J Glob Optim (2007) 37:325–355
DOI 10.1007/s10898-006-9052-x

ORIGINAL ARTICLE

Constrained optimization using multiple objective programming

Kathrin Klamroth · **Jørgen Tind**

Received: 30 May 2005 / Accepted: 1 June 2006 /
Published online: 21 July 2006

Abstract In practical applications of mathematical programming it is frequently observed that the decision maker prefers apparently suboptimal solutions. A natural explanation for this phenomenon is that the applied mathematical model was not sufficiently realistic and did not fully represent all the decision makers criteria and constraints. Since multicriteria optimization approaches are specifically designed to incorporate such complex preference structures, they gain more and more importance in application areas as, for example, engineering design and capital budgeting. The aim of this paper is to analyze optimization problems both from a constrained programming and a multicriteria programming perspective. It is shown that both formulations share important properties, and that many classical solution approaches have correspondences in the respective models. The analysis naturally leads to a discussion of the applicability of some recent approximation techniques for multicriteria programming problems for the approximation of optimal solutions and of Lagrange multipliers in convex constrained programming. Convergence results are proven for convex and nonconvex problems.

Keywords Constrained optimization · Multiple objective programming · Lagrange multipliers · Convergence

K. Klamroth (✉)
Institute of Applied Mathematics, University of Erlangen-Nuremberg, Martensstr. 3,
91058 Erlangen, Germany
e-mail: klamroth@am.uni-erlangen.de

J. Tind
Department of Applied Mathematics and Statistics, Institute for Mathematical Sciences,
University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø,
Denmark
e-mail: tind@math.ku.dk

1 Introduction

In many practical applications of constrained programming (CP) the constraints are based on estimated values like, for example, the amount of resources that will be available for a production process. A slight violation of one or several constraints is often acceptable if this results in a considerable improvement of the objective function. If, on the other hand, a lower consumption of some of the resources can be achieved, the overall utility of a solution may be improved even though this is not reflected in the objective function. Following this line of thought, some of the constraints of a CP may be interpreted as additional objective functions on which a minimum requirement was set, but violations as well as improvements are of interest to the decision maker. This is frequently the case in engineering design problems where, for example, the stiffness of a structure should be maximized while minimizing the amount of material used, and in environmental applications with soft constraints (Wierzbicki et al. 2000).

In the field of evolutionary algorithms, the difficulties in incorporating constraints into the fitness function of a genetic algorithm have lead to similar considerations. A recent review on the application of multicriteria approaches to handle constraints in genetic algorithms, including a numerical comparison of several such constraint-handling techniques, is contained in Mezura-Montes and Coello Coello (2002). In this context, Camponogara and Talukdar (1997) and similarly Osyczka et al. (2000) proposed a bicriterion method for solving CPs. The method transforms the given CP into a bicriteria problem such that one criterion equals the original objective function, and the other criterion is the sum of constraint violations. The resulting bicriteria problem can then be solved using (multicriteria) evolutionary methods that approach the nondominated set with a whole set of solutions (individuals) containing a good estimate of an optimal (and thus feasible) solution of the given CP among them. Jiménez et al. (2002) proposed an algorithm using Pareto dominance inside a preselection scheme to solve, among others, global optimization problems. In this approach, a given CP is reformulated as an unconstrained multiobjective optimization problem, and different priorities are assigned to the objective functions (feasible solutions with a good original objective value get the highest priority). Surry and Radcliffe (1997) used a combination of the vector evaluated genetic algorithm (VEGA) method and Pareto ranking to handle constraints in an approach called constrained optimization by multi-objective genetic algorithms (COMOGA). Individuals (corresponding to solutions of a given CP) are ranked depending on their sum of constraint violations, while fitness evaluations are based on (adaptively chosen) weightings of the two criteria "original objective" and "sum of constraint violations". Coello Coello (2000) suggested to use Pareto dominance selection, ranking feasible individuals higher than infeasible ones and assigning corresponding fitness values, to handle constraints in a genetic algorithm. Moreover, Coello Coello and Mezura-Montes (2002) developed a Niched-Pareto genetic algorithm (NPGA) to handle constraints in single-objective CPs, in which individuals are selected through a tournament based on Pareto dominance.

The close relationship between constrained programming and multicriteria optimization was also observed in other contexts. Among others, Wierzbicki (1977, 1980) introduced a scalarization method for multicriteria optimization problems that was motivated by penalty methods in constrained programming, Fletcher and Leyffer (2002) considered the objective function and the sum of constraint violations as two criteria and suggest a filter SQP method for nonlinear programming problems that uses dominance relations with respect to these two criteria, Carosi et al. (2003) discussed the connections between semidefinite optimization and vector optimization, and Boyd and Vandenberghe (2004) related Lagrangian relaxation to multicriteria optimization.

Motivated by these considerations, we can conclude that a given CP can be interrelated with one or several related multiobjective programming problems (MOPs), where some (or all) of the constraints of the CP are moved into the set of objective functions. It is easy to see that an optimal solution of a CP always is an efficient (or weakly efficient) solution of the related MOP. Consequently, methods designed for the determination of efficient solutions of MOPs can be adapted to approximate—or exactly determine—an optimal solution of the underlying CP. As a matter of fact, we will show that some of the classical approaches for the solution of CPs are directly related to scalarization approaches applied in MOP and vice versa. This connection gives rise to a new and surprisingly simple interpretation of some of the classical results in constrained programming. Moreover, it suggests the application of a whole range of solution and approximation methods for MOPs for the solution—or approximation—of CPs.

The remainder of the paper is organized as follows: after a formal statement of CPs and their associated MOPs in the following section, the relationship between well-known scalarization approaches for MOPs and problem relaxations for CPs are analysed in Sect. 3. Section 4 discusses an approximation approach for MOPs that appears suitable for the solution of CPs. Interrelations to Lagrangian relaxation are highlighted for the case of convex problems, and convergence results are proven for convex as well as non-convex problems. The paper is concluded with a short summary and some hints to further application areas in Sect. 5.

2 Problem formulation

The following notation will be used throughout the paper. Let $u, w \in \mathbb{R}^k$ be two vectors. We denote components of vectors by subscripts and enumerate vectors by superscripts. $u > w$ denotes $u_i > w_i$ for all $i = 1, \ldots, k$. $u \geq w$ denotes $u_i \geq w_i$ for all $i = 1, \ldots, k$, but $u \neq w$. $u \geqq w$ allows equality. The symbols $<, \leq, \leqq$ are used accordingly. Let $\mathbb{R}^k_{\leqq} := \{x \in \mathbb{R}^k : x \leqq 0\}$. The set $\mathbb{R}^k_{\geqq}$ is defined accordingly and the set $u + \mathbb{R}^k_{\geqq} := \{(u + x) \in \mathbb{R}^k : x \in \mathbb{R}^k_{\geqq}\}$, where $u \in \mathbb{R}^k$, is referred to as a dominating cone at u.

2.1 Constrained programming problems

In this and in the following sections we focus our discussion on inequality constrained programming problems. Nevertheless, most of the results immediately transfer to the general case of inequality and equality constrained problems. We consider the following constrained programming problem CP:

$$\begin{aligned} \max\ & f(x) \\ \text{s.t.}\ & g_i(x) \geq 0. \quad \forall i \in \{1, \ldots, k\}, \\ & x \in S, \end{aligned} \tag{1}$$

where $S \subseteq \mathbb{R}^n$ is the feasible set that may, for example, be given by further equality and/or inequality constraints and that may also include integrality constraints on some of the variables. We assume that all functions $f(x)$ and $g_i(x)$, $i = 1, \ldots, k$, are real-valued. For simplicity we will assume that an optimal solution exists for all problems considered. This assumption can of course be weakened depending on the particular problem formulation.

2.2 Multiple objective programming problems

Relaxing the constraints in the CP (1) and interpreting them as additional objective functions, we can formulate a related MOP as:

$$\begin{aligned} &\max\ f(x)\\ &\max\ g_1(x)\\ &\qquad\vdots\\ &\max\ g_k(x)\\ &\text{s.t.}\quad x \in S. \end{aligned} \tag{2}$$

We will refer to problem (2) as the MOP *associated to* the CP (1).

For notational convenience and to facilitate interchanges between objective functions and constraint functions, we denote a feasible criterion vector by $z(x) = \big(z_0(x), z_1(x), \ldots, z_k(x)\big)^T := \big(f(x), g_1(x), \ldots, g_k(x)\big)^T \in \mathbb{R}^{k+1}$. Using this notation, the *set of all feasible criterion vectors* Z, the *set of all (globally) nondominated criterion vectors* N and the *set of all efficient points* E of (2) are defined as follows:

$$\begin{aligned} Z &= \{z \in \mathbb{R}^{k+1} : z = z(x),\ x \in S\},\\ N &= \{z \in Z : \nexists \tilde{z} \in Z \text{ s.t. } \tilde{z} \geq z\},\\ E &= \{x \in S : z(x) \in N\}. \end{aligned}$$

We assume that the set Z is bounded and $\mathbb{R}^{k+1}_{\leqq}$-closed, i.e., the set $Z + \mathbb{R}^{k+1}_{\leqq}$ is closed, and that the sets N and E are nonempty. A point $\bar{x} \in S$ is called *weakly efficient* if there does not exist another point $\hat{x} \in S$ such that $z(\hat{x}) > z(\bar{x})$. The point $z^* \in \mathbb{R}^{k+1}$ with $z_i^* = \max\{z_i(x) : x \in S\}$, $i = 0, \ldots, k$ is called the *ideal criterion vector*, and the point $z^{**} \in \mathbb{R}^{k+1}$ with $z_i^{**} = z_i^* + \varepsilon_i$, $i = 0, \ldots, k$, where $\varepsilon = (\varepsilon_0, \ldots, \varepsilon_k) > 0$ are small positive numbers, is called the *utopian criterion vector*.

The set of properly nondominated solutions is defined according to Geoffrion (1968): a point $\bar{z} \in N$ is called *properly nondominated*, if there exists $M > 0$ such that for each $i = 0, \ldots, k$ and each $z \in Z$ satisfying $z_i > \bar{z}_i$ there exists an index $j \neq i$ with $z_j < \bar{z}_j$ and

$$\frac{\bar{z}_i - z_i}{z_j - \bar{z}_j} \leq M.$$

Otherwise $\bar{z} \in N$ is called *improperly nondominated*.

2.3 Interrelating CP and MOP

The following result is an immediate consequence of the application of the *e*-constraint approach to the associated MOP which will be described in detail in Sect. 3.1. A proof of this result can, for example, be found in Steuer (1986); it follows also from Theorem 2 below.

Theorem 1 *The set of optimal solutions of the CP* (1) *always contains an efficient solution of the associated MOP* (2)*, and all optimal solutions of* (1) *are weakly efficient for* (2).

Consequently, an optimal solution of the CP (1) can be determined as a specific efficient solution of the associated MOP (2). If the set of all efficient solutions of the associated MOP was known, that solution with the smallest value of f satisfying all the constraints in the other objectives could be easily identified. Clearly, this is in general not an efficient approach for the solution of CP since it involves the determination of a whole set of solutions rather than just one most preferred solution. However, if appropriate approximation approaches are used to approximate the efficient set of the associated MOP, near-optimal solutions of CP can be found in a very efficient way. How such approaches could be implemented will be discussed in Sect. 4. Moreover, as will be discussed in detail in the following section, Theorem 1 is the basis for an insightful comparison of different relaxations of a CP and corresponding scalarization approaches for the associated MOP.

3 Scalarization approaches for MOPs and their relation to constrained programming

A common approach for the solution of MOPs is to transform the original multicriteria problem into a series of scalarized, single criterion subproblems which are then solved using classical methods from constrained or unconstrained programming. Note that this is a principally different approach than the application of genetic algorithms for the approximation of the complete nondominated set of, for example, an associated MOP as mentioned in the introduction. However, most of the constraint handling techniques applied in the context of genetic algorithms can be interpreted as a combination of scalarization approaches for parts of the objectives (in this case the constraints which are typically combined in a weighted-sums scalarization, cf. Sect. 3.2) with an approximation of the nondominated set of the remaining bi or multiobjective, then unconstrained problem.

The aim of this section is to give a consistent review of the similarities and differences between scalarization approaches for MOPs and (sometimes only partial) relaxations and penalty approaches for CPs. For this purpose, the most commonly used scalarization approaches for MOPs are reviewed in the light of a given CP (1) and its associated MOP (2). As may be expected, very similar difficulties arise for both ways of interpreting the problem. Nevertheless, the comparison suggests also alternative relaxation strategies for CPs. One of them, which has its roots in the relation between weighted-sums scalarizations and Lagrangian relaxation (cf. Sect. 3.2), is described in detail in Sect. 4.

3.1 *e*-Constraint approach

The *e*-constraint approach reveals most directly the close relationship between CP and a scalarization of its associated MOP.

Let $i \in \{0, \ldots, k\}$ and $e_j \in \mathbb{R}, j \in J_i := \{0, \ldots, k\} \setminus \{i\}$. Then the ith *objective e-constraint program* introduced in Haimes et al. (1971) (see also Chankong and Haimes 1983), can be formulated as

$$\begin{aligned} \max\ & z_i(x) \\ \text{s.t.}\ & z_j(x) \geq e_j, \quad \forall j \in J_i, \\ & x \in S. \end{aligned} \tag{3}$$

We will assume in the following that the lower bounds e_j, $j \in J_i$ are always chosen such that (3) is feasible.

The following results on the *e*-constraint approach are well-known but included here for completeness.

Theorem 2 (Chankong and Haimes 1983)

(1) *Every optimal solution of* (3) *is a weakly efficient solution of* (2), *and the set of all optimal solutions of* (3) *contains at least one efficient solution of* (2).
(2) *If $\bar{x} \in S$ is an efficient solution of* (2), *then there exists an index $i \in \{0, \ldots, k\}$ and lower bounds $e_j \in \mathbb{R}$, $j \in J_i$ such that $\bar{x}$ is an optimal solution of* (3).

Theorem 2 can be strengthened in the case of unique optimal solutions:

Corollary 1 (Chankong and Haimes 1983) *If there exists an index i and lower bounds $e_j \in \mathbb{R}$, $j \in J_i$ such that $\bar{x}$ is the unique optimal solution of* (3), *then $\bar{x}$ is efficient for* (2).

Selecting $i = 0$ in (3), i.e., $z_i(x) = f(x)$, and $e_j = 0, j = 1, \ldots, k$, Theorem 1 immediately follows from Theorem 2. In other words, the *e*-constraint program (3) for $i = 0$ and $e = 0$ is nothing else than the original CP (1).

3.2 Weighted-sums approach

The weighted-sums approach was suggested by Gass and Saaty (1955) and is maybe the most commonly used scalarization technique for MOPs:

Let $\Lambda := \{\lambda \in \mathbb{R}^{k+1} : \lambda > 0, \sum_{i=0}^{k} \lambda_i = 1\}$ be the set of all strictly positive weighting vectors, and let $\Lambda_0 := \{\lambda \in \mathbb{R}^{k+1} : \lambda \geq 0, \sum_{i=0}^{k} \lambda_i = 1\}$ be the set of all non-negative weighting vectors. Then for a fixed $\bar{\lambda}$ in Λ or Λ_0, respectively, the *composite* or *weighted-sums program* corresponding to (2) is given by

$$\begin{aligned} \max\ & \bar{\lambda}^T z(x) \\ \text{s.t.}\ & x \in S. \end{aligned} \tag{4}$$

The following results on the weighted-sums scalarization are well-known:

Theorem 3 (see, e.g., Steuer 1986)

(1) *If $\bar{\lambda} \in \Lambda_0$, then an optimal solution $\bar{x}$ of* (4) *is weakly efficient for* (2), *and if $\bar{\lambda} \in \Lambda$, then an optimal solution $\bar{x}$ of* (4) *is efficient for* (2).
(2) *If Z is convex and if $\bar{x}$ is a properly efficient solution of* (2), *then there exists $\bar{\lambda} \in \Lambda$ such that $\bar{x}$ is optimal for* (4).

As was also observed, for example, in Boyd and Vandenberghe (2004), it is easy to see that for $\bar{\lambda} \in \Lambda_0$ with $\bar{\lambda}_0 \neq 0$, (4) is equivalent to the Lagrangian relaxation of the original CP (1) given by

$$\begin{aligned} \max\ & f(x) + \sum_{i=1}^{k} \tilde{\lambda}_i^T g_i(x) \\ \text{s.t.}\ & x \in S \end{aligned} \tag{5}$$

with Lagrange multipliers $\tilde{\lambda} \in \mathbb{R}^k_{\geqq}$. The corresponding transformation between weighting coefficients $\bar{\lambda}$ and Lagrange multipliers $\tilde{\lambda}$ is obtained by setting $\tilde{\lambda} := \left(\bar{\lambda}_1, \ldots, \bar{\lambda}_k\right)/\bar{\lambda}_0 \in \mathbb{R}^k_{\geqq}$ in case that $\bar{\lambda} \in \Lambda_0$ with $\bar{\lambda}_0 \neq 0$ is given, or, conversely, by setting $\bar{\lambda} := \left(1, \tilde{\lambda}_1, \ldots, \tilde{\lambda_k}\right)/||\left(1, \tilde{\lambda}_1, \ldots, \tilde{\lambda_k}\right)|| \in \mathbb{R}^{k+1}_{\geq}$ in case that $\tilde{\lambda} \in \mathbb{R}^k_{\geqq}$ is given. Consequently, the optimal solution of a Lagrangian relaxation of the CP (1) can alternatively be found as the optimal solution of the corresponding weighted-sums scalarization (4) of the MOP (2).

In case of convex problems we obtain the following necessary optimality condition for CP (see Rockafellar 1970):

Theorem 4 *Let $f, g_1, \ldots, g_k$ be convex functions and let S be a convex set. Let I denote the set of indices for which $g_i(x)$ are nonaffine functions, and assume that problem* (1) *has at least one feasible solution in* $\mathrm{ri}(S)$ *which satisfies with strict inequality all inequality constraints for $i \in I$, i.e. $\{x \in \mathrm{ri}(S)\colon\, g_i(x) > 0,\, i \in I \,\wedge\, g_i(x) \geq 0,\, i \notin I\} \neq \emptyset$. Furthermore, let $\bar{x}$ be an optimal solution of the CP* (1). *Then there exist weighting coefficients $\bar{\lambda} \in \Lambda_0$ with $\bar{\lambda}_0 \neq 0$ and Lagrange multipliers $\tilde{\lambda} \in \mathbb{R}^k_{\geqq}$, respectively, such that $\bar{x}$ is optimal for* (4) *and for* (5).

For any choice of weighting coefficients $\bar{\lambda} \geqq 0$ with $\bar{\lambda}_0 > 0$ and Lagrange multipliers $\tilde{\lambda} \geqq 0$ the optimal objective values of problems (4) and (5) are upper bounds on the optimal objective value of CP. Hence the *weighted sums dual* of CP

$$\min_{\bar{\lambda} \geqq 0,\, \bar{\lambda}_0 > 0} \; \max_{x \in S} \; \bar{\lambda}_0 f(x) + \sum_{i=1}^{k} \bar{\lambda}_i^T g_i(x) \tag{6}$$

is equivalent to the Lagrangean dual of CP

$$\min_{\tilde{\lambda} \geqq 0} \max_{x \in S} f(x) + \sum_{i=1}^{k} \tilde{\lambda}_i^T g_i(x). \tag{7}$$

In the convex case, Theorem 4 implies that both dual problems yield an optimal solution of CP.

3.3 Weighted Chebyshev approach and augmented weighted Chebyshev approach

Let $w \in \mathbb{R}^{k+1}_{\geq}$ be a set of nonnegative weights and let z^{**} be the utopian point of (2). Then the *weighted Chebyshev program* corresponding to (2), originally suggested by Bowman (1976) (see also Lin (2005), for a recent survey on this and on related methods) can be written as

$$\begin{array}{ll} \min & \|z^{**} - z(x)\|_\infty^w \\ \text{s.t.} & x \in S, \end{array}, \tag{8}$$

where $\|z^{**} - z(x)\|_\infty^w = \max_{i=0,\ldots,k} w_i(z_i^{**} - z_i(x))$ with weights $w_i > 0$, $i = 0, \ldots, k$, is the weighted Chebyshev distance between the utopian point z^{**} and the point $z(x) \in Z$. Note that due to the definition of the utopian point we have $z_i^{**} - z_i(x) > 0$ for all $x \in S$, $i = 0, \ldots, k$. The following results on the weighted Chebyshev scalarization are again well-known:

Theorem 5 (see, e.g., Ehrgott 2000)

(1) *Every optimal solution of* (8) *is weakly efficient for* (2)*, and the set of all optimal solutions of* (8) *contains at least one efficient solution of* (2)*. If the optimal solution of* (8) *is unique, then it is efficient for* (2).
(2) *If* $\bar{x}$ *is an efficient solution of* (2)*, then there exists* $w > 0$ *such that* $\bar{x}$ *is optimal for* (8).

Theorem 5 together with Theorem 1 immediately implies the following result interrelating the CP (1) with the weighted Chebyshev program (8):

Theorem 6 *If* $\bar{x}$ *is an optimal solution of the CP* (1)*, then there exists a weighting vector* $w \in \mathbb{R}^{k+1}_{>}$ *such that* $\bar{x}$ *is also an optimal solution of the weighted Chebyshev program* (8).

Note that different from the weighted-sums approach, no convexity assumptions are needed in Theorems 5 and 6.

The relationship between (1) and (8) can be better understood if (8) is rewritten as

$$\begin{aligned} \min\ & \alpha \\ \text{s.t.}\ & \alpha \geq w_i(z_i^{**} - z_i(x)), \quad \forall i \in \{0, \ldots, k\}. \\ & x \in S. \end{aligned} \tag{9}$$

Suppose that $(\bar{\alpha}, z(\bar{x}))$ is an optimal solution of (9). Then there is at least one index $i \in \{0, \ldots, k\}$ such that the constraint $\alpha \geq w_i(z_i^{**} - z_i(x))$ is binding at $(\bar{\alpha}, z(\bar{x}))$, and that $z_i(\bar{x})$ is the maximum possible value such that $\bar{\alpha} \geq w_j(z_j^{**} - z_j(\bar{x}))$ for all $j \in J_i = \{0, \ldots, k\} \setminus \{i\}$. Replacing α by $\bar{\alpha}$ in (9) and maximizing over $z_i(x)$ hence yields the following problem formulation that has the same optimal solution $\bar{x}$ as problem (9):

$$\begin{aligned} \max\ & z_i(x) \\ \text{s.t.}\ & \bar{\alpha} \geq w_j(z_j^{**} - z_j(x)), \quad \forall j \in J_i. \\ & x \in S. \end{aligned} \tag{10}$$

Problem (10) is easily recognized as an *e*-constraint program with objective $z_i(x)$ and right-hand side values $e_j := z_j^{**} - \frac{\bar{\alpha}}{w_j}$ for $j \in J_i$:

$$\begin{aligned} \max\ & z_i(x) \\ \text{s.t.}\ & z_j(x) \geq z_j^{**} - \frac{\bar{\alpha}}{w_j}, \quad \forall j \in J_i. \\ & x \in S. \end{aligned} \tag{11}$$

Moreover, if $(\bar{\alpha}, z(\bar{x}))$ is a unique optimal solution of (9), all of the constraints $\alpha \geq w_i(z_i^{**} - z_i(x))$ must be binding at $(\bar{\alpha}, z(\bar{x}))$, and for all $i \in \{0, \ldots, k\}$, $z_i(\bar{x})$ is the maximum possible value such that $\bar{\alpha} \geq w_j(z_j^{**} - z_j(\bar{x})), j \in J_i$. In this case an arbitrary index $i \in \{0, \ldots, k\}$ can be selected for the reformulations (10) and (11). Selecting $i = 0$ yields

$$\begin{aligned} \max\ & z_0(x) = f(x) \\ \text{s.t.}\ & z_j(x) \geq z_j^{**} - \frac{\bar{\alpha}}{w_j}, \quad \forall j \in \{1, \ldots, k\}, \\ & x \in S \end{aligned} \tag{12}$$

which is equivalent to the original CP (1) if $z_j^{**} - \frac{\bar{\alpha}}{w_j} = 0, j \in \{1, \ldots, k\}$, i.e., if $\frac{\bar{\alpha}}{w_j} = z_j^{**}$, $j \in \{1, \ldots, k\}$.

The reformulation (12) shows that the optimal solution $\bar{x}$ of the weighted Chebyshev problem (8) may yield an upper bound or a lower bound $z_0(\bar{x})$ on the optimal objective value of CP, depending on the choice of the weights w. Consequently, the formulation of a corresponding dual of CP is meaningless in this case.

Another drawback of the weighted Chebyshev approach (8) is that it may generate weakly efficient solutions. In order to overcome this difficulty, Steuer and Choo (1983) formulated an *augmented weighted Chebyshev program* as

$$\begin{array}{ll} \min & \|z^{**} - z(x)\|_\infty^w + \rho\|z^{**} - z(x)\|_1 \\ \text{s.t.} & x \in S, \end{array} \tag{13}$$

where $\|z^{**} - z(x)\|_1 = \sum_{i=0}^{k}(z_i^{**} - z_i(x))$ is the l_1 distance between the utopian point z^{**} and the point $z(x) \in Z$, and $\rho \geq 0$ is a (small) nonnegative scalar. As was shown in Steuer and Choo (1983), if $\rho > 0$ then every optimal solution of (13) is properly efficient for (2), and all efficient solutions of (2) can be found for appropriately selected values of w and ρ. Using the same arguments and notation as above (13) has the same optimal solution $\bar{x}$ as

$$\begin{array}{ll} \max & z_i(x) + \rho \sum_{j=0}^{k} z_j(x), \\ \text{s.t.} & z_j(x) \geq z_j^{**} - \dfrac{\bar{\alpha}}{w_j}, \quad \forall j \in J_i, \\ & x \in S, \end{array} \tag{14}$$

cf. (11). This analysis shows that the augmented weighted Chebyshev program (13) corresponds to a relaxation of the CP (1), where a Lagrangian penalty term is added to the objective function while the constraints are only partially relaxed. This becomes even more obvious in the case of unique optimal solutions and if $i = 0$ is chosen in (14) (cf. (12)).

3.4 Reference point approach

The reference point approach was introduced and discussed in Wierzbicki (1977, 1980, 1986), see also Stewart (1992) for a comparative discussion and Wierzbicki et al. (2000) for an application oriented analysis. It is based on the idea that, intuitively, decision makers may want to attain certain reference levels which can be used to define a reference point. Such a reference point may in fact be located anywhere in $\mathbb{R}^{k+1}$, i.e., it is not necessarily feasible nor infeasible for (2). The goal is then to minimize measures of under-achievement and to maximize over-achievement with respect to this reference point as far as possible in a balanced way.

Let $z^0 \in \mathbb{R}^{k+1}$ be such a *reference point* that could, for example, represent an aspiration level or desirable goals in the respective objectives. For example, in relation to the CP (1) a natural choice for $z_1^0, \ldots, z_k^0$ could be $z_i^0 = 0, i = 1, \ldots, k$. A continuous scalarizing function $s\colon \mathbb{R}^{k+1} \to \mathbb{R}$ is called an *order-representing achievement function* if it is strictly monotone in z, i.e., if $z^1 < z^2$ implies $s(z^1 - z^0) < s(z^2 - z^0)$, and if $\{z \in \mathbb{R}^{k+1} : s(z - z^0) > 0\} = z^0 + \mathbb{R}^{k+1}_{>}$ for any fixed $z^0 \in \mathbb{R}^{k+1}$. Analogously, s is called an *order-approximating achievement function* if, for any fixed $z^0 \in \mathbb{R}^{k+1}$, it is strongly monotone in z, i.e., if $z^1 \leq z^2$ implies $s(z^1 - z^0) < s(z^2 - z^0)$, and if for some small $\varepsilon > \bar{\varepsilon} \geq 0$, $z^0 + (\mathbb{R}^{k+1}_{\geqq})_{\bar{\varepsilon}} \subset \{z \in \mathbb{R}^{k+1} : s(z - z^0) \geq 0\} \subseteq z^0 + (\mathbb{R}^{k+1}_{\geqq})_{\varepsilon}$, where $(\mathbb{R}^{k+1}_{\geqq})_{\varepsilon}$ is an ε-conical neighborhood of $\mathbb{R}^{k+1}_{\geqq}$, i.e., $(\mathbb{R}^{k+1}_{\geqq})_{\varepsilon} = \{z \in \mathbb{R}^{k+1} : \operatorname{dist}(z, \mathbb{R}^{k+1}_{\geqq}) < \varepsilon||z||\}$.

As an example for an order-approximating achievement function, consider

$$s(z - z^0) = \left(\min_{i=0,\dots,k} (z_i - z_i^0) \right) + \alpha \left(\sum_{i=0}^{k} (z_i - z_i^0) \right) \tag{15}$$

with a scalar $\alpha > 0$ that is sufficiently small as compared to ε and large as compared to $\bar{\varepsilon}$. The *reference point approach* is based on the solution of

$$\begin{aligned} &\max \; s(z(x) - z^0) \\ &\text{s.t.} \quad x \in S \end{aligned} \tag{16}$$

with an arbitrary reference point $z^0 \in \mathbb{R}^{k+1}$. Note that an advantage of formulation (16) is that it always has a feasible solution if $S \neq \emptyset$. Moreover, all constraints $g_i(x) \geq 0$, $i = 1, \dots, k$ of the associated CP (1) are relaxed in this formulation. In case of linear problems, problem (16) with achievement function (15) can be rewritten as a linear programming problem. In general, however, (15) must be represented by additional constraints, or nondifferentiable optimization techniques have to be applied for the solution of (16).

Theorem 7 (Wierzbicki 1986)

(1) *If* $s \colon \mathbb{R}^{k+1} \to \mathbb{R}$ *is order-representing and if* $z(\bar{x})$ *is an optimal solution of* (16)*, then* $\bar{x}$ *is a weakly efficient solution of* (2).
If $s \colon \mathbb{R}^{k+1} \to \mathbb{R}$ *is order-approximating and if* $z(\bar{x})$ *is an optimal solution of* (16)*, then* $\bar{x}$ *is an efficient solution of* (2).

(2) *If* $\bar{x} \in S$ *is a weakly efficient solution of* (2) *and if* $s \colon \mathbb{R}^{k+1} \to \mathbb{R}$ *is order-representing, then the optimum of* (16) *with* $z^0 = z(\bar{x})$ *is attained at* $z(\bar{x})$ *and the optimal objective value is* 0.
If $\bar{x} \in S$ *is an* ε*-properly efficient solution of* (2) *and if* $s : \mathbb{R}^{k+1} \to \mathbb{R}$ *is order-approximating, then the optimum of* (16) *with* $z^0 = z(\bar{x})$ *is attained at* $z(\bar{x})$ *and the optimal objective value is* 0.

Combining this result with Theorem 1, we immediately obtain:

Theorem 8 *If* $\bar{x}$ *is an optimal solution of the CP* (1) *and if* $s \colon \mathbb{R}^{k+1} \to \mathbb{R}$ *is an order-representing achievement function, then there exists a reference point* $z^0 \in \mathbb{R}^{k+1}$ *such that* $\bar{x}$ *is also an optimal solution of the reference point approach* (16).

Note that again no convexity assumptions are needed in Theorems 7 and 8.

The relationship between (16) and penalty methods in constrained optimization where all constraints are relaxed and constraint violations are penalized in the objective function, including the weighted-sums approach (4) and, equivalently, Lagrangian relaxation (cf. Sect. 3.2), was already discussed in Wierzbicki (1977, 1980, 1986). Theorem 7 is based on the separation of sets which constitutes the close relationship to penalty methods, the concrete formulation of which depends on the concrete choice of an achievement function $s \colon \mathbb{R}^{k+1} \to \mathbb{R}$. For differentiable, monotone achievement functions, corresponding weighting coefficients at an optimal solution $z(\bar{x})$ of (16) can be computed as

$$\bar{\lambda} = \frac{\partial s(z(\bar{x}) - z^0)}{\partial z} \Bigg/ \left\| \frac{\partial s(z(\bar{x}) - z^0)}{\partial z} \right\|_1 .$$

In the case of convex problems, an optimal solution of (4) with this choice of weights can thus also be found as an optimal solution of (16) and vice versa.

If the utopian point is selected as reference point, i.e., if $z^0 = z^{**}$, then the objective function of the augmented weighted Chebyshev program (13) is order-approximating and hence the augmented weighted Chebyshev approach (cf. Sect. 3.3) can be interpreted as a special case of the reference point approach. In fact, (16) with achievement function (15) is in this case equivalent to (13).

Similarly, the direction method (cf. Sect. 3.5) is based on the selection of an appropriate starting point (or reference point) z^0, starting from which under- and over-achievement are optimized based on a predefined search direction.

3.5 Direction method

An early proposal of a direction method goes back to Boldur (1970) (see Roy 1971). He computes a steplength of a direction and a starting point based on a linear interpolation of two extreme positions of the criteria. Let $z^0 \in \mathbb{R}^{k+1}$ be a given starting point and consider a search direction $d \in \mathbb{R}^{k+1}$. Moreover, let $p \in \mathbb{R}^{k+1}_{\geqq}$ be given penalty parameters. We shall here consider the direction method as it is presented by Pascoletti and Serafini (1984) (for the case that $p = 0$), given by

$$\begin{aligned} \max\ & \alpha + \sum_{i=0}^{k} p_i q_i \\ \text{s.t.}\ & z(x) = z^0 + \alpha d + q \\ & q \in \mathbb{R}^{k+1}_{\geqq} \\ & x \in S. \end{aligned} \tag{17}$$

Note that, depending on the choice of z^0 and d, problem (17) may be infeasible or unbounded. Reasonable choices are to use $z^0 \in Z + \mathbb{R}^{k+1}_{\leqq}$ as a starting solution, which in fact requires the a priori knowledge of the feasible set Z, and search along a direction $d \in \mathbb{R}^{k+1} \setminus \mathbb{R}^{k+1}_{\leqq}$. If the penalty term in the objective function is omitted (i.e., if $p = 0$), Pascoletti and Serafini (1984) showed that for every optimal solution $(\bar{\alpha}, z(\bar{x}), \bar{q})$ of (17), $\bar{x}$ is a weakly efficient solution of (2) (and an efficient solution in case of uniqueness), and all efficient solutions of (2) can be generated by appropriate choices for $z^0 \in Z + \mathbb{R}^{k+1}_{\leqq}$ and $d \in \mathbb{R}^{k+1} \setminus \mathbb{R}^{k+1}_{\leqq}$ in (17). As discussed in Schandl et al. (2002b), the generation of weakly efficient solutions can be avoided if an appropriate penalty term is added in the objective function, which can, for example, be realized by selecting $p > 0$ in (17).

Theorem 9 (Pascoletti and Serafini 1984; Schandl et al. 2002b)

(1) *If* $(\bar{\alpha}, z(\bar{x}), \bar{q})$ *is an optimal solution of* (17) *with* $p \in \mathbb{R}^{k+1}_{>}$, *then* $\bar{x}$ *is an efficient solution of* (2). *If* $p \in \mathbb{R}^{k+1}_{\geqq}$, *then* $\bar{x}$ *is at least a weakly efficient solution of* (2).

(2) *If* $\bar{x} \in S$ *is an efficient solution of* (2), *then there exist penalty coefficients* $p \in \mathbb{R}^{k+1}_{\geqq}$, *a starting point* $z^0 \in Z + \mathbb{R}^{k+1}_{\leqq}$ *and a search direction* $d \in \mathbb{R}^{k+1} \setminus \mathbb{R}^{k+1}_{\leqq}$ *such that* $\bar{x}$ *is an optimal solution of* (17).

Note that $p = 0$ is always an appropriate choice in Theorem 9(2). Theorems 1 and 9 can again be used to derive the following result:

Theorem 10 *If $\bar{x}$ is an optimal solution of the CP* (1), *then there exists $p \in \mathbb{R}^{k+1}_{\geqq}$, a starting point $z^0 \in Z + \mathbb{R}^n_{\leqq}$ and a search direction $d \in \mathbb{R}^{k+1} \setminus \mathbb{R}^{k+1}_{\leqq}$ such that $\bar{x}$ is also an optimal solution of the direction method* (17).

Note that again no convexity assumptions are needed in Theorems 9 and 10.

To analyze the interrelation between (1) and (17), consider the case that (17) has an optimal solution $(\bar{\alpha}, z(\bar{x}), \bar{q})$ such that $\bar{x}$ is efficient for (2). Since $(\bar{\alpha}, z(\bar{x}), \bar{q})$ is feasible for (17) it satisfies $z_i(\bar{x}) = z_i^0 + \bar{\alpha} d_i + \bar{q}_i$ for all $i = 0, \ldots, k$. Moreover, the efficiency of $\bar{x}$ implies that $z_i(\bar{x})$ is maximal with the property that $z_j(\bar{x}) = z_j^0 + \bar{\alpha} d_j + \bar{q}_j$ for all $j \in J_i = \{0, \ldots, k\} \setminus \{i\}$. Similar to the case of the weighted Chebyshev approach this implies that if $\bar{\alpha}$ and $\bar{q}$ are fixed, $\bar{x}$ also solves the following problem:

$$\begin{aligned} \max\ & z_0(x) = f(x) \\ \text{s.t.}\ & z_j(x) = z_j^0 + \bar{\alpha} d_j + \bar{q}_j, \quad \forall j \in \{1, \ldots, k\} \\ & x \in S. \end{aligned} \tag{18}$$

This reformulation is equivalent to the original CP (1) if $z_j^0 - \bar{\alpha} d_j + \bar{q}_j = 0$ for all $j \in \{1, \ldots, k\}$, and also shows the similarity between the direction method and the weighted Chebyshev approach, cf. (12).

An alternative reformulation of (17) is obtained by using the equality $q = z(x) - z^0 - \alpha d$ in the objective function:

$$\begin{aligned} \max\ & \left(1 - \sum_{i=0}^{k} p_i d_i\right) \alpha \ + \ \sum_{i=0}^{k} p_i z_i(x) \ - \ \underbrace{\sum_{i=0}^{k} p_i z_i^0}_{\text{constant}} \\ \text{s.t.}\ & z(x) - z^0 - \alpha d \geqq 0. \\ & x \in S. \end{aligned} \tag{19}$$

In this reformulation, similarities are also visible to the weighted sums approach and hence to a modification of Lagrangian relaxation applied to the original CP (1).

In this context one may also consider the various interactive trade-off procedures that have been proposed. An early reference along this line is the work of Geoffrion et al. (1972). In these procedures a vector of trade-offs is constructed via an interaction with a decision maker or by marginal substitution rates determined by an implicitly given utility function. This defines a vector of preferred improvements which could be used as the direction d in the above framework.

3.6 Benson's method

Let $z^0 \in Z$ be a feasible starting point. *Benson's method* (Benson 1978) for the solution of (2) is based on the maximization of the sum of differences between objective values of feasible points and the starting point z^0:

$$\begin{aligned} \max\ & \sum_{i=0}^{k} \varepsilon_i \\ \text{s.t.}\ & z(x) = z^0 + \varepsilon \\ & \varepsilon \in \mathbb{R}^{k+1}_{\geqq} \\ & x \in S. \end{aligned} \tag{20}$$

Note that problem (20) can be viewed as a special case of the direction method (17) with $d = 0$ and $p = (1, \dots, 1) \in \mathbb{R}^{k+1}_{>}$. Therefore the following result is also a consequence of Theorem 9 above.

Theorem 11 (Benson 1978)

(1) *If* $(\bar{\varepsilon}, z(\bar{x}))$ *is an optimal solution of* (20)*, then* $\bar{x}$ *is efficient for* (2).
(2) *If* $\bar{x} \in S$ *is an efficient solution of* (2)*, then there exists a starting point* $z^0 \in Z$ *such that* $\bar{x}$ *is an optimal solution of* (20).

As observed in Ehrgott (2006, in press), setting $\varepsilon := z(x) - z^0$ in (20) yields the equivalent formulation

$$\begin{aligned} \max\ & \sum_{i=0}^{k} z_i(x) \\ \text{s.t.}\ & z_i(x) \geq z_i^0, \quad \forall i \in \{0, \dots, k\}, \\ & x \in S, \end{aligned} \tag{21}$$

highlighting the fact that (20) is basically a combination of a special weighted-sums scalarization and the *e*-constraint approach. In view of the CP (1) Benson's method hence corresponds to a modification of a specific Lagrangian relaxation where the constraints are kept, but with possibly modified right-hand side values. Due to these modified and varying constraints, problem (20) does not provide meaningful bounds for CP and hence dual formulations as for the weighted sums approach are not useful here.

Nevertheless, Theorem 1 implies that (20) can be used to formulate an optimality condition for CP in the following sense:

Theorem 12 *If* $\bar{x}$ *is an optimal solution of the CP* (1)*, then there exists a starting point* $z^0 \in Z$ *such that* $\bar{x}$ *is an optimal solution of* (20).

3.7 Method of elastic constraints

Let $i \in \{0, \dots, k\}$, $p \in \mathbb{R}^k$ and $e_j \in \mathbb{R}$, $j \in J_i := \{0, \dots, k\} \setminus \{i\}$. Then the *Method of Elastic Constraints* introduced by Ehrgott and Ryan (2003) can be formulated as

$$\begin{aligned} \max\ & z_i(x) - \sum_{j \in J_i} p_j s_{l_j} \\ \text{s.t.}\ & z_j(x) + s_{l_j} - s_{p_j} = e_j, \quad \forall j \in J_i \\ & s_l, s_p \in \mathbb{R}^k_{\geqq} \\ & x \in S \end{aligned} \tag{22}$$

with slack and surplus variables s_{l_j} and s_{p_j} associated with the bound e_j on objective j, $j \in J_i$.

Theorem 13 (Ehrgott and Ryan 2003)

(1) *If* $p > 0$ *and if* $(z(\bar{x}), \bar{s}_l, \bar{s}_p)$ *is an optimal solution of* (22)*, then* $\bar{x}$ *is an efficient solution of* (2).
(2) *If* $\bar{x} \in S$ *is an efficient solution of* (2)*, then for each* $i \in \{0, \dots, k\}$ *there exist lower bounds* $e_j \in \mathbb{R}$*,* $j \in J_i$ *and penalty coefficients* p^i_j*,* $j \in J_i$ *such that* $\bar{x}$ *is an optimal solution of* (22) *for all penalty vectors* $p \in \mathbb{R}^k$ *satisfying* $p \geqq p^i$.

As in the previous sections, Theorem 13 combined with Theorem 1 immediately implies the following:

Theorem 14 *If $\bar{x}$ is an optimal solution of the CP* (1)*, then there exist lower bounds $e_j \in \mathbb{R}$, $j \in J_i$ and penalty coefficients p^i_j, $j \in J_i$ such that $\bar{x}$ is also an optimal solution of the method of elastic constraints* (22).

The connection between (1) and (22) becomes clear by first noting that for $p > 0$, an optimal solution $(z(\bar{x}), \bar{s}_l, \bar{s}_p)$ of (22) satisfies $\bar{s}_{l_j} = \max\{0, e_j - z_j(\bar{x})\}$ and $\bar{s}_{p_j} = \max\{0, z_j(\bar{x}) - e_j\}$ for all $j \in J_i$. Selecting in addition $i = 0$ implies that problem (22) can be written equivalently as

$$\begin{aligned} \max\ & z_0(x) - \sum_{j=1}^{k} p_j \cdot \max\{0, e_j - z_j(x)\} \\ \text{s.t.}\ & x \in S, \end{aligned} \tag{23}$$

which, for $e_j = 0, j = 1, \ldots, k$, corresponds to an exact absolute value penalty function approach applied to the original CP (1) (see, e.g., Bazaraa et al. (1993). Since for any choice of $p \in \mathbb{R}^k_>$ and $e_j \leq 0$ for all $j = 1, \ldots, k$ the optimal objective value of (23) is an upper bound on the optimal objective value of CP, an *elastic constraints dual* of CP can be formulated as

$$\min_{p \in \mathbb{R}^k_>, e \in \mathbb{R}^k_{\leqq}} \ \max_{x \in S} \ z_0(x) - \sum_{j=1}^{k} p_j \cdot \max\{0, e_j - z_j(x)\}. \tag{24}$$

From a practical point of view, this dual formulation for CP may be useful since it has no dualtiy gap while giving more flexibility to the right-hand sides through the parameters $e_j, j \in J_i$, leading to more flexible and thus easier relaxed subproblems of the form (23).

3.8 Comparison

Considering a CP (1) and its associated MOP (2), we have shown that the most common scalarization approaches applied to (2) yield constrained or unconstrained programming problems that are closely related to some of the classical relaxation approaches applied to (1). The corresponding reformulations of (2) and their interpretation in view of (1) are summarized in Table 1.

4 Approximation of the nondominated set applied to constrained programming

In order to iteratively approach an optimal—or most preferred—solution of a CP (1), nondominated solutions of the associated MOP (2) can be used. Iteratively improved piecewise linear approximations of the nondominated set of the associated MOP can be a powerful tool to simultaneously obtain an overview of the alternatives and an estimate for the optimal solution value of a CP in an efficient way.

For this purpose, a variety of different methods could be applied including, for example, genetic algorithms as described in the introduction. In this section, we focus on a particular method recently suggested by Klamroth et al. (2002) that generates

Table 1 Comparison of the most common scalarization approaches in multicriteria programming and their relation to constrained programming

	Objective	Constraints
e-constraint approach	$z_i(x) \in \{f(x), g_1(x), \ldots, g_k(x)\}$ (one objective or constraint selected)	$z_j(x) \geq e_j,\ j \in J_i$ (right-hand sides possibly modified)
Weighted-sums approach	$f(x)+\tilde{\lambda}^T g(x)$ (Lagrangian relaxation of CP)	relaxed
Weighted Chebyshev approach	$z_i(x) \in \{f(x), g_1(x), \ldots, g_k(x)\}$ (one objective or constraint selected)	$z_j(x) \geq z_j^* - \frac{\bar{\alpha}}{w_j},\ j \in J_i$ (right-hand sides possibly modified)
Augmented weighted Chebyshev approach	$z_i(x) + \rho \sum_{j=0}^{k} z_j(x)$ (one objective or constraint selected plus Lagrangian penalty term)	$z_j(x) \geq z_j^* - \frac{\bar{\alpha}}{w_j},\ j \in J_i$ (right-hand sides possibly modified)
Reference point approach	$s(z(x)-z^0)$ (penalty function approach)	relaxed
Direction method	$z_i(x) \in \{f(x), g_1(x), \ldots, g_k(x)\}$ (one objective or constraint selected)	$z_j(x) \geq z_j^0 + \bar{\alpha} d_j + \bar{q}_j,\ j \in J_i$ (right-hand sides possibly modified)
Benson's method	$f(x) + \sum_{i=1}^{k} g_i(x)$ (specific Lagrangian relaxation of CP)	$z_i(x) \geq z_i^0,\ i \in \{0, \ldots, k\}$ (right-hand sides modified and not relaxed)
Method of elastic constraints	$f(x) - \sum_{i=1}^{k} p_i \max\{0, e_i - g_i(x)\}$ (exact absolute value penalty function for CP)	relaxed

such approximations in a deterministic, yet problem-dependent way by utilizing polyhedral distance functions to construct the approximation and evaluate its quality. The functions automatically adapt to the problem structure and scaling which makes the approximation process unbiased and self-driven.

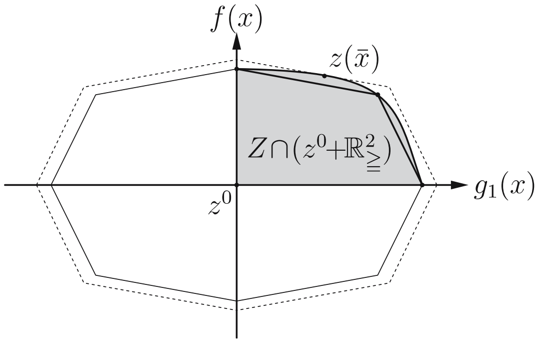
For the sake of completeness, the basic definitions and results of Klamroth et al. (2002) are briefly reviewed in Sects. 4.1 (for the convex case) and 4.5 (for the nonconvex case). They are supplemented by new and extended convergence results (Sects. 4.2 and 4.5) which have implications also for Langrangian relaxation methods in constrained programming (Sect. 4.4), and they are discussed with respect to their applicability for the solution of constrained programming problems in general (Sects. 4.3 and 4.5).

4.1 Global approximation for convex problems

Let Z be $\mathbb{R}^{k+1}_{\leqq}$ -convex (i.e., $Z + \mathbb{R}^{k+1}_{\leqq}$ is convex) with $\operatorname{int} Z \neq \emptyset$, and suppose that a reference point z^0 is given that satisfies $N \subseteq z^0 + \mathbb{R}^{k+1}_{\geqq}$. For a polyhedral gauge $\gamma : \mathbb{R}^{k+1} \to \mathbb{R}$, consider the problem

$$\begin{aligned} \max\ & \gamma(z(x) - z^0) \\ \text{s.t.}\ & z(x) \in z^0 + \mathbb{R}^{k+1}_{\geqq} \\ & x \in S. \end{aligned} \tag{25}$$

Fig. 1 Illustration of a simple oblique norm applied in problem (25) to an MOP (2) with the two objectives $z_0(x) = f(x)$ and $z_1(x) = g_1(x)$



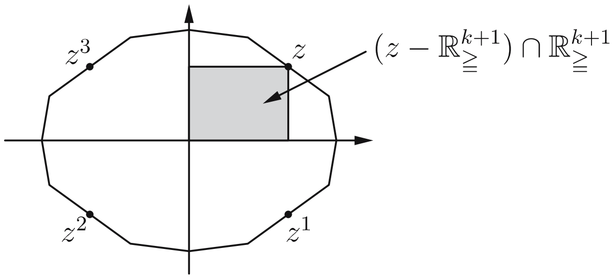


Fig. 2 Example of the unit ball of an oblique norm

If γ is defined by a symmetrical unit ball B centered at the origin that is obtained by symmetrically extending a given piecewise linear inner approximation of N in the objective space (see Fig. 1 for an example) (25) finds a feasible point $z(\bar{x}) \in Z$ that maximizes the problem-dependent γ-distance from the current approximation in the objective space. Schandl et al. (2002a) showed that in this case γ is an oblique norm, i.e., it is absolute ($\gamma(w) = \gamma(u) \quad \forall w \in R(u) := \{w \in \mathbb{R}^{k+1} : |w_i| = |u_i| \, \forall i = 0, \dots, k\}$) and it satisfies $(z - \mathbb{R}^{k+1}_{\geqq}) \cap \mathbb{R}^{k+1}_{\geqq} \cap \partial B = \{z\} \, \forall z \in (\partial B)_{\geqq}$. An example of an oblique norm in $\mathbb{R}^2$ is given in Fig. 2; $R(z) = \{z, z^1, z^2, z^3\}$ in this example.

Theorem 15 (Schandl et al. 2002b)

(1) *If γ is an oblique norm then every optimal solution $\bar{x}$ of* (25) *is an efficient solution of the MOP* (2).
(2) *If Z is $\mathbb{R}^{k+1}_{\leqq}$ -convex and if $\bar{x}$ is a properly efficient solution of* (2)*, then there exists an oblique norm γ such that $\bar{x}$ solves* (25).

Formulation (25) combines ideas from several classical scalarization methods, among them the weighted-sums approach (Sect. 3.2), the augmented weighted Chebyshev approach (Sect. 3.3) and the reference point approach (Sect. 3.4). It is, however, more general than the weighted-sums approach (cf. reformulation (28) below) and the augmented weighted Chebyshev approach (since it allows for a larger class of distance measures and applies them, intuitively speaking, from the opposite side). Moreover, it differs from the reference point approach by keeping the reference point fixed and requiring its selection such that $N \subseteq z^0 + \mathbb{R}^{k+1}_{\geqq}$, while varying the scalarizing function (which in fact may not be order-representing nor order-approximating, cf. Sect. 3.4).

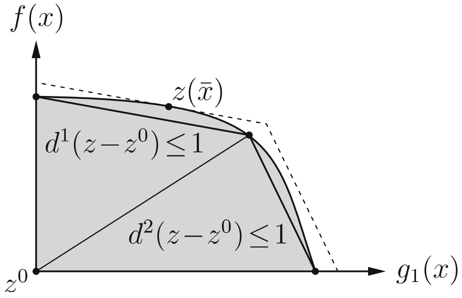


Fig. 3 Illustration of problem (26) for an MOP (2) with the two objectives $z_0(x) = f(x)$ and $z_1(x) = g_1(x)$

Let $d^1, \ldots, d^s \in \mathbb{R}^{k+1}$ be the normal vectors of the facets of the unit ball B of a polyhedral gauge γ such that $\{z \geqq 0 : d^j z \leq 1, j = 1, \ldots, s\} = B \cap \mathbb{R}^{k+1}_{\geqq}$ and $\left\{z \geqq 0 \ : \ d^j(z - z^0) \leq 1, \ j = 1, \ldots, s\right\} \subseteq Z_{\leqq} = Z + \mathbb{R}^{k+1}_{\leqq}$. Then problem (25) can be formulated as the following disjunctive programming problem:

$$\begin{aligned} &\max \ \lambda \\ &\text{s.t.} \ \ \textstyle\bigvee_{j=1}^{s} \left(d^j(z(x) - z^0) \geq \lambda \wedge x \in S\right). \\ &\qquad \lambda \in \mathbb{R}. \end{aligned} \tag{26}$$

Figure 3 shows an example with two facets represented by the normal vectors d^1 and d^2. The point $z(\bar{x})$ corresponds to an optimal λ in (26).

Problem (26) may be reconsidered in two directions via alternative formulations. One formulation brings the program (26) into a more conventional setting by application of the linearizing technique of Balas (1979). When (26) is used in the approximation procedure the result of (26) will always be positive. Hence we may assume that λ is positive. An equivalent problem, although with reciprocal value, is then

$$\begin{aligned} &\min \ \tfrac{1}{\lambda} \\ &\text{s.t.} \ \ \textstyle\bigvee_{j=1}^{s} \tfrac{1}{\lambda} d^j(z(x) - z^0) \geq 1, \\ &\qquad x \in S, \lambda > 0. \end{aligned}$$

Substituting $\frac{1}{\lambda}$ by μ gives the problem

$$\begin{aligned} &\min \ \mu \\ &\text{s.t.} \ \ \textstyle\bigvee_{j=1}^{s} \mu d^j(z(x) - z^0) \geq 1, \\ &\qquad x \in S, \mu \geq 0. \end{aligned}$$

Let $\mu_j \in \mathbb{R}$ for $j = 1, \ldots, s$. An equivalent formulation is

$$\begin{aligned} &\min \ \sum_{j=1}^{s} \mu_j \\ &\text{s.t.} \ \ \mu_j d^j(z(x) - z^0) \geq y_j, \qquad \forall j \in \{1, \ldots, s\}, \\ &\qquad \sum_{j=1}^{s} y_j = 1, \\ &\qquad y_j \geq 0, \mu_j \geq 0, \qquad\qquad \forall j \in \{1, \ldots, s\}, \\ &\qquad x \in S. \end{aligned} \tag{27}$$

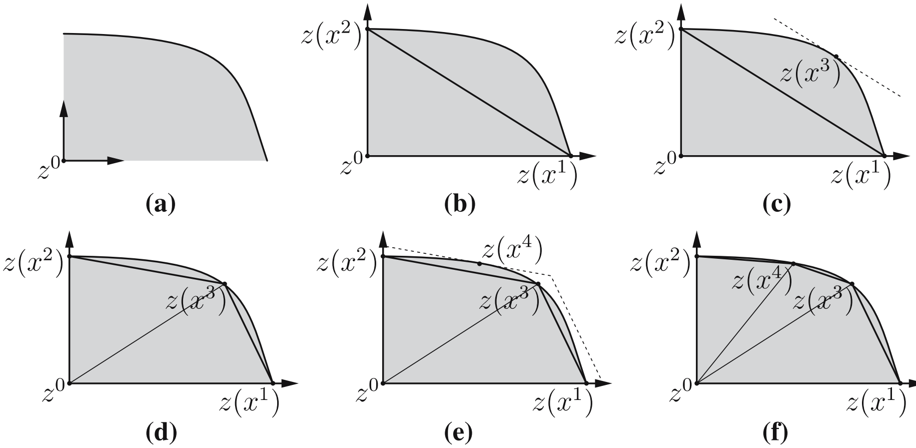


Fig. 4 Inner approximation algorithm

Problem (26) has in this way been put into a conventional compound form, which may be best suitable for direct application of a standard optimization routine.

The other formulation decomposes problem (26) into multiple subproblems, each of a particularly simple structure. For this purpose, let B be the unit ball of γ and denote by $C_1, \ldots, C_s$ the fundamental cones of $B \cap \mathbb{R}^{k+1}_{\geqq}$. If d^j is the normal vector of the facet of the cone $C_j, j = 1, \ldots, s$, then (25) can be decomposed into s subproblems $(P^j), j = 1, \ldots, s$, of the form

$$\begin{aligned} \lambda_j &= \max\ d^j(z(x) - z^0) = \sum_{i=0}^{k} d_i^j\,(z_i(x) - z_i^0) \\ &\text{s.t.}\quad x \in S \end{aligned} \tag{28}$$

from which the maximum value of λ_j, $j = 1, \ldots, s$ must be selected to obtain an overall optimal solution of (26). Note that each subproblem (28) corresponds to a weighted-sums scalarization of the MOP (2) (cf. Sect. 3.2) and thus has a Lagrangian type objective function and contains only the problem dependent constraints $x \in S$. If the given problem (2) is the associated MOP of a CP (1), we can generally assume that all the complicating constraints of (1) are already contained in the objectives of (2) and that the feasible set S has a very simple structure, making (28) easily solvable.

The approximation algorithm suggested in Klamroth *et al.* (2002) now iteratively solves problem (25) by computing optimal solutions of all newly generated subproblems of the form (28), starting from an initial approximation that can be generated, for example, by solving $\min\{z_i(x)\colon\ x \in S\}$ for all $i = 0, \ldots, k$.

In each iteration, the point of "worst" approximation is added to the current approximation which leads to an adaptive update of the polyhedral gauge γ and thus to the generation of a new set of "active" subproblems (28) in the updated cones. Figure 4 shows the procedure at the example of the inner approximation for a convex problem. Outer approximations can be constructed in a similar way.

Note that this approach is related to several other approximation approaches for convex MOPs, most of all probably to the *NISE method* (Cohon 1978) and to the *estimate refinement method* (Lotov 1985, Lotov et al. 2004). The main difference lies

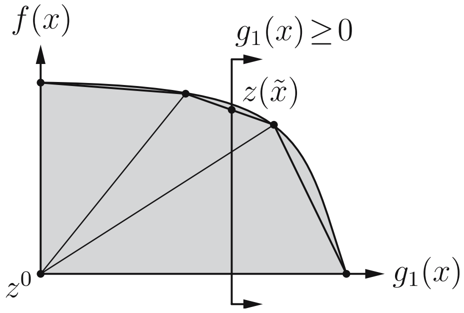


Fig. 5 Inner approximation algorithm applied to a CP (1) with one constraint $g_1(x) \geq 0$

in the fact that the quality of the current approximation is measured based on the approximation itself and not, for example, by the Hausdorff distance. Consequently, different approximation results are obtained since different points may be selected for the addition to the approximation, even if the individual subproblems solved are of the same or of a similar type as the problems (28) above. The main advantage of using a problem-dependent distance measure can be seen in the achieved scale-independence which is particularly important if the MOP is based on an underlying CP where objective function and constraints may model completely different things.

Figure 5 shows the point $z(\tilde{x})$ obtained from the intersection of the piecewise linear approximation of the nondominated set of (MOP) with the constraint set $g_i(x) \geq 0$, $i = 1, \ldots, k$, for an example with $k = 1$.

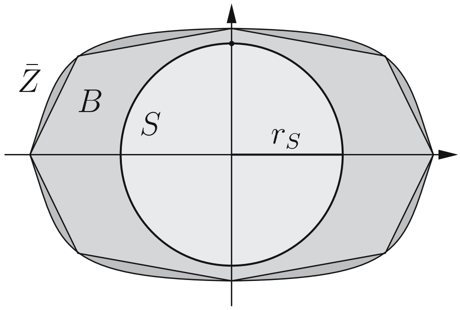
Note that in the bicriteria case each iteration of the approximation algorithm involves the solution of only one weighted-sums scalarization (28) of (2). Particularly in this case, this approach thus yields an efficient method to approximate the optimal Lagrange multipliers as well as the optimal solution of constrained programming problems (see also Sects. 4.2–4.4). Even though the number of active subproblems per iteration may theoretically be larger for $k \geq 2$, it can be expected not to have a considerable impact on the average time needed to find the next iterate. Recall also that, as an alternative to the decomposition of problem (26) into subproblems (28), one compound problem (27) could be solved in each iteration of the approximation method.

4.2 Convergence rate for convex problems

Based on a result of Rote (1992) on the convergence rate of sandwich approximations of convex functions, Klamroth et al. (2002) showed that for convex bicriteria problems the approximation error after m iterations of the approximation algorithm described in Sect. 4.1 measured by the approximating gauge γ, decreases by the order of $O(\frac{1}{m^2})$ which is optimal. Using similar relations between the gauge distances γ and the classical Hausdorff distance as in Klamroth et al. (2002), combined with more general results of Kamenev (1992, 1994) and Lotov et al. (2004), we will show in the following that a generalization is possible also for $(k+1)$-criteria problems, $k+1 \geq 2$, yielding a convergence rate of $O(\frac{1}{m^{2/k}})$ in this case.

For this purpose, suppose that the unit ball B of the approximating gauge γ of a problem in $\mathbb{R}^{k+1}$ after m iterations of the approximation algorithm is given by the reflection set of $B \cap \mathbb{R}^{k+1}_{\geqq}$, i.e., $B = R(B \cap \mathbb{R}^{k+1}_{\geqq})$. Moreover, let $\bar{Z}$ be the reflection

Fig. 6 Comparing the problem dependent gauge distance to the Hausdorff distance



set of $(Z - z^0)_{\leqq} \cap \mathbb{R}^{k+1}_{\geqq}$, i.e., $\bar{Z} = R((Z - z^0)_{\leqq} \cap \mathbb{R}^{k+1}_{\geqq})$ is the set obtained from $Z_{\leqq} = Z + \mathbb{R}^{k+1}_{\leqq}$ by moving its reference point z^0 to the origin and extending the resulting set from $\mathbb{R}^{k+1}_{\geqq}$ symmetrically around the origin. Then the Hausdorff distance $d_H(B, \bar{Z})$ between the compact convex set $\bar{Z}$ and its polyhedral approximation $B \subseteq \bar{Z}$ is given by

$$d_H(B, \bar{Z}) = \sup_{z \in \bar{Z}} \inf_{b \in B} \| z - b \|_2,$$

where $\| z - b \|_2$ denotes the Euclidean distance between the two points b and z in $\mathbb{R}^{k+1}$. Let r_S be the radius of a sphere S centered at the origin that is completely contained in B. If we denote the norm with unit ball S by $\| \bullet \|_S$, we have $\| u \|_2 = r_S \cdot \| u \|_S$ for all $u \in \mathbb{R}^{k+1}$. Moreover, $\| u \|_S \geq \gamma(u)$ for all $u \in \mathbb{R}^{k+1}$ since $S \subseteq B$ (see Fig. 6 for an example). Hence,

$$d_H(B, \bar{Z}) = r_S \cdot \sup_{z \in \bar{Z}} \inf_{b \in B} \| z - b \|_S \geq r_S \cdot \sup_{z \in \bar{Z}} \inf_{b \in B} \gamma(z - b) = r_S \cdot |\gamma(\bar{z} - z^0) - 1|\,,$$

where $\bar{z}$ is an optimal solution of (25). Observe that the above relations are true for the approximating gauge γ and unit ball B at every iteration of the approximation algorithm following the mth iteration.

Kamenev (1992, 1994) (see also Lotov et al. 2004) showed that if the estimate refinement method is applied to iteratively construct inscribed polyhedra P^m, $m \geq 1$, to approximate a compact convex set $\bar{Z}$ in $\mathbb{R}^{k+1}$, then for any $0 < \varepsilon < 1$ there exists $M \in \mathbb{N}$ such that for all $m \geq M$ the Hausdorff distance between the approximating polyhedron and the set $\bar{Z}$ after m iterations can be bounded by

$$d_H(P^m, \bar{Z}) \leq (1 + \varepsilon) \cdot 16 R_{\bar{Z}} \cdot \left(\frac{(k+1)\pi_{k+1}}{\pi_k} \right)^{2/k} \cdot m^{-2/k},$$

where π_k is the volume of the unit sphere in $\mathbb{R}^k$ and $R_{\bar{Z}}$ denotes the radius of a sphere circumscribed around $\bar{Z}$.

Exactly the same points that are also found by solving problem (25) are generated by the estimate refinement method: new points are found by maximizing the Euclidean distance between supporting hyperplanes of P^m and $\bar{Z}$, respectively, over all possible normal directions. Since P^m is a polyhedron, this maximum is attained at a supported point of $\bar{Z}$ solving (25) in some fundamental cone of P^m. These points are then added to the convex hull of the approximation in a possibly different order which may be nonoptimal with respect to γ, and hence we have

$$\sup_{z\in\bar{Z}} \inf_{b\in B} \gamma(z-b) \le \sup_{z\in\bar{Z}} \inf_{p\in P^m} \gamma(z-p).$$

Thus, the approximation error of the approximation algorithm described in Sect. 4.1 after m iterations, $m \ge M$, can be bounded by

$$\begin{aligned} |\gamma(\bar{z}-z^0)-1| &= \sup_{z\in\bar{Z}} \inf_{b\in B} \gamma(z-b) \le \sup_{z\in\bar{Z}} \inf_{p\in P^m} \gamma(z-p) \\ &\le \frac{1}{r_S}\, d_H(P^m,\bar{Z}) \le K m^{-2/k} \end{aligned} \tag{29}$$

for some constant K that can be chosen independently of m. (Note that the radius r_S of the sphere inscribed into B at iteration m can be chosen such that S is already inscribed into the initial approximation, i.e., for $m = 0$.)

Theorem 16 *Consider a CP* (1) *with k constraints and its associated MOP* (2) *with* $k+1$ *criteria,* $k \ge 1$. *Then the approximation error after m iterations of the approximation algorithm described in Sect. 4.1, measured by the adaptive polyhedral gauge* γ, *decreases by the order of* $O(m^{-2/k})$.

Proof Since the set $\bar{Z}$ is compact and convex by assumption, the convergence rate of $O(m^{-2/k})$ follows from Kamenev (1992, 1994) and (29). □

Note that this convergence rate is in general best possible since, for example, the unit sphere in $\mathbb{R}^{k+1}$ can be approximated by an inscribed polyhedron at most with an accuracy of this order (see, e.g., Gruber 1992).

4.3 Local approximation for convex problems

The method presented in Sect. 4.1 is designed to generate an approximation of the complete nondominated set N of the MOP (2). If the MOP associated with a given CP is considered, one may, however, be interested in an approximation of only that region of N that contains optimal solutions of the underlying CP (1). For the problem depicted in Fig. 5 this could be realized, for example, by refining the approximation only in that candidate cone that potentially contains an optimal solution $\bar{z} = z(\bar{x})$ of (1).

To avoid trivial situations we will assume in the following that at least one constraint is binding at any optimal solution of the given CP (1). Note that otherwise the problem could be solved by simply solving the relaxed problem $\max_{x\in S} f(x)$. (For the associated MOP (2) this would imply that the ideal point is a feasible solution.)

For CPs with a single constraint and the associated bicriteria MOPs, cones that potentially contain an optimal solution $\bar{z} = z(\bar{x})$ of the CP (1) can be easily identified by checking the constraint $g_1(x) \ge 0$ for the fundamental vectors (corresponding to the extreme points of B) that generate the respective cones (see Fig. 5). In fact, we can restrict our search to candidate cones that are generated by two extreme points $z(x)$ and $z(y)$, one of which satisfies the constraint, i.e., $g_1(x) \ge 0$ while the other violates the constraint, i.e., $g_1(y) < 0$.

For CPs with two or more constraints and the related multicriteria MOPs, the situation is, however, more complicated as the following example with two constraints shows: suppose that an approximating polyhedron $z^0 + B$ with reference point $z^0 = (f(x^0), g_1(x^0), g_2(x^0))^T = (0,0,0)^T$ is given by the extreme points $z^1 = (1.5,0,0)^T$,

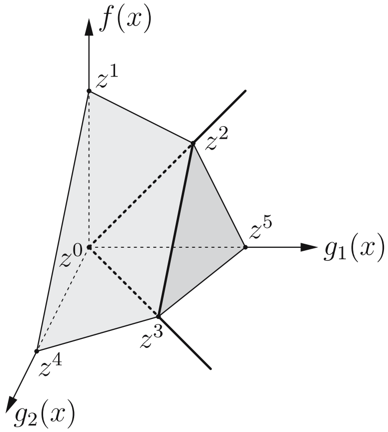


Fig. 7 The cone defined by $z^1 = z(x^1), \ldots, z^4 = z(x^4)$ may contain solutions with $g_1(x) > 1$ even though $g_1(x^i) \leq 1 \forall i = 1, \ldots, 4$

$z^2 = (1, 1, 0)^T$, $z^3 = (0, 1, 1)^T$, $z^4 = (0, 0, 1.5)^T$, and $z^5 = (0, 1.5, 0)^T$ (see Fig. 7 for an illustration).

Then the facet of the fundamental cone $z^0 + C$ spanned by $z^1, \ldots, z^4$ has the normal vector $d = (1, 0.5, 1) \geqq 0$. If now the bound on g_1 is given by $g_1(x) \geq 1 + \varepsilon$ where $\varepsilon > 0$ is a sufficiently small number, we can easily construct an example problem where the intersection of the feasible set Z with the face of the cone C defined by z^2 and z^3 (shown in bold lines in Fig. 7) contains feasible points $z(x)$ with $g_1(x) \geq 1 + \varepsilon$. However, all extreme points defining the considered cone violate the constraint $g_1(x) \geq 1 + \varepsilon$. Thus, even if all extreme points of a cone violate one of the constraints, this cone may still contain an optimal solution of the CP (1).

Unless other constraints on the location of an optimal solution of a CP with $k \geq 2$ are known, the approximation must therefore in general be refined in all fundamental cones of $z^0 + B$.

4.4 Application in Lagrangian relaxation

The convergence rate outlined is Sect. 4.2 has a useful application in the determination of the optimal Lagrange multipliers associated with the contraints of CP (1).

As well-known, Lagrange multipliers play a fundamental role in the analysis and solution of nonlinear, in particular convex optimization problems. An important class of solution methods fall into the category of subgradient optimization, where the selection of an efficient updating scheme of the Lagrange multipliers is important. Thus, the convergence of a subgradient procedure, theoretically and in practice, depends on the updating scheme. Also, in various decomposition procedures the updates of Lagrange multipliers are important. In addition to the classical Dantzig–Wolfe decomposition other schemes have been successful, for example methods based on interior point algorithms (see Elhedhli et al. 2001), or mixed schemes like the mean value methods (see Holmberg 1997). Some of the schemes have known convergence rates, others have not.

The approximation method presented here can be viewed as a specific updating procedure of the Lagrange multipliers. In each iteration the vector d^j of the fundamental cone containing the optimal solution is an approximation of the optimal Lagrange multipliers (upon normalization, in which the first component corresponding to the objective function is set to 1).

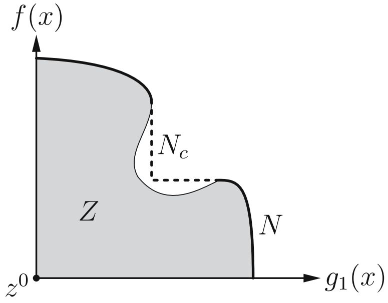


Fig. 8 Connected extension N_c of the components of N for a nonconvex example problem

Note, that this immediately gives us a procedure for the determination of the Lagrange multipliers with a prescribed (and in a worst case sense the best possible) convergence rate as given by Theorem 16.

4.5 Nonconvex problems

Consider a CP (1) and its associated MOP (2) with a nonconvex feasible set Z. We assume that $Z \subseteq \mathbb{R}^{k+1}$ is $\mathbb{R}^{k+1}_{\leqq}$ -closed with $\operatorname{int}(Z) \neq \emptyset$, and that a reference point $z^0 \in Z_{\leqq} = Z + \mathbb{R}^{k+1}_{\leqq}$ is given. Since the nondominated set N may be nonconnected in general, a piecewise linear approximation can only aim at approximating the set $N_c := \{z \in Z_{\leqq} \,:\, \nexists\, \tilde{z} \in Z_{\leqq} \text{ s.t. } \tilde{z} \geq z\}$ (see Fig. 8 for an example).

In order to extend the ideas from Sect. 4 to the nonconvex case, Klamroth et al. (2002) suggested to replace the convex unit ball of a distance measuring gauge γ by a nonconvex set B that is constructed from the intersection of dominating cones. B is then used as a unit set to define a distance measuring function γ as $\gamma(z) := \min\{\lambda \,:\, z \in \lambda B\}$.

The basic idea for an approximation procedure is—similar to the convex case—to minimize the maximum γ-distance between a nondominated point in Z and the boundary of B. However, since the weighted sums method is in general not suitable for nonconvex problems, in this case variants of the Chebyshev method (cf. Sect. 3.3) are used for the generation of candidate points which are iteratively added to the approximation.

Let $d^1, \ldots, d^s \in \mathbb{R}^{k+1}_{\geqq}$ be a nonempty and finite set of vectors spanning the nonnegative orthant, and let B be defined by $B = \operatorname{cl}(\mathbb{R}^{k+1}_{\geqq} \setminus \bigcup_{j=1,\ldots,s}(d^j + \mathbb{R}^{k+1}_{\geqq}))$. We assume that B is bounded, has nonempty interior and that $B \subseteq ((Z - z^0)_{\leqq} \cap \mathbb{R}^{k+1}_{\geqq})$. As in the convex case, B could be symmetrically extended to all orthants of the coordinate system, yielding a compact set $\hat{B}$ that contains the origin in its interior.

If we interpret the vectors $z^0 + d^1, \ldots, z^0 + d^s$ as local nadir points, they define a corresponding set of local ideal points $z^0 + v^1, \ldots, z^0 + v^s$ which can be computed by

$$z_i^0 + v_i^j = \max\left\{z_i\colon z_l = z_l^0 + d_l^j \ \forall l \neq i,\ l \in \{0, \ldots, k\};\ z \in Z_{\leqq}\right\}, \quad i = 0, \ldots, k.$$

Each pair $(z^0 + d^j, z^0 + v^j)$, $j = 1, \ldots, s$ in the initial approximation is assumed to generate a $(k+1)$-dimensional axis-parallel rectangular box which can be used to define the weights for a local application of the Chebyshev method (see Fig. 9 for an example). As candidates for the extension of the approximation we consider those points that

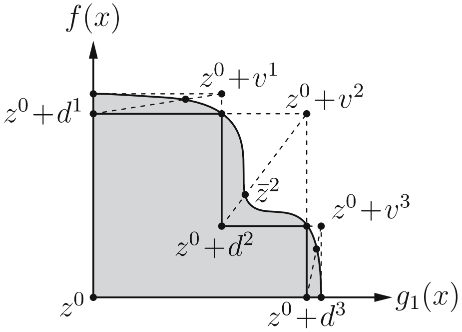


Fig. 9 Approximation in the nonconvex case. The optimal solution of (31) is attained at $\bar{z}^2$ where $(z^0+d^2)+\bar{\lambda}\cdot\frac{v^2-d^2}{\gamma(v^2)-\gamma(d^2)}=\bar{z}^2$

are generated by the corresponding local Chebyshev method in these boxes, and a resulting point $v\in N_c$ that is currently worst approximated with respect to the distance measure γ is added to the approximation. Consequently, v can be generated by solving the disjunctive programming problem

$$\begin{aligned}&\max\ \gamma(v-z^0)\\ &\text{s.t.}\ \ \textstyle\bigvee_{j=1}^{s}\left((z^0+d^j)+\lambda_j(v^j-d^j)=v\ \wedge\ \lambda_j\geq 0\ \wedge\ v\leqq z(x)\ \wedge\ x\in S\right).\end{aligned}\tag{30}$$

Note that within a cone $(z^0+d^j)+\mathbb{R}^{k+1}_{\geqq}$ with $d^j<v^j$, $j\in\{1,\ldots,s\}$, solving (30) is equivalent to an application of the weighted Chebyshev method (8) with the ideal point z^0+v^j and with the weights $w_i^j:=\frac{1}{v_i^j-d_i^j}$, $i=0,\ldots,k$. Moreover, problem (30) can be simplified to

$$\begin{aligned}&\max\ \lambda\\ &\text{s.t.}\ \bigvee_{j=1}^{s}\left((z^0+d^j)+\lambda\cdot\frac{v^j-d^j}{\gamma(v^j)-\gamma(d^j)}\leqq z(x)\ \wedge\ x\in S\right),\\ &\qquad \lambda\geq 0.\end{aligned}\tag{31}$$

In this formulation, the distance information between the current approximation (given by B) and a point $(z^0+d^j)+\lambda\cdot\frac{v^j-d^j}{\gamma(v^j)-\gamma(d^j)}$ is captured in the value of λ. In particular, the optimal solutions $\bar{v}$ of (30) and $\bar{\lambda}$ of (31) satisfy $\gamma(\bar{v}-z^0)=1+\bar{\lambda}$ (see Klamroth *et al.* (2002)). Figure 9 shows problem (31) and its optimal solution.

Theorem 17 (Klamroth et al. 2002) *Let $\bar{\lambda}$ be an optimal solution of* (31), *let J be the index set of all constraints of* (31) *that are satisfied and binding at optimality, and let $\bar{v}^j:=d^j+\bar{\lambda}\cdot\frac{v^j-d^j}{\gamma(v^j)-\gamma(d^j)}$, $j\in J$.*

(1) *The set B' defined as*

$$B':=\left(B\cup\bigcup_{j\in J}(\bar{v}^j-\mathbb{R}^{k+1}_{\geqq})\right)\cap\mathbb{R}^{k+1}_{\geqq}$$

satisfies $B\subseteq B'\subseteq((Z-z^0)_{\leqq}\cap\mathbb{R}^{k+1}_{\geqq})$, i.e., the set B' yields again an inner approximation of the set N_c.

(2) *If $z^0+\bar{v}^j\leqq\bar{z}^j\in Z$, then $\bar{z}^j$ is weakly nondominated. Moreover, if Z is strictly $\mathrm{int}\mathbb{R}^{k+1}_{\geqq}$-convex, then the solution $\bar{z}^j$ is properly nondominated.*

In Sect. 4.1 the disjunctive problem (26) was converted into the compound form (27). Similarly (31) can be converted into a compound form. Let $\lambda_j \in \mathbb{R}$ for $j = 1 \dots s$ and consider

$$\begin{aligned} \max \ & \sum_{j=1}^{s} \lambda_j \\ \text{s.t.} \ & \lambda_j \cdot \frac{v^j - d^j}{\gamma(v^j) - \gamma(d^j)} \leqq y_j(z(x) - (z^0 + d^j)), \quad \forall j \in \{1, \dots, s\}, \\ & \sum_{j=1}^{s} y_j = 1, \\ & y_j \geq 0, \lambda_j \geq 0, \quad \forall j \in \{1, \dots, s\}, \\ & x \in S. \end{aligned} \tag{32}$$

From an algorithmic point of view we would like to iteratively solve problem (31) and include the resulting solution into the approximation as indicated in Theorem 17(1). Note that, if for some $j \in J$ an optimal solution $\bar{v}^j$ of problem (31) satisfies $\bar{v}^j = d^j$, the approximation in the corresponding box is tight and it can be discarded in further iterations. Otherwise, i.e., if $\bar{v}^j > d^j$ (and $\bar{\lambda} > 0$), the point $\bar{v}^j$ lies on the diagonal of the jth local Chebyshev box, and the $k+1$ sub-boxes obtained from computing $B' = (B \cup (\bar{v}^j - \mathbb{R}^{k+1}_{\geqq})) \cap \mathbb{R}^{k+1}_{\geqq}$ are again $(k+1)$-dimensional. Using similar ideas as in Hamacher et al. (to appear), who suggested a similar approximation algorithm for bicriteria discrete problems based on repetitive applications of the e-constraint approach (see Sect. 3.1) rather than local applications of the weighted Chebyshev method, the sum of the volumes V' of these sub-boxes can be computed as

$$\begin{aligned} V' &= V^j - \prod_{i=0}^{k} \bar{\lambda}(v_i^j - d_i^j) - \prod_{i=0}^{k} (1 - \bar{\lambda})(v_i^j - d_i^j) \\ &= V^j - \bar{\lambda}^{k+1} V^j - (1 - \bar{\lambda})^{k+1} V^j, \end{aligned}$$

where $V^j = \prod_{i=0}^{k} (v_i^j - d_i^j)$ denotes the volume of the original box before its partition (see Fig. 10b and c for an example where $\bar{v}^1 = \bar{z}^1$). Since V' is maximal if $\bar{\lambda} = \frac{1}{2}$, we obtain the following bound on the rate at which the volume of the local Chebyshev boxes used for the approximation decreases:

Theorem 18 *Consider a nonconvex CP* (1) *with k constraints and its associated MOP* (2) *with* $k+1$ *criteria,* $k \geq 1$. *In each iteration of the approximation algorithm for nonconvex problems described in this section, the volume V of at least one local Chebyshev box is decreased at least by an amount of* $(\frac{1}{2})^k V$.

As an alternative to a repeated solution of (31) or of its reformulation (32), the problem can be decomposed into subproblems (P^j), $j = 1, \dots, s$, that resemble individual applications of the direction method (17) and that can be formulated as

$$\begin{aligned} \delta_j = \max \ & \lambda \\ \text{s.t.} \ & (z^0 + d^j) + \lambda \cdot \frac{v^j - d^j}{\gamma(v^j) - \gamma(d^j)} \leqq z(x) \\ & \lambda \in \mathbb{R}, \ x \in S. \end{aligned} \tag{33}$$

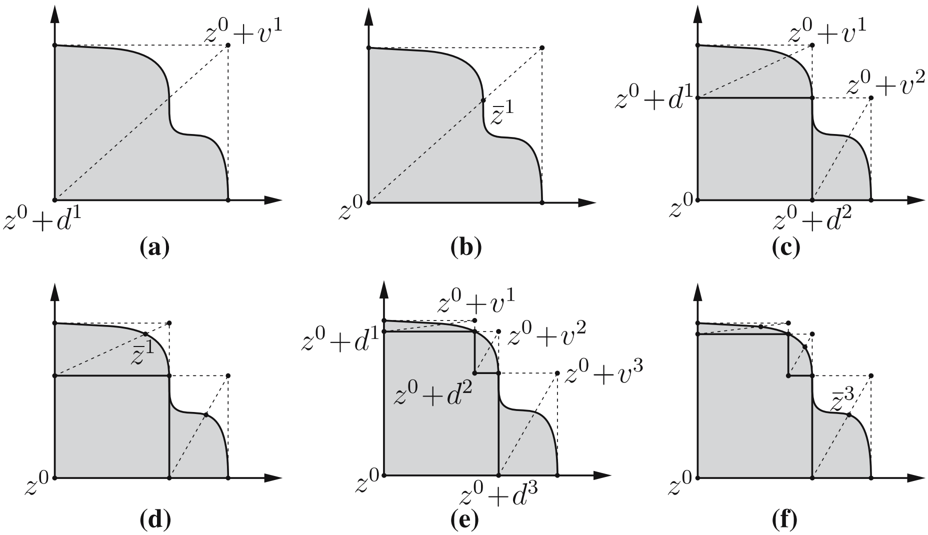


Fig. 10 Inner approximation algorithm for general nonconvex problems: **a** and **b** generation of an initial approximation and **c** – **f** iterations

The optimal $\bar{\lambda}$ of (33) determines the vector

$$\bar{v} = (z^0+d^j) + \bar{\lambda} \cdot \frac{v^j - d^j}{\gamma(v^j) - \gamma(d^j)}.$$

Then the optimal solution value of (31) equals the maximum value of δ_j, $j = 1,\dots,s$ which also yields the related vector $\bar{v}$ as indicated above. The resulting approximation algorithm for nonconvex problems is shown in Fig. 10.

Note that, as in the convex case, each update of the set B generates a new set of active subproblems (33) in the modified local Chebyshev boxes, while the subproblems remain unchanged in the remaining boxes.

Theorem 19 *Consider a nonconvex CP* (1) *with k constraints and its associated MOP* (2) *with* $k+1$ *criteria,* $k \geq 1$. *Then the approximation error of the approximation algorithm outlined in this section, measured by the adaptive distance measuring function* γ, *converges to zero.*

Proof According to Theorems 17 and 18, the adaptively generated unit sets B_m, $m \geq 1$, of the distance measuring function γ_m satisfy $B_m \subseteq B_{m+1}$ and $B_m \subseteq ((Z - z^0)_{\leqq} \cap \mathbb{R}^{k+1}_{\geqq})$ for all $m \geq 1$. Consequently, the sequence of approximating sets $(B_m)_{m\geq 1}$ is convergent to some limiting set $\bar{B}$ which satisfies $\bar{B} \subseteq ((Z - z^0)_{\leqq} \cap \mathbb{R}^{k+1}_{\geqq})$, i.e., $z^0 + \bar{B}$ is an inner approximation of the set $Z_{\leqq} \cap \mathbb{R}^{k+1}_{\geqq}$. Moreover, for all stages of the approximation, all parts of the set $N_c \subseteq \partial(Z_{\leqq})$ (which is to be approximated) are contained in the union of the local Chebyshev boxes given by the pairs of local nadir and local ideal points $(z^0 + d^{j,m}, z^0 + v^{j,m})$, $j = 1,\dots,s_m$, induced by the approximation B_m in iteration m, $m \geq 1$.

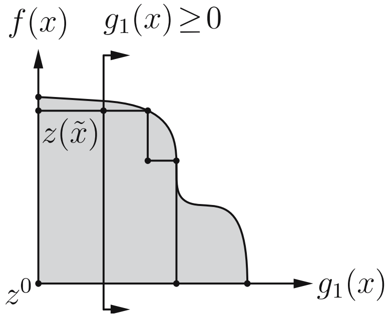


Fig. 11 Approximation of an optimal solution of the CP (1) for $k = 1$

Similarly, $(\bar{\lambda}_m)_{m \geq 1}$ is a sequence with $0 \leq \bar{\lambda}_m \leq \max_{j=1,\ldots,s_m}\{\gamma_m(v^{j,m}) - \gamma_m(d^{j,m})\} \leq \max_{j=1,\ldots,s_1}\{\gamma_1(v^{j,1}) - \gamma_1(d^{j,1})\} =: K$ for all $m \geq 1$. Thus, $(\bar{\lambda}_m)$ has a subsequence $(\bar{\lambda}_{m_l})$ that is convergent to some value $\bar{\lambda} \geq 0$.

The boundary of the limiting set $\bar{B}$ may consist of parts where it coincides with the set N_c (i.e., the volume $\prod_{i=0}^{k}(v_i^j - d_i^j)$ of the corresponding Chebyshev boxes has converged to zero and we thus have the desired convergence in these areas) and other parts where the corresponding Chebyshev boxes still satisfy $\prod_{i=0}^{k}(v_i^j - d_i^j) \geq \delta > 0$. Note that there may only be finitely many such boxes, and denote their local nadir and local ideal points by $(z^0 + d^j, z^0 + v^j)$, $j = 1, \ldots, \bar{s}$.

Now suppose that $\bar{\lambda} > 0$. Since $\lim_{l\to\infty} \bar{\lambda}_{m_l} = \bar{\lambda}$ and $\lim_{l\to\infty} B_{m_l} = \bar{B}$, we have that the optimal solution of

$$\begin{aligned} &\max\ \gamma_{\bar{B}}(v - z^0) \\ &\text{s.t.}\ \ \bigvee_{j=1}^{\bar{s}} \left((z^0 + d^j) + \lambda_j(v^j - d^j) = v\ \wedge\ \lambda_j \geq 0\ \wedge\ v \leqq z(x)\ \wedge\ x \in S\right) \end{aligned}$$

is $\gamma_{\bar{B}}(\bar{v} - z^0) = 1 + \bar{\lambda} > 1$, where $\gamma_{\bar{B}}(z) = \min\{\lambda\ :\ z \in \lambda\bar{B}\}$, cf. problem (30). Consequently, there exists a local Chebyshev box defined by (d^l, v^l), $l \in \{1, \ldots, \bar{s}\}$ such that $(z^0 + d^l) + \bar{\lambda}(v^l - d^l) = \bar{v}$. Since the volume $\prod_{i=0}^{k}(v_i^l - d_i^l)$ of the corresponding box is non-zero, we have that $v_i^l > d_i^l$ for all $i \in \{0, \ldots, k\}$, and thus $\bar{v}_i > (z^0 + d^l)_i$ for all $i \in \{0, \ldots, k\}$. But then the corresponding Chebyshev box should have been split at some stage of the approximation procedure, contradicting the fact that $\lim_{l\to\infty} B_{m_l} = \bar{B}$. We can conclude that $\lim_{l\to\infty} B_{m_l} = ((Z - z^0)_{\leqq} \cap \mathbb{R}^{k+1}_{\geqq}) = \lim_{m\to\infty} B_m$, and thus $\lim_{m\to\infty} \bar{\lambda}_m = 0$. □

Intersecting the final approximation with the constraint set $g_i(x) \geqq 0$, $i = 1, \ldots, k$, yields an approximation of the optimal solution of the original CP (1) the quality of which can be estimated through the value of γ, see Fig. 11 for an example with $k = 1$. In the light of Sects. 3.3, 3.5, and 4.4, an application of this approximation method to nonconvex CPs could be interpreted as an adaptive scheme for the update of partial problem relaxations (or bound modifications), or for the adaptation of search directions (line searches) in the objective space of the associated MOP.

Theorem 18 also suggests a simple, nonadaptive way of selecting the next subproblem (33) and the next local Chebyshev box that is split in the approximation procedure: If always a box with maximal volume among all active boxes is chosen, then the expected volume reduction according to Theorem 18 is maximized. Using again a similar analysis as Hamacher et al. (to appear), a bound on the convergence rate can be derived also for this method:

Table 2 Approximate values of $\frac{\ln(C)}{\ln(k+1)}$ for $k = 1, \ldots, 6$

k	1	2	3	4	5	6
$\frac{\ln(C)}{\ln(k+1)}$	−1	−0.2619	−0.0963	−0.0401	−0.0177	−0.0081

Starting with an initial approximation consisting of one single box (see Fig. 10 a) with volume V^0, after one iteration of the procedure the total volume of the at most $k+1$ resulting active boxes is reduced to an amount of at most $V^1 \leq V^0(1-(\frac{1}{2})^k)$. After $k+1$ further iterations, either each of these $k+1$ boxes has been split one more time, yielding an active volume of at most $V^{1+(k+1)} \leq V^1(1-(\frac{1}{2})^k)$, or an even better reduction could be achieved by selecting other boxes with larger volumes. In general, after $m_q = \sum_{r=0}^{q}(k+1)^r = \frac{(k+1)^{q+1}-1}{k}$ iterations, $q \geq 0$, we have

$$V^{m_q} \leq V^{m_{q-1}}\left(1-\left(\frac{1}{2}\right)^k\right) \leq V^0\left(1-\left(\frac{1}{2}\right)^k\right)^{q+1} = V^0\left(1-\left(\frac{1}{2}\right)^k\right)^{\frac{\ln(km_q+1)}{\ln(k+1)}}$$
$$= V^0 \cdot C^{\frac{\log_C(km_q+1)}{\log_C(k+1)}} = V^0 \cdot (km_q+1)^{\frac{1}{\log_C(k+1)}} = V^0 \cdot (km_q+1)^{\frac{\ln(C)}{\ln(k+1)}}$$

with $C := \frac{2^k-1}{2^k} < 1$. Thus, also for this selection rule, the sequence of approximating sets $(B_m)_{m\geq 1}$ converges to $((Z-z^0)_{\leqq} \cap \mathbb{R}^{k+1}_{\geqq})$ for all $k \geq 1$. Some values of $\frac{\ln(C)}{\ln(k+1)}$ for small k are given in Table 2. In particular, for bicriteria problems (i.e., $k=1$), we obtain a convergence rate of $O(\frac{1}{m})$.

5 Conclusions

Multicriteria optimization is most often considered as a generalization of single criterion optimization, and at least historically single criterion optimization has been the basis for the development of multicriteria optimization. However, with the vast increase over the last years of research in multicriteria optimization some feedback can be gained into single criterion optimization.

The purpose of the current paper is to point out the intimate link between some classical and nonclassical procedures in the two areas, even though they have been conceptually developed independently of each other.

Emphasis has been put on the comparison of some approximation schemes in multicriteria optimization and iterative solution methods in nonlinear, and in particular convex programming.

An approximation procedure for the analysis of multicriteria problems is shown to have intimate links to similar approximation methods for the determination of Lagrange multipliers in convex programming. A convergence result has been obtained for the suggested approximation scheme and an application of this result has been discussed and transferred into the determination of optimal Lagrange multipliers.

A discussion of the nonconvex case has been carried out as well and, as may be expected, with somewhat weaker results. However, approximation methods based on the Chebyshev approach have appeared useful in the multicriteria context, and more interesting results could be expected to be obtained in single criterion optimization by transfer of results obtained in the multicriteria setting.

Perhaps a final comment could be made here with respect to the class of optimization problems developed in data envelopment analysis, Cooper et al. (2000). This class of problems lie in some sense between single and multicriteria optimization. On one side the goal is to find, in a multicriteria sense, an efficient solution, and on the other side to optimize a certain real value (the so-called efficiency score). The problems are constructed in a technical, nonbiased way based on direct information about known data (for production and consumption). No preferences are built into the modelling. Moreover, in the present context it should be noted that the models themselves are polyhedral approximations.

Similarly, the approximation procedure presented here gives a polyhedral and nonbiased description. The procedure iteratively improves the determination of the set of feasible solutions and values without any prior knowledge about preferences. Some assumptions about convexity are usually done for the models used in data envelopment analysis. However, these assumption may not be fulfilled in practice. Further studies of nonconvex approximation procedures should open up for some interesting results in this and other areas.

References

Balas, E.: Disjunctive programming. Ann. Discrete Math. **5**, 3–51 (1979)

Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming—Theory and Algorithms, 2nd edn. Wiley, New York (1993)

Benson, H.P.: Existence of efficient solutions for vector maximization problems. J. Optim. Theory Appl. **26**, 569–580 (1978)

Boldur, G.: Linear programming problems with complex decision conditions. In: Proceeding of the 7th Mathematical Programming Symposium, The Hague, September 1970

Bowman V. J. Jr.: On the relationship of the Tchebycheff norm and the efficient frontier of multiple-criteria objectives. Lect. Notes Econ. Math. Syst. **130**, 76–86 (1976)

Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)

Camponogara, E., Talukdar, S.N.: A genetic algorithm for constrained and multiobjective optimization. In: Alander, J.T. (ed.) 3rd Nordic Workshop on Genetic Algorithms and their Applications, pp. 49–62. University of Vaasa, Finland (1997)

Carosi, L., Jahn, J., Martein, L.: On the connections between semidefinite optimization and vector optimization. J. Interdisciplinary Math. **6**, 219–229 (2003)

Chankong, V., Haimes, Y.Y.: Multiobjective Decision Making: Theory and Methodology. Elsevier Science Publishing, New York (1983)

Coello Coello, C.A.: Constraint-handling using an evolutionary multiobjective optimization technique. Civil Eng Environ Syst. **17**, 319–346 (2000)

Coello Coello, C.A., Mezura-Montes, E.: Handling constraints in genetic algorithms using dominance-based tournaments. In: Parmee, I. C. (ed.) Proceedings of the 5th International Conference on Adaptive Computing Design and Manufacture (ACDM 2002), pp. 273–284. Springer-Verlag, Berlin (2002)

Cohon, J.L.: Multiobjective Programming and Planning. Academic Press, New York (1978)

Cooper, W.W., Seiford, L.M., Tone, K.: Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software. Kluwer, Boston (2000)

Ehrgott, M.: A Discussion of Scalarization Techniques for Multiple Objective Integer Programming. Annals of Operations Research. (To appear 2006)

Ehrgott, M.: Multicriteria Optimization. Lecture Notes in Economics and Mathematical Systems. Springer-Verlag, Berlin (2000)

Ehrgott, M., Ryan, D.M.: The method of elastic constraints for multiobjective combinatorial optimization and its application in airline crew scheduling. In: Tanino, T., Tanaka, T., Inuiguchi, M. (eds.) Multi-Objective Programming and Goal Programming, pp. 117–122. Springer, Berlin (2003)

Elhedhli, S., Goffin, J.-L., Vial, J.-P.: Cutting plane methods for differentiable optimization. In: Pardalos, P., Floudas, C. (eds.) Encyclopedia of Optimization, vol. 4, pp. 40–45. Kluwer, Dordrecht (2001)

Fletcher, R., Leyffer, S.: Nonlinear programming without a penalty function. Math. Program. **91**, 239–269 (2002)

Gass, S., Saaty, T.: The computational algorithm for the parametric objective function. Naval Res. Logistics Quart. **2**, 39 (1955)

Geoffrion, A.M.: Proper efficiency and the theory of vector maximization. J. Math. Anal. Appl. **22**(3), 618–630 (1968)

Geoffrion, A.M., Dyer, J.S., Feinberg, A.: An interactive approach for multicriterion optimization, with application to the operation of an academic department. Manage. Sci. **19**, 357–368 (1972)

Gruber, P.M.: Approximation of convex bodies. In: Gruber, P.M., Wills, J.M. (eds.) Handbook of Convex Geometry, pp. 321–345. North-Holland, Amsterdam (1992)

Haimes, Y.Y., Lasdon, L.S., Wismer, D.A.: On a bicriterion formulation of the problems of integrated system identification and system optimization. IEEE Trans. Sys. Man Cybern. **1**, 296–297 (1971)

Hamacher, H.W., Pederson, C.R., Ruzika, S.: Finding Representative Systems for Discrete Bicriteria Optimization Problems by Box Algorithms. Operations Research Letters (to appear)

Holmberg, K.: Mean value cross decomposition applied to integer programming problems. Eur. J. Oper. Res. **97**, 124–138 (1997)

Jiménez, F., Gómez-Skarmeta, A.F., and Sánchez, G.: How evolutionary multi-objective optimization can be used for goals and priorities based optimization. In: Alba, E., Fernández, F., Gómez, J. A., Hidalgo, J. I., Lanchares, J., Merelo, J. J., Sánchez, J. M. (eds.) Primer Congreso Español de Algoritmos Evolutivos y Bioinspirados (AEB'02), pp. 460–465. Universidad de la Extremadura, España (2002)

Kamenev, G.K.: A class of adaptive algorithms for the approximation of convex bodies by polyhedra. Zh. Vychisl. Mat. Mat. Fiz., **32**(1), 136–152. In Russian; English translation in Comput. Maths. Math. Phys. **32**(1), 114–127 (1992)

Kamenev, G.K.: Investigation of an algorithm for the approximation of convex bodies. Zh. Vychisl. Mat. Mat. Fiz., **34**(4), 608–616. In Russian; English translation in Comput. Maths. Math. Phys. **34**(4), 521–528 (1994)

Klamroth, K., Tind, J., Wiecek, M.M.: Unbiased approximation in multicriteria optimization. Math. Meth. Oper. Res. **56**, 413–437 (2002)

Lin, J.: On min-norm and min-max methods of multi-objective optimization. Math. Program. **103**, 1–33 (2005)

Lotov, A.V.: Methods for Analysis of Mathematical Models of Controlled Systems on the Basis of Constructing the Set of Feasible Values for the Criteria of Control Quality. Dr.Hab. Thesis, Computing Center of the USSR Academy of Sciences, Moscow (1985) (In Russian)

Lotov, A.V., Bushenkov, V.A., Kamenev, G.K.: Interactive Decision Maps. Approximation and Visualization of Pareto Frontier. Kluwer, Dordrecht (2004)

Mezura-Montes, E., Coello Coello, C.A.: A numerical comparison of some multiobjective-based techniques to handle constraints in genetic algorithms. Technical Report EVOCINV-03-2002, Evolutionary Computation Group at CINVESTAV-IPN, Mexico, D.F. 07300 (2002)

Osyczka, A., Krenich, S., Tamura, H., Goldberg, D.: A bicriterion approach to constrained optimization problems using genetic algorithm. Evol. Optim. Int. J. Internet **2**, 41–53 (2000)

Pascoletti, A., Serafini, P.: Scalarizing vector optimization problems. J. Optim. Theory Appl. **42**, 499–524 (1984)

Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton, NJ (1970)

Rote, G.: The convergence rate of the sandwich algorithm for approximating convex functions. Computing, **48**, 337–361 (1992)

Roy, B.: Problems and methods with multiple objective functions. Math. Program. 1, 239–266 (1971)

Schandl, B., Klamroth, K., Wiecek, M.M.: Introducing oblique norms into multiple criteria programming. J. Global Optim. **23**, 81–97 (2002a)

Schandl, B., Klamroth, K., Wiecek, M.M.: Norm-based approximation in multicriteria programming. Comput. Math. Appl. **44**, 925–942 (2002b)

Steuer, R.E.: Multiple Criteria Optimization: Theory, Computation, and Application. Wiley, New York (1986)

Steuer, R.E., Choo, E.U.: An interactive weighted Tchebycheff procedure for multiple objective programming. Math. Program. **26**, 326–344 (1983)

Stewart, T.J.: A critical survey on the status of multiple criteria decision making, theory and practice. OMEGA Int. J. Manage. Sci. **20**, 569–586 (1992)

Surry, P.D., Radcliffe, N. J.: The COMOGA method: Constrained optimization be multiobjective genetic algorithms. Control Cybern. **26**, 391–412 (1997)
Wierzbicki, A.P.: Basic properties of scalarizing functionals for multiobjective optimization. Math. Operationsforschung und Statistik, Ser. Optim. **8**, 55–60 (1977)
Wierzbicki, A.P.: The use of reference objectives in multiobjective optimization. In: Fandel, G., Gal, T. (eds.) Multiple Criteria Decision Making, Theory and Applications. Lecture Notes in Economics and Mathematical Systems vol. 177, pp. 468–486. Springer, Berlin (1980)
Wierzbicki, A.P.: On the completeness and constructiveness of parametric characterizations to vector optimization problems. OR Spektrum **8**, 73–87 (1986)
Wierzbicki, A.P., Makowski, M., Wessels, J. (eds.): Model-Based Decision Support Methodology with Environmental Applications. Kluwer, Dordrecht (2000)