent works include Bolintinéanu and El Maghri (1998), Bigi and Castellani (2000), Jimenez and Novo (2002).

There is some literature on the connectedness of nondominated sets. Bitran and Magnanti (1979) show $\mathcal{Y}_N$ and $\mathcal{Y}_{pN}$ are connected if $\mathcal{Y}$ is compact and convex. Luc (1989) proves connectedness results for $calY_{wN}$ if $\mathcal{Y}$ is $\mathcal{C}$-compact and convex. Danilidis *et al.* (1997) consider problems with three objectives, and Hirschberger (2002) shows that the convexity is not essential: if $calC$ and $\mathcal{Y}$ are closed, $\mathcal{Y}_N$ and $\mathcal{Y}_{pN}$ are connected. $\mathcal{Y}_{wN}$ is connected if in addition $\mathcal{Y}_N$ is nonempty.

Exercises

3.1. Prove that if $\mathcal{Y}$ is closed then $\operatorname{cl}\mathcal{S}(\mathcal{Y}) \subset \mathcal{S}_0(\mathcal{Y})$. Hint: Choose sequences λ_k, y^k such that $y^k \in \operatorname{Opt}(\lambda_k, \mathcal{Y})$ and show that $\lambda_k \to \hat{\lambda}$ and $y^k \to \hat{y}$ with $\hat{y} \in \operatorname{Opt}(\hat{\lambda}, \mathcal{Y}), \hat{\lambda} \geq 0$.

3.2. Prove Proposition 3.8, i.e. show that if $\hat{y}$ is the unique element of $\operatorname{Opt}(\lambda, \mathcal{Y})$ for some $\lambda \in \mathbb{R}^p_{\geq}$ then $\hat{y} \in \mathcal{Y}_N$.

3.3. Give one example of a set $\mathcal{Y} \in \mathbb{R}^2$ for each of the following situations:

1. $\mathcal{S}_0(\mathcal{Y}) \subset \mathcal{Y}_{wN}$ with strict inclusion.
2. $\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_N \subset \mathcal{S}_0(\mathcal{Y})$ with both inclusions strict,
3. $\mathcal{S}(\mathcal{Y}) \cup \mathcal{S}'_0(\mathcal{Y}) = \mathcal{Y} = \mathcal{S}_0(\mathcal{Y})$, where

$$\mathcal{S}'_0(\mathcal{Y}) = \left\{ y' \in \mathcal{Y} : y' \text{ is the unique element of } \operatorname{Opt}(\lambda, \mathcal{Y}), \lambda \in \mathbb{R}^p_{\geq} \right\}.$$

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

1. Show that $\hat{y} = (-1, 0)$ is properly nondominated in Benson's sense, i.e.

$$(\operatorname{cl}(\operatorname{cone}(\mathcal{Y} + \mathcal{C} - \hat{y}))) \cap (-\mathcal{C}) = \{0\}.$$

2. Show that $\hat{y} \in \operatorname{Opt}(\lambda, \mathcal{Y})$ for some $\lambda \notin \mathbb{R}^p_{>}$ and verify that this $\lambda \in \mathcal{C}^{s\circ}$, where

$$\mathcal{C}^{s\circ} = \{\mu \in \mathbb{R}^p : \langle \mu, d \rangle > 0 \text{ for all } d \in \mathcal{C}\}.$$

This result shows that proper nondominance is related to weighted sum scalarization with weighting vectors in $\mathcal{C}^{s\circ}$.

3.5 (Tamura and Arai (1982)). Let

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

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

3.6. Let $\mathcal{C} \subset \mathbb{R}^p$ be a cone. The polar cone $\mathcal{C}^\circ$ of $\mathcal{C}$ is defined as follows:

$$\mathcal{C}^\circ := \{y \in \mathbb{R}^p : \langle y, d \rangle \geq 0 \text{for all } d \in \mathcal{C} \setminus \{0\}\}.$$

Prove the following:

1. $\mathcal{C}^\circ$ is a closed convex cone containing 0.
2. $\mathcal{C} \subset (\mathcal{C}^\circ)^\circ =: \mathcal{C}^{\circ\circ}$.
3. $\mathcal{C}_1 \subset \mathcal{C}_2 \Rightarrow \mathcal{C}_2^\circ \subset \mathcal{C}_1^\circ$.
4. $\mathcal{C}^\circ = (\mathcal{C}^{\circ\circ})^\circ$.

3.7. This exercise is about comparing weighted sum scalarizations with weighting vectors from polar cones and $\mathcal{C}$-nondominance. Let $\mathcal{C}$ be a convex pointed cone and $\lambda \in \mathcal{C}^\circ$ and define

$$\mathrm{Opt}_{\mathcal{C}}(\lambda, \mathcal{Y}) := \left\{ \hat{y} \in \mathcal{Y} : \langle \lambda, \hat{y} \rangle = \min_{y \in \mathcal{Y}} \langle \lambda, y \rangle \right\}.$$

1. Show that
$$\mathcal{S}_{\mathcal{C}^\circ}(\mathcal{Y}) := \bigcup_{\lambda \in \mathcal{C}^\circ \setminus \{0\}} \mathrm{Opt}(\lambda, \mathcal{Y}) \subset \mathcal{Y}_{\mathcal{C}wN},$$
where $\hat{y} \in \mathcal{Y}_{\mathcal{C}wN}$ if $(\mathcal{Y} + \operatorname{int} \mathcal{C} - \hat{y}) \cap (-\operatorname{int} \mathcal{C}) = \emptyset$
2. Let $\mathcal{C}^{s\circ}$ be as in Exercise 3.6. Show
$$\mathcal{S}_{\mathcal{C}^{s\circ}}(\mathcal{Y}) := \bigcup_{\lambda \in \mathcal{C}^{s\circ}} \mathrm{Opt}(\lambda, \mathcal{Y}) \subset \mathcal{Y}_{\mathcal{C}N}.$$

Hint: Look at the proofs of Theorems 3.4 and 3.7, respectively.

3.8 (Wiecek (1995)). Consider the problem

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

Use the conditions of Theorem 3.25 to find at least one candidate for a properly efficient solution $\hat{x}$ (in the sense of Kuhn and Tucker). Try to determine all candidates.

3.9. Prove that $\hat{x} \in \mathcal{X}$ is efficient if and only if the optimal value of the optimization problem

$$\begin{aligned} \min \sum_{k=1}^{p} f_k(x) & \\ \text{subject to } f_k(x) &\le f_k(\hat{x}) \\ x &\in \mathcal{X} \end{aligned}$$

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

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.